Genetic epidemiological approaches in complex neurological disorders

Jouke- Jan Hottenga

Genetic epidemiological approaches in complex neurological disorders

Jouke- Jan Hottenga

Jouke- Jan Hottenga

Genetic epidemiological approaches in complex neurological disorders
Thesis, University of Leiden, November 10th, 2005

ISBN-10: 9090197974
ISBN-13: 9789090197975

© 2005 Jouke- Jan Hottenga except (parts of) the following chapters:
Chapter 2: Bohn Stafleu van Loghum Bv
Chapter 4, 8: Blackwell Publishing
Chapter 7: Springer-Verlag GmbH

No part of this thesis may be reproduced in any form, by print, photocopy, digital file, internet, or any means without written permission of the copyright owner.

Printed by: Printgraph, Brzesko, Poland
Cover design: Zofia Kazimiera Mordarska

Genetic epidemiological approaches in complex neurological disorders

Proefschrift

Ter verkrijging van de graad van Doctor aan de Universiteit van Leiden, op gezag van Rector Magnificus Dr. D. D. Breimer, hoogleraar in de falculteit der Wiskunde en Natuurwetenschappen en die der Geneeskunde, volgens besluit van het college voor promoties te verdedigen op 10 november 2005 klokke 15.15 uur.

door

Jouke- Jan Hottenga

Geboren te Leiden in 1974

Promotiecommissie

Promotores	Prof.dr. R.R. Frants
	Prof.dr. C.M. van Duijn (Erasmus Universiteit Rotterdam)
Co promotor	Dr. A.M.J.M. van den Maagdenberg
Referent	Prof.dr. P.E. Slagboom
Overige Leden	Prof.dr. D.I. Boomsma (Vrije Universiteit Amsterdam)
	Prof.dr. M.H. Breuning
	Prof.dr. J.C. van Houwelingen

The studies described in this thesis were performed at the department of Epidemiology & Biostatistics of the Erasmus University Medical Center in Rotterdam and at the departments of Neurology and Human Genetics of the Leiden University Medical Center in Leiden, the Netherlands. This work was supported by the Netherlands Organization for Scientific Research (NWO).

Financial support for the printing of this thesis has been generously provided by Nederlandse Hoofdpijn Vereniging, Pfizer and Furore.

To Mirosława and Edward-Jan

In memory of Lodewijk A. Sandkuijl

Contents

CHAPTER 1

Introduction

JJ Hottenga

Complex neurological disorders

The prevalence of various complex neurological disorders, like migraine and Alzheimer's disease, is high in the general population[1-4]. Although complex neurological disorders are different in pathology and clinical manifestation, the impact on the quality of life of patients and the socio-economic level of the population is undoubtedly substantial[5-8]. The quality of life is reduced by the, often progressive, nature of the disorders and the lack of adequate treatment. Society and economy are burdened by costs of treatment, hospitalizations and loss of active working days of affected people. Findings that may help to reduce this impact are therefore of high importance. However, for most neurological disorders the pathophysiology, biochemical pathways and causative factors are complex and still largely unknown.

Neurological disorders are complex in various ways. A simple limitation is that brain tissue is difficult to study and research questions often have to be answered by other study designs. Another more important complexity is the often multifactorial nature of these disorders[9]. Multiple risk factors, environmental as well as genetic, contribute to the disorder individually or by means of interaction. Each independent risk factor increases the susceptibility, but not all risk factors are required to cause the trait. The multifactorial aspect also applies to genetic risk factors. As a result, the disorders cluster in families, but contrary to Mendelian monogenic disorders, there is often no clear mode of inheritance[10].

Gene identification in complex neurological disorders

In many complex neurological disorders a substantial part of the aetiology can be ascribed to genetic factors. In migraine, epilepsy and Alzheimer for example, the estimated heritability or proportion of variance explained by genetic factors, is ~46%, ~70% and ~48%, respectively[11-13]. Identification of these genetic factors is frequently an initial key step in understanding the

pathophysiology. Positional cloning is an often used method to identify genes. It involves essentially two steps, namely the identification of the region on the genome involved in the disorder (locus mapping through genome scans), followed by identification of the causative gene. In a genome scan, a narrow grid of markers evenly spaced over the genome is tested. For this purpose highly polymorphic microsatellite - repeat markers are used that have between two and thirty repeats (alleles), each consisting of two to six nucleotides. The marker alleles are subsequently correlated with the segregation of the disorder, leading to the identification of the genomic region(s) harboring the disease gene(s). Next, candidate genes in these regions are prioritized for further analysis. For Mendelian, monogenic, disorders, candidate genes are analyzed (for instance by sequencing) to identify high-impact mutations (missense, non-sense, deletions, insertions etc.). In the case of complex traits, one has to identify low-impact variants (polymorphisms). To this end, a denser grid of single nucleotide polymorphism (SNP) markers (bi-allelic) can be tested by association studies, followed by functional validation such as analysis of changed expression of the causative gene in affected individuals.

An alternative is the candidate gene approach; directly selecting candidate genes without prior genome scan experiments. The selection of a candidate gene is based on pathological, biochemical or molecular knowledge of the disorder. The candidate gene approach thus provides an opportunity to quickly assess the involvement of genes. This is useful to exclude known genes or to confirm / replicate findings of other studies. Nowadays, candidate genes can also come from for instance transcriptomics and proteomics studies.

In this thesis the main focus will be on the use of techniques involved in positional cloning. In recent years, the use of positional cloning has exponentially increased the number of genes known to be involved in human monogenic diseases[14]. For complex genetic traits including many neurological disorders, the successes have been more limited.

Problems in gene identification of complex traits

Trait definition

In neurological disorders there is often a lack of biological markers and diagnosis is based mainly on the presence of clinical symptoms. Although international diagnostic criteria for many disorders have been established, problems remain with their implementation in genetic studies[15-18]. There can be large variation in the expression of a disorder in patients of a family, making the inclusion or exclusion of these individuals as being affected for the study difficult. Diagnostic criteria such as 'severity' can be interpreted differently by patients and physicians. There can also be heterogeneity when patients have different subsets and/or frequency of clinical symptoms. For example, the presence and frequency of vomiting and phonophobia in migraine patients can be different[16]. Additional variation in phenotype can be caused by co-morbidity and clinical overlap of symptoms. In Alzheimer's disease for example, there is a large overlap with other dementias like vascular dementia and Parkinson[19,20]. Patients with epilepsy can sometimes be characterized with more than one syndrome. Therefore, the definition of neurological traits as phenotypes to be analyzed in genetic studies is in many cases not optimal.

Genetic Heterogeneity

Genetic heterogeneity is a major reason why neurological disorders are complex[9,10,14]. In linkage analysis genetic heterogeneity is often categorized in allelic - and locus heterogeneity. Allelic heterogeneity refers to the situation that multiple alleles of a single gene are related to an increased risk of the disorder, whereas locus heterogeneity refers to multiple genes involved in the disorder. Genetic heterogeneity may obscure the mode of inheritance, when autosomal recessive (2 risk alleles required for a trait) -, dominant (a single risk allele sufficient for a trait) - as well as chromosome X linked genes are involved. More important, in gene-mapping studies, affected families or

persons not sharing the same genetic variant (phenocopy) contribute negatively to the study outcome. Across populations, heterogeneity will cause difficulties for study replication, as it remains a question whether the genes found in one population are also risk factors in another[21]. For complex diseases, failure of detecting or exclusion of a specific risk factor in a given family does not mean that it is not a risk factor in other families. Heterogeneity has been reported for many traits including rare Mendelian disorders. An example is familial hemiplegic migraine (FHM), in which at least three genes lead to the development of this trait[22-25].

Interactions

Frequently risk alleles of multiple genes are required to cause a complex disorder, therefore, gene-gene interactions should be taken into account. For example, the Apolipoprotein ε4 allele (APOE*4) is an established risk factor for Alzheimer's disease, which is currently frequently included as a covariate in association studies[26,27]. Likewise, environmental factors can alter the effects of genes; gene-environment interaction. Without interactions, the risk of genes is considered to be additive; the risk for a subsequent harmful allele is increasing the total risk of the disorder independent of other risk factors. However, the risk of the allele can also be related to the presence of other risk factors, where the risk is much higher or lower than the expected risk based on the individual risk factors (non-linear effects, interaction). In a more extreme case, a disorder may be present only when multiple risk factors are present simultaneously (gene-epistasis). Currently a few statistical linkage methods can be employed to take multiple genes or environmental covariates into account and these are infrequently applied[28-30]. The sample sizes required for detecting interactions are substantial and may become prohibitive[31]. Genes interacting with the environment may be detected in specific populations only. Like with heterogeneity, this hampers study replication, which is considered good evidence for true causality[10,32,33].

Methods for identifying complex disease genes

As mentioned in the previous section, the exact strategy to identify the causative gene defect in monogenic disorders may differ from that in complex traits, but both strategies make use of positional cloning of genes (gene-mapping) and the analysis of candidate genes. For gene-mapping in complex diseases, linkage and sib-pair analysis are more suited, while association studies and transmission disequilibrium tests are more frequently employed to study candidate genes. Furthermore, the methods can be employed to study both dichotomous traits as well as quantitative trait loci (QTLs) in which the trait is a continuous variable[34].

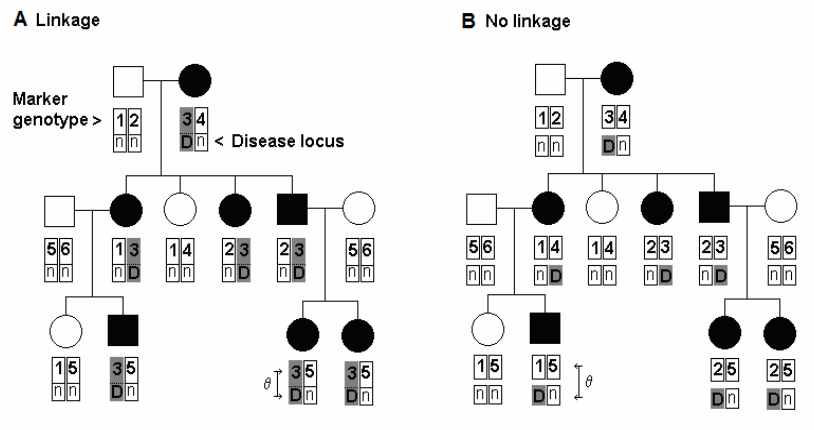
Linkage studies in extended families

Hallmark of linkage analysis is a process called recombination. During meiosis, recombination occurs between homologous chromosomes in either parent leading to two new hybrid chromosomes (gametes) that are transmitted to the offspring[35]. In case one of the parents carries a risk gene, only a part of the chromosome and marker alleles close to this gene will remain 'linked' to the gene over several generations (linkage disequilibrium). When the distance on the chromosome between the risk gene and the tested marker increases, the probability of recombination increases as well, and linkage disequilibrium diminishes. Testing for linkage in a family means that one evaluates to what extend the disorder co-segregates with a tested marker allele (single point analysis) or multiple marker haplotype (multipoint analysis)[36]. Under the null hypothesis the maximum likelihood of the observed marker data assuming no linkage with the disorder is calculated (recombination probability $\theta = 0.50$) (figure 1). This likelihood is subsequently compared with the maximum likelihood under the assumption that the given marker data (an allele or haplotype) is linked to the disorder ($\theta < 0.50$). The 10-log likelihood ratio, or LOD score, is calculated to indicate if the alternative hypothesis (i.e. the presence of linkage) is better or worse than the hypothesis assuming no

linkage. A LOD score above 3.3 is generally considered significant evidence for linkage in genome scans[33]. In addition to testing single families, the same approach can be applied to test multiple families at once. The marker of choice for linkage analysis is often the microsatellite marker as it has the highest informativeness (heterozygosity) in the parental transmission of alleles.

Figure 1

The principle of linkage presented in a single family.



To test the hypothesis of linkage the segregation of the disease locus *D* is correlated with genotypes of a multi- allelic marker. The likelihood of the family data is maximized for the recombination probability θ. In the linked family the dominant disorder is fully co-segregating with maternal allele 3 (figure 1A). The maximum likelihood is found at θ = 0.00 as no recombinations were observed between allele 3 and disease locus *D*. In the other family there is no linkage between any of the marker alleles and the disorder (figure 1B). There is no consistent co-segregation and several recombinations should have taken place in order to maintain linkage evidence. The maximum likelihood is found at θ = 0.50 equaling the null hypothesis of no linkage.

Statistical analysis for linkage can be done with parametric (model-based) or non-parametric (model-free) methods[37]. In the model-based approach several parameters have to be specified in order to calculate the maximum likelihood for the linkage statistic[36,37]. These are the gene frequency of the disorder, the phenocopy probability and the probabilities of being affected while carrying

one - or two copies of the risk allele (penetrances). With the parameters, the model and mode of inheritance are fixed. The correctness of this model may influence the study outcome[38-40]. Studies have shown that the effect of wrong specification of the linkage model in single point analysis is generally low, except for the mode of inheritance[41-43]. Segregation analysis can be used to find the best fitting mode of inheritance and model parameters[44-46]. For parametric linkage analysis several programs, such as FASTLINK or MENDEL are available[47-49]. In the model-free analysis the likelihood ratio is based only on the sharing of alleles between affected and non-affected individuals. These are compared with the expected random segregation of alleles. As a result the non-parametric approach is less susceptible to spurious results due to wrong specification of the model. The cost of using model-free methods is often a reduction in power to detect linkage as compared to a *correctly* specified model-based method[42,43]. Non-parametric linkage for dichotomous traits or QTLs can be tested with programs like MENDEL, GENEHUNTER, MERLIN and SOLAR[49-53].

Linkage analysis is sensitive to genetic heterogeneity[54]. A way to reduce this heterogeneity is to select a more homogenous sample of families. High-impact risk factors do exist for complex traits; there are families in which the disorder and risk alleles seem to follow a Mendelian pattern of inheritance (i.e. with an almost one-to-one correlation between the gene and the disorder). Often the phenotype of patients within these families is more consistent; symptoms may have an earlier age at onset or additional characteristics may be present[16,55,56]. Selecting these families, thereby reducing the heterogeneity, and applying linkage analysis has often been a successful first step into the molecular biology of complex neurological disorders[55,57,58]. Another approach to analyze a larger sample of families, is to take the heterogeneity of loci into account with programs like HOMOG, or to analyze the data using liability classes[36,59]. Finally, locus homogeneity of studies may also be improved by selecting a sample from more homogenous isolated populations[60,61].

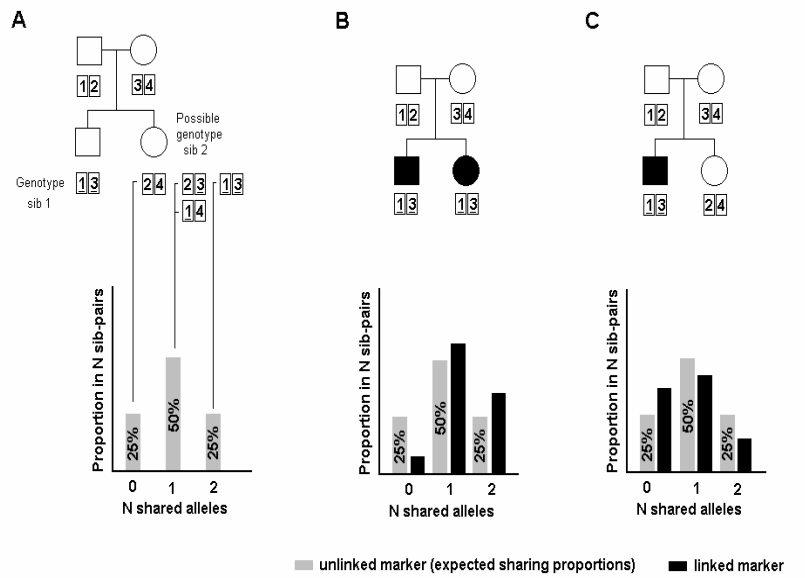
Sib-pair studies

In sib-pair studies the sharing of alleles between two siblings is studied in relation to the phenotype / disorder[62]. At a given locus, each sibling receives two of the four alleles that can be transmitted by the parents. As a result, a sib-pair will share 0, 1 or 2 alleles for a locus as shown in figure 2. The sharing is called identity-by-descent (IBD) in case the genotypes of the parents are discrete and the alleles that the siblings share can be scored exactly. In case the parents' genotypes are ambiguous and the exact sharing of the alleles (phase) cannot be determined, the sharing is called identity-by-state (IBS). With the IBS/IBD status of the individual pairs, a summation of the probabilities sharing 0, 1 or 2 alleles for all pairs can be calculated. For a random marker not related to the disorder these expected sharing probabilities are 25%, 50% and 25% (figure 2A). When linkage is present between the marker and the disorder, excess sharing of alleles is expected in affected (concordant) sib-pairs (figure 2B). A Z-score statistic, comparing the expected with the observed sharing probabilities for a marker, can be used as a test for linkage. Since no prior genetic model for the allele segregation needs to be assumed, sib-pair analysis is a non-parametric test for linkage. The marker of choice for sib-pair analysis is the microsatellite repeat marker, as multiple alleles give the most information about the parental transmission.

In addition to affected sib-pair analysis, other types of sib-pair analyses are possible. One is testing discordant sib-pairs; where only one sib is affected, in which the assumption is made that sib-pairs share less alleles than expected (figure 2C)[63,64]. Also QTLs can be studied where the trait variance between sibs is correlated with the number of shared alleles[65-67]. Affected sib-pair -, discordant sib-pair - and QTL analysis are implemented in various software packages like MAPMAKER SIBS, GENEHUNTER, MERLIN, MENDEL or SOLAR[49-53,62]. These will calculate the IBD probabilities as well as the various LOD score statistics.

Figure 2

The principle and expected sharing proportions of alleles in sib-pairs for a concordant - and discordant sib-pair study design given an unlinked and linked marker for a (dominant) disorder.



For N sib-pairs the expected proportions of 0, 1 or 2 alleles are 25%, 50% and 25% when there is no linkage, represented by the grey bars (figure 2A). A hypothetical marker linked to the disorder is shown in the black bars. In case of analyzing a sample of concordant sib-pairs this marker will show increased sharing proportions of 1 and 2 alleles (figure 2B). When analyzing a sample of discordant sib-pairs the marker will show a decreased sharing of 1 and 2 alleles (figure 2C). The heights of the black bars are potential outcomes of such analyses.

The sib-pair design is one of the most robust designs for gene mapping. Unlike association studies this design is not affected by confounding of population stratification. Also, as compared to the parametric or model-based linkage methods in extended families, they are less susceptible to large effects of heterogeneity, non-penetrance and phenocopies in single families[10]. Unfortunately, the power to detect loci in complex disorders for this design is often low[42,68]. When a locus is detected, the shared region on the genome

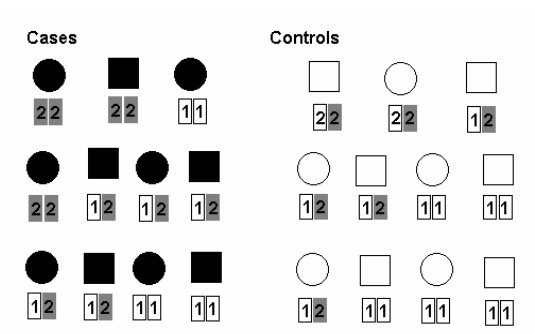
between two sibs is generally much larger compared to family or association studies hampering subsequent gene identification[69].

Association studies

In an association study the frequency of (a) specific marker allele(s) is compared between a group of unrelated patients (cases) and a group of unaffected individuals (controls) (figure 3). The assumption made is that the studied allele encodes a variant that increases the risk for the disorder. Compared to family-based designs, association studies have more power to detect genes with a relatively small influence on the disorder[68]. The use of SNP markers is preferred, as the power to detect gene effects is optimal for bi-allelic markers with a high gene frequency and the mutation rate of SNPs is generally lower[70,71]. Association studies can be applied to test single candidate genes and for genome scans testing up to 100 000 SNPs. Currently, the application of the association study for genome scans is still limited, however with the maturing of rapid and cheap SNP genotyping technology, the introduction of the HapMap project and advancing statistical methods this is about to change[72-74].

Figure 3

The principle of an association study.



The frequency of alleles for a tested marker is compared between affected cases and unaffected controls. In the case of association (in this case for allele 2) there is a substantial difference in frequencies.

Compared to family linkage studies the collection of data for association studies is simple and cost-effective. For late-onset disorders, like Alzheimer's disease, it may be the favored method of choice because relatives like parents and siblings are often not available anymore. Selection of cases and controls can be done using preferably large epidemiological studies[75]. Cases and controls are preferably matched for age, gender, population origin and other risk factors to control for confounding variables. For the statistical analysis of association studies many classic epidemiological methods can be applied[76]. These methods include the Pearson χ^2 statistic, odds ratio and relative risk analysis, logistic regression, survival analysis and ANOVA tests for QTLs. Before commencing on testing differences in allele frequencies, it is advisable to test for Hardy-Weinberg equilibrium (HWE) in cases and controls[77]. This can exclude large influences of selection bias, population stratification and genotyping errors.

Association between a marker and a disease will be found in four situations. In the first 'lucky' situation the tested marker is directly the functional polymorphism that causes the disorder. In this case, follow-up studies should aim at studying the gene effects preferably using other methods in independent study samples[32,75]. In the second situation the marker is in close linkage disequilibrium with the gene-variant that influences the disorder. Recent studies have shown that the linkage disequilibrium between several SNPs in candidate genes is variable and may extend to only a few kilobases[78,79]. The expected shared genomic regions between cases are likely to be very small[69]. Testing other SNPs in the same gene and studying for instance gene expression is therefore required for identification of the functional variant(s). The third reason for finding a positive association is confounding. A frequently mentioned problem is population stratification[10,80,81]. Here, the cases and controls are ascertained from two populations, which differ in gene frequencies and disease risk. In the case and control groups the representation of these populations is therefore unequal, and tested markers that have a

different gene frequency in both populations will be associated with the disorder. The fourth reason for finding association is that the result is a statistical false-positive[75,81-83]. Given a significance level of 0.05, which is frequently used for association studies, the probability of false-positive results is substantial. Given that up to 15 million variants and about 30 000 genes are present in the human genome, the probability of selecting the right SNP(s) *a priori* is extremely small[14,70,84,85] . This problem may be reduced by careful selection of candidate genes, but a recent review showed that many candidate gene associations may be false-positives[70].

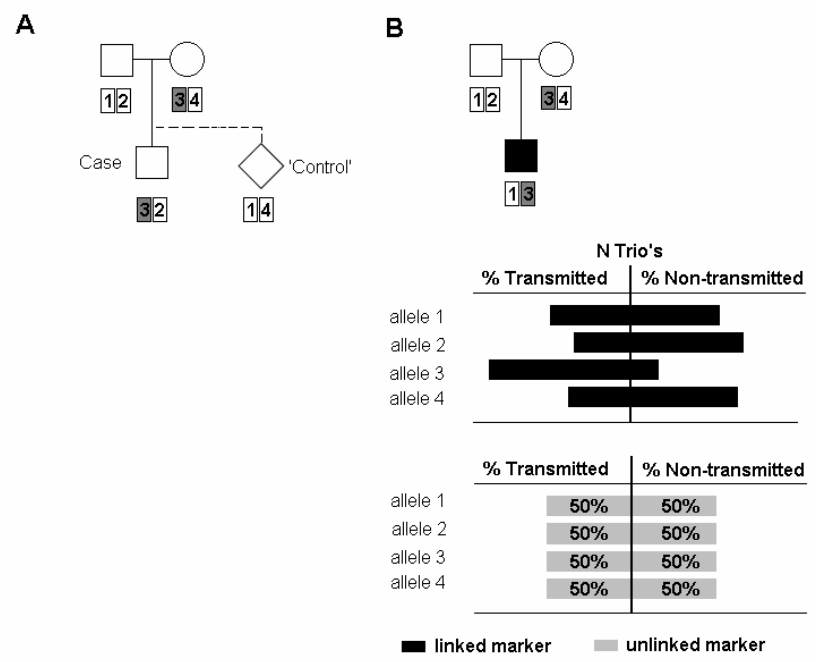
Several suggestions have been made to improve association study designs. These include testing for population stratification and other possible confounders, and to increase the significance level for reporting associations[14,75,86,87]. Also the study sample sizes, given the relative risks found for various associations, should be sufficiently high[70,88]. Taking into account the restrictions of the design, the ease of data collection, epidemiological analysis and the high power to pick up genes with relatively small effect size make the design a useful tool to study complex neurological disorders.

Family-based association studies

Family-based association studies are good alternatives for the straightforward case-control design to maintain the flexibility of the case-control approach without the confounding of population stratification. Several methods have been developed. The first was the haplotype relative risk (HRR) method[89,90]. Here, the two parents of a patient are also genotyped and the transmission of alleles to the case is compared with a pseudo-control assuming to have the alleles not transmitted to the case (diamond in figure 4A). Although this approach reduced the effects of population stratification, it could not eliminate them completely[91]. Another approach to the population stratification problem was the transmission disequilibrium test (TDT) (figure 4B)[92].

Figure 4

The Haplotype Relative Risk method and Transmission Disequilibrium Test principles.



A) In the haplotype relative risk method the non-transmitted alleles of the heterozygous parents are considered to be the genotype of the putative control (diamond). Standard analysis of association can subsequently be applied to test the hypothesis of association. B) The TDT approach compares the transmission of alleles from parents to offspring with the expected random Mendelian transmission (grey bars). In case of association the transmission of a specific allele (3) is increased while others are decreased (2,4), as shown for a hypothetical linked marker in the black bars.

The rationale behind the TDT is that the alleles are assumed to be transmitted randomly from parents to offspring. The TDT compares the number of times each allele was transmitted or not transmitted to an affected offspring by means of a χ^2 statistic. In case a marker allele is related to the studied disorder, the transmission of this allele will be increased in cases. The TDT test can be applied to study association of alleles as well as linkage, and is therefore useful for fine mapping of disease genes. For association testing, only one trio

should be taken per family if the original TDT statistic is applied, because cases are otherwise not independent[91,93]. For testing linkage, extended families can be tested as well.

Several extensions to the TDT have been proposed over the recent years. One was to use markers with multiple alleles accounting for the loss of information caused by parental homozygosity, while maintaining the advantage of correction for population stratification[94-97]. Furthermore, the use of haplotypes / multiple markers with - or without known haplotype data of the parents has been proposed[94,97]. Other adjustments were made by various authors to incorporate QTLs or covariates like age and sex[98-101]. However, most extensions were made to account for the TDT requirement to have both parents available, a substantial problem in late-onset disorders. The use of other family members, most notably siblings, was implemented in various tests to account for missing parent data[102-104]. Family members were used both for reconstruction of parental genotypes/haplotypes, as well as to test the transmission over different family members[94,102,105-107]. With the inclusion of family members, the use of the affection status of these members was also considered, increasing the sample size and information per family. As previously mentioned the association of a marker then becomes dependent on the number of family members present in the sample. Various statistics handling this problem have been developed and this has led to the current situation in which these methods have become a hybrid analysis of association, sib-pair and /or linkage that can be applied to numerous family constellations[93,108-111].

Genetics of neurological disorders studied in this thesis

In this thesis genetic epidemiological methods were applied to various neurological disorders. Here, short summaries of the disorders and their main genetic findings are presented in order to provide some background of their complexity.

Alzheimer's disease

Alzheimer's disease (AD) is characterized by a gradual onset of decline of memory and problems in at least one other area of cognition. Additional characteristics are a gradually progressive course of the disorder with a preserved level of consciousness. AD is a frequent late onset disorder, going from a male and female prevalence of 1.2% in people between 65 and 69 years old, to a prevalence of 33% in people aged up to 90 and older[3,4,112,113]. Diagnosis of AD is made based on extensive clinical anamnesis following the NINCDS-ADRDA criteria[17]. The diagnosis can sometimes be ambiguous, as both vascular dementia and Parkinson's disease have a large clinical overlap with AD[19,20]. The pathology of AD shows extra cellular plaques mainly composed of amyloid β peptide and intracellular neurofibrillary tangles containing hyperphosphorylated protein[114]. AD is also heterogeneous in age at onset and is often divided into groups with early-onset AD and late-onset AD for research and clinical purposes. The exact age which distincts early- from late-onset AD is fixed at 65 years, but remains a matter of discussion.

Particularly for early-onset families, but also for late-onset AD, twin and familial studies have shown that there is a strong heritable component for AD[13,115,116]. Exactly how much of the pathology of Alzheimer's disease can be explained by genetic factors is somewhat ambiguous; heritability estimates range from 29 to 78% [115]. This is mainly due to the variable late onset of the disorder, since persons might still become affected or are censored because of mortality. Segregation analysis of early-onset families has shown that there is not only a large single genetic component as the multifactorial model fits optimally[116].

For the early-onset Mendelian forms of AD several genes are known. The first gene that was found using linkage analysis in early-onset AD families was the transmembrane amyloid precursor protein (APP) on chromosome 21q21 [117,118]. Subsequently, mutations in two other genes, Presenilin-1 and Presenilin-2

(PSEN1 and PSEN2), were identified on chromosomes 14q24 and 1q42, respectively[119-121]. Although mutations in these three genes are frequently found in families with AD, these are accounting for only a few percent of the total number of AD cases in the general population. Another gene variant APOE*4, is accountable for a more substantial part of the population AD cases. The APOE*4 allele is an established risk factor for AD and is one of the most replicated associations studied[26,70]. New loci for late-onset AD have been found on chromosomes 10p11.23-q22.3, 12p12.3-q13.13 and 20p11.23-q12, but no consistent results of mutations related to AD have been found in these areas[122-124]. Gene-gene interaction and gene-environment interaction, especially with APOE*4 are frequently studied[125-127]. The interactions as well as the large heterogeneity make AD a complex disorder to study.

Migraine

Migraine is a common neurovascular disorder manifested by attacks of severe disabling headache. Anyone may have a migraine attack sometimes but the frequency of the attacks makes the disorder. The lifetime prevalence of migraine is up to 6% of men and 18% of women in the general population[1,2]. Diagnosis is made on the basis of a patient's history and is categorized in attack types using standardized diagnostic criteria defined by the International Headache Society (IHS)[16]. Attacks of migraine without aura (MO) are characterized by severe, often unilateral, throbbing headache that is aggravated by physical activity and is accompanied by other disabling neurological symptoms like vomiting, nausea, photophobia and/or phonophobia. One third of the migraine patients also develops visual aura symptoms, which are preceding or accompanying the headache; migraine with aura (MA).

Migraine is a complex disorder in which both environmental as well as genetic factors are involved[128,129]. The estimated heritability for the common types of migraine is 46% [11]. In addition, migraine can also be a part of autosomal dominant cerebrovascular syndromes, such as cerebral autosomal dominant

arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) and hereditary vascular retinopathy (HVR)[130-132]. Gene identification in the common forms of migraine has been extremely difficult, mainly because of the high prevalence, genetic heterogeneity and variable expression of the disorder. Furthermore, the consideration of patients with MA and / or MO attack types as being affected in families for linkage is an unresolved issue.

Mapping of migraine genes was initiated in Familial Hemiplegic Migraine; a rare autosomal dominant form of MA where patients additionally develop one-sided hemiparesis during attacks[16]. Two genes have been identified using this approach. The first gene (FHM1), CACNA1A, is located on chromosome 19p13 and encodes the $Ca_v2.1$ (formerly α1A) calcium channel subunit of P/Q-type calcium channels[22,58]. The second FHM gene (FHM2), ATP1A2, was identified on chromosome 1q23.2 and encodes the Na^+/K^+ ATPase α2 subunit[23,24]. Genome scans have also revealed several loci for the common types of migraine MA and MO. Loci identified in various single and multiple families were reported on chromosomes 1q31, 4q24, 6p12.2-p21.1, 11q24, 14q21.2-q22.3 and Xq24-q28 [133-139]. Recently, the Finnish MA locus on chromosome 4q24 has been replicated in MO families from Iceland[140]. Unfortunately, for none of the loci involved in the common types of migraine the causative gene has been identified yet.

Epilepsy

Epilepsy is characterized by recurrent unprovoked seizures with an abnormal electrical activity in the brain that leads to stereotype alterations in behavior[141]. The active epilepsy prevalence is 0.5% and is most often found in children and adolescents[142,143]. Epilepsy is a broad category of symptom complexes that arise from a large number of structural and functional brain disorders[144]. Epilepsy syndromes can be classified according to aetiology and seizure characteristics[18]. Different forms of seizures are: (1) *myoclonic seizures* during which a patient stares for a few seconds and sometimes blinks, (2) *atonic*

seizures during which a patient falls limply to the ground, (3) *tonic-clonic seizures* during which a patient becomes stiff and falls after which he has convulsions, and (4) *tonic seizures* which equal the tonic-clonic seizures except for the convulsions. Based on the aetiology, epilepsies can be put into the categories symptomatic, idiopathic and cryptogenic. Symptomatic are those epilepsies, which have a known underlying disorder, such as a stroke or tumors, and account for 20 to 40% of the epilepsy cases[141]. Idiopathic epilepsies are defined as epilepsies, which have no known underlying cause other than a hereditary predisposition. Cryptogenic are the epilepsies without any known associated risk factors and without presence of a familial predisposition. The epilepsy syndromes are characterized by combinations of clinical features like seizure types, age of onset and electroencephalogram (EEG) abnormalities.

Like for AD and migraine, familial studies and twin studies have shown that epilepsy is a disorder with genetic and environmental risk factors involved[145,146]. The estimated heritability of epilepsy ranges between 61 and 77% [12]. Of course, the contribution of genetic risk factors can vary with different epilepsy syndromes. Gene mapping studies have therefore focused on the idiopathic syndromes, which are frequently the rare monogenic variants of epilepsy syndromes. Positional cloning of the genes involved in these disorders has led to a multitude of mutations responsible for epilepsy[141]. Currently, nearly all known genes responsible for the epilepsy syndromes encode ion channels or functionally related structures. Examples are benign familial neonatal convulsions (BFNC) in which mutations have been found in the KCNQ2, KCNQ3 voltage gated potassium channels, or generalized epilepsy with febrile seizures (GEFS+) in which mutations have been described in the voltage gated sodium channels SCN1A, SCN1B and SCN2A[147-151]. However, for many other epilepsy syndromes the responsible genes have not been identified yet[152-154].

A part of the complexity of epilepsy syndromes is the overlap between various epilepsy syndromes that are described in literature. For example, in chapter eight a family is described, which fulfills the criteria of both autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE) as well as familial partial epilepsy with variable foci (FPEVF)[155,156]. Furthermore, the variable expression of the syndrome(s) in patients, the reduced penetrance of the Mendelian forms of epilepsy and the substantial heterogeneity within the syndromes make the mapping of these genes a challenge.

Scope of the thesis

Complex neurological disorders are frequent in the population and have a substantial impact on health care, socio–economic level and quality of life. Finding genetic risk factors involved in these disorders may clarify the pathophysiology and biochemical pathways, and may boost knowledge about the disorder and possible treatment. The finding of genetic risk factors in complex neurological disorders is nonetheless often difficult. In this thesis, some methodological issues involved in studying complex neurological traits with association studies were addressed. In addition, family-based mapping techniques were applied to an assortment of pedigrees with complex neurological traits. In the ***first chapter*** a general introduction of the complex trait, its related problems with gene-mapping and the current methodology are discussed. The ***second chapter*** focuses on a problem that may be encountered in association studies: population stratification. A simple overview of methodology to test and, if necessary, circumvent population stratification is provided. Furthermore, the probability of finding false-positive association was studied in relation to population diversity and genetic drift. In the ***third chapter***, an approach is presented to evaluate false-positive gene-gene interactions found in association studies. This approach may greatly improve the study findings and detect statistical fluctuations in results. In ***chapter four*** the comorbidity and risk of migraine and Raynaud Phenomenon was studied

with a locus involved in Hereditary Vascular Retinopathy. A TDT approach was applied in a single family to study if the HVR haplotype would increase the susceptibility for both disorders. In ***chapter five*** segregation analyses were used to study how migraine attacks with - and without aura are inherited in Dutch migraine families. The effect of including patients with MA and MO in extended MO families was studied as well. In ***chapter six***, linkage analysis in seven large Dutch MO families was performed, which aimed at locating novel loci for migraine without aura. An interesting conclusion from this study is the confirmation of the Finnish locus on chromosome 4q24 known to be involved in MA. This study also showed the difficulties of linkage analysis in complex disease, as the heterogeneity of the disorder affected the linkage findings even under a homogeneous selection of families. In ***chapter seven*** heterogeneity of familial cortical tremor with epilepsy was shown with the exclusion of a Japanese locus on chromosome 8q23.3-q24.1. The mapping and replication of a locus for familial partial epilepsy with variable foci on chromosome 22q11-q12 in ***chapter eight*** shows that parametric linkage analysis in extended pedigrees can be a useful tool for mapping genes in more rare and less heterogeneous complex neurological disorders.

References

1. Lipton RB, Stewart WF. Migraine headaches: epidemiology and co morbidity. Clin Neurosci. 1998;5:2-9.
2. Launer LJ, Terwindt GM, Ferrari MD. The prevalence and characteristics of migraine in a population-based cohort: the GEM study. Neurology. 1999;53:537-42.
3. Fitzpatrick AL, Kuller LH, Ives DG, et al. Incidence and prevalence of dementia in the Cardiovascular Health Study. J Am Geriatr Soc. 2004;52:195-204.
4. Bachman DL, Wolf PA, Linn R et al. Prevalence of dementia and probable senile dementia of the Alzheimer type in the Framingham Study. Neurology. 1992;42:115-9.
5. Breslau N, Rasmussen BK. The impact of migraine: Epidemiology, risk factors, and co-morbidities. Neurology. 2001; 56 Suppl 1:S4-12.
6. Witthaus E, Ott A, Barendregt JJ et al. Burden of mortality and morbidity from dementia. Alzheimer Dis Assoc Disord. 1999;13:176-81.
7. Baker GA, Nashef L, van Hout BA. Current issues in the management of epilepsy: the impact of frequent seizures on cost of illness, quality of life, and mortality. Epilepsia. 1997;38 Suppl 1:S1-8.

8. Bishop M, Allen CA. The impact of epilepsy on quality of life: a qualitative analysis. Epilepsy Behav. 2003;4:226-33.
9. Khoury MJ, Beaty TH, Cohen BH. Fundamentals of genetic epidemiology. Oxford university press. Oxford. 1993. ISBN 0195052889.
10. Lander ES, Schork NJ. Genetic dissection of complex traits. Science. 1994;265:2037-48.
11. Mulder EJ, Van Baal C, Gaist D et al. Genetic and environmental influences on migraine: a twin study across six countries. Twin Res. 2003;6:422-31.
12. Kjeldsen MJ, Kyvik KO, Friis ML, Christensen K. Genetic and environmental factors in febrile seizures: a Danish population-based twin study. Epilepsy Res. 2002;51:167-77.
13. Pedersen NL, Gatz M, Berg S, Johansson B. How heritable is Alzheimer's disease late in life? Findings from Swedish twins. Ann Neurol. 2004;55:180-5.
14. Risch NJ. Searching for genetic determinants in the new millennium. Nature. 2000; 15;405:847-56.
15. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR. 4th edition. Arlington. 2000. ISBN 0890420254.
16. "International Headache Society". Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. Cephalalgia 1988;supplement 7 8:1-96.
17. McKhann G, Drachman D, Folstein M et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984;34:939-44.
18. Commission on Classification and Terminology of the International League Against Epilepsy. Proposal for revised classification of epilepsies and epileptic syndromes. Epilepsia. 1989;30:389-99.
19. Roman GC, Tatemichi TK, Erkinjuntti T et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. Neurology. 1993;43:250-60.
20. Perl DP, Olanow CW, Calne D. Alzheimer's disease and Parkinson's disease: distinct entities or extremes of a spectrum of neurodegeneration? Ann Neurol. 1998;44 3 Suppl 1:S19-31.
21. Page GP, George V, Go RC et al. "Are we there yet?": Deciding when one has demonstrated specific genetic causation in complex diseases and quantitative traits. Am J Hum Genet. 2003;73:711-19.
22. Ophoff RA, Terwindt GM, Vergouwe MN et al. Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. Cell. 1996;1;87:543-52.
23. De Fusco M, Marconi R, Silvestri L et al. Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. Nat Genet. 2003;33:192-6.
24. Ducros A, Joutel A, Vahedi K et al. Mapping of a second locus for familial hemiplegic migraine to 1q21-q23 and evidence of further heterogeneity. Ann Neurol. 1997;42:885-90.
25. Joutel A, Ducros A, Vahedi K et al. Genetic heterogeneity of familial hemiplegic migraine. Am J Hum Genet. 1994;55:1166-72.
26. Farrer LA, Cupples LA, Haines JL et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer's disease. A meta-analysis. APOE and Alzheimer's Disease Meta Analysis Consortium. JAMA. 1997;278:1349-56.
27. Kim KW, Jhoo JH, Lee KU et al. No association between presenilin 1 (PS1) intronic polymorphism and sporadic Alzheimer's disease in Koreans. J Neural Transm. 2000;107:1191-200.

28. Lucek P, Hanke J, Reich J et al. Multi-locus nonparametric linkage analysis of complex trait loci with neural networks. Hum Hered. 1998;48:275-84.
29. Beekman M, Heijmans BT, Martin NG et al. Two-locus linkage analysis applied to putative quantitative trait loci for lipoprotein(a) levels. Twin Res. 2003;6:322-4.
30. Liu X, Fallin MD, Kao WL. Genetic dissection methods: designs used for tests of gene-environment interaction. Curr Opin Genet Dev. 2004;14:241-45.
31. Gauderman WJ. Sample size requirements for association studies of gene-gene interaction. Am J Epidemiol. 2002;155:478-84.
32. Hirschhorn JN, Altshuler D. Once and again-issues surrounding replication in genetic association studies. J Clin Endocrinol Metab. 2002;87:4438-41.
33. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet. 1995;11:241-7.
34. Blangero J. Localization and identification of human quantitative trait loci: king harvest has surely come. Curr Opin Genet Dev. 2004;14:233-40.
35. Strachan T, Read A. Human Molecular Genetics. 3rd edition. Garland Science/Taylor & Francis Group. New York. 2003. ISBN 0815341822.
36. Ott J, Terwilliger JD. Handbook of human genetic linkage. Baltimore, MD: Johns Hopkins University Press. 1994. ISBN 0801848032.
37. Ott J, Hoh J. Statistical approaches to gene mapping. Am J Hum Genet. 2000;67:289-94.
38. Weeks DE, Lehner T, Squires-Wheeler E, Kaufmann C, Ott J. Measuring the inflation of the lod score due to its maximization over model parameter values in human linkage analysis. Genet Epidemiol. 1990;7:237-43.
39. Clerget-Darpoux F, Babron MC, Bonaiti-Pellie C. Power and robustness of the linkage homogeneity test in genetic analysis of common disorders. J Psychiatr Res. 1987;21:625-30.
40. Risch N, Giuffra L. Model misspecification and multipoint linkage analysis. Hum Hered. 1992;42:77-92.
41. Greenberg DA, Abreu P, Hodge SE. The power to detect linkage in complex disease by means of simple LOD-score analyses. Am J Hum Genet. 1998;63:870-9.
42. Abreu PC, Greenberg DA, Hodge SE. Direct power comparisons between simple LOD scores and NPL scores for linkage analysis in complex diseases. Am J Hum Genet. 1999;65:847-57.
43. Hodge SE. Model-free vs. model-based linkage analysis: a false dichotomy? Am J Med Genet. 2001;105:62-4.
44. Morton NE, Rao DC, Lalouel JM. Methods in genetic epidemiology. S. Karger Publishing. New York. 1983. ISBN 3805536682.
45. Lalouel JM, Morton NE. Complex segregation analysis with pointers. Hum Hered. 1981;31:312-21.
46. Lalouel JM, Rao DC, Morton NE, Elston RC. A unified model for complex segregation analysis. Am J Hum Genet. 1983;35:816-26.
47. Lathrop GM, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci U S A. 1984;81:3443-6.
48. Lathrop GM, Lalouel JM. Easy calculations of lod scores and genetic risks on small computers. Am J Hum Genet. 1984;36:460-5.
49. Lange K, Weeks D, Boehnke M. Programs for Pedigree Analysis: MENDEL, FISHER, and dGENE. Genet Epidemiol. 1988;5:471-2.
50. Li H, Schaid DJ. GENEHUNTER: application to analysis of bipolar pedigrees and some extensions. Genet Epidemiol. 1997;14:659-63.

51. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet. 1996;58:1347-63.
52. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet. 1998;62:1198-211.
53. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet. 2002;30:97-101.
54. Chiano MN, Yates JR. Linkage detection under heterogeneity and the mixture problem. Ann Hum Genet. 1995;59:83-95.
55. Bowcock AM, Anderson LA, Friedman LS et al. THRA1 and D17S183 flank an interval of < 4 cM for the breast-ovarian cancer gene (BRCA1) on chromosome 17q21. Am J Hum Genet. 1993;52:718-22.
56. Terwindt GM, Ophoff RA, Haan J et al. Familial hemiplegic migraine: a clinical comparison of families linked and unlinked to chromosome 19. DMGRG. Cephalalgia. 1996;16:153-5.
57. Brzustowicz LM, Hodgkinson KA, Chow EW et al. Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22. Science. 2000;28:288:678-82.
58. Joutel A, Bousser MG, Biousse V et al. A gene for familial hemiplegic migraine maps to chromosome 19. Nat Genet. 1993;5:40-5.
59. Bhat A, Heath SC, Ott J. Heterogeneity for multiple disease loci in linkage analysis. Hum Hered 1999;49:229-231.
60. Peltonen L, Palotie A, Lange K. Use of population isolates for mapping complex traits. Nat Rev Genet. 2000;1:182-90.
61. Vaessen N, Heutink P, Houwing-Duistermaat JJ et al. A genome-wide search for linkage-disequilibrium with type 1 diabetes in a recent genetically isolated population from the Netherlands. Diabetes. 2002;51:856-9.
62. Kruglyak L, Lander ES. Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet. 1995;57:439-54.
63. Risch N, Zhang H. Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science. 1995;268:1584-9.
64. Zhao H, Zhang H, Rotter JI. Cost-effective sib-pair designs in the mapping of quantitative-trait loci. Am J Hum Genet. 1997;60:1211-21.
65. Chen WM, Broman KW, Liang KY. Quantitative trait linkage analysis by generalized estimating equations: unification of variance components and Haseman-Elston regression. Genet Epidemiol. 2004;26:265-72.
66. Elston RC, Buxbaum S, Jacobs KB, Olson JM. Haseman and Elston revisited. Genet Epidemiol. 2000;19:1-17.
67. Putter H, Sandkuijl LA, van Houwelingen JC. Score test for detecting linkage to quantitative traits. Genet Epidemiol. 2002;22:345-55.
68. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science. 1996;273:1516-7.
69. Houwen RH, Baharloo S, Blankenship K et al. Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. Nat Genet. 1994;8:380-6.
70. Lohmueller KE, Pearce CL, Pike M et al. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet. 2003;33:177-82.
71. Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nat Gen. 2003;33: Supplement 228-37.

72. Kennedy GC, Matsuzaki H, Dong S et al. Large-scale genotyping of complex DNA. Nat Biotechnol. 2003;21:1233-7.
73. International HapMap Consortium. Integrating ethics and science in the International HapMap Project. Nat Rev Genet. 2004;5:467-75.
74. John S, Shephard N, Liu G et al. Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. Am J Hum Genet. 2004;75:54-64.
75. Little J, Bradley L, Bray MS et al. Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. Am J Epidemiol. 2002;156:300-10.
76. Rothman J, Greenland S. Modern epidemiology. 2nd edition. Lippincott Williams & Wilkins. Philadelphia. 1998. ISBN 0316757764.
77. Sasieni PD. From genotypes to genes: doubling the sample size. Biometrics. 1997;53:1253-61.
78. Long JR, Zhao LJ, Liu PY et al. Patterns of linkage disequilibrium and haplotype distribution in disease candidate genes. BMC Genet. 2004;5:11.
79. North BV, Curtis D, Martin ER, et al. Further Investigation of Linkage Disequilibrium SNPs and their Ability to Identify Associated Susceptibility Loci. Ann Hum Genet. 2004;68:240-8.
80. Knowler WC, Williams RC, Pettitt DJ, Steinberg AG. Gm3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. Am J Hum Genet. 1988;43:520-6.
81. Daly AK, Day CP. Candidate gene case-control association studies: advantages and potential pitfalls. Br J Clin Pharmacol. 2001;52:489-99.
82. Baron M. The search for complex disease genes: fault by linkage or fault by association? Mol Psychiatry. 2001;6:143-9.
83. Ioannidis JP. Genetic associations: false or true? Trends Mol Med. 2003;9:135-8.
84. Sachidanandam R, Weissman D, Schmidt SC et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature. 2001;409:928-33.
85. Lander ES, Linton LM, Birren B et al. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature. 2001;409:860-921.
86. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999;55:997-1004.
87. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet. 1999;65:220-8.
88. Ioannidis JP, Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG. Genetic associations in large versus small studies: an empirical assessment. Lancet. 2003;361:567-71.
89. Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. Ann Hum Genet. 1987;51:227-33.
90. Schaid DJ. Transmission disequilibrium, family controls, and great expectations. Am J Hum Genet. 1998;63:935-41.
91. Spielman RS, Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet. 1996;59:983-9.
92. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet. 1993;52:506-16.
93. Martin ER, Kaplan NL, Weir BS. Tests for linkage and association in nuclear families. Am J Hum Genet. 1997;61:439-48.

94. Lazzeroni LC, Lange K. A conditional inference framework for extending the transmission / disequilibrium test. Hum Hered. 1998;48:67-81.
95. Bickeboller H, Clerget-Darpoux F. Statistical properties of the allelic and genotypic transmission / disequilibrium test for multiallelic markers. Genet Epidemiol. 1995;12:865-70.
96. Kaplan NL, Martin ER, Weir BS. Power studies for the transmission/disequilibrium tests with multiple alleles. Am J Hum Genet. 1997;60:691-702.
97. Zhao H, Zhang S, Merikangas KR et al. Transmission / disequilibrium tests using multiple tightly linked markers. Am J Hum Genet. 2000;67:936-46.
98. Allison DB. Transmission-disequilibrium tests for quantitative traits. Am J Hum Genet. 1997;60:676-90.
99. Rabinowitz D. A transmission disequilibrium test for quantitative trait loci. Hum Hered. 1997;47:342-50.
100. Monks SA, Kaplan NL. Removing the sampling restrictions from family-based tests of association for a quantitative-trait locus. Am J Hum Genet. 2000;66:576-92.
101. Ghosh S, Reich T. The sib TDT adjusted for age of disease onset. Ann Hum Genet. 2004;68:249-56.
102. Curtis D. Use of siblings as controls in case-control association studies. Ann Hum Genet. 1997;61:319-33.
103. Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission / disequilibrium test. Am J Hum Genet. 1998;62:450-8.
104. Horvath S, Laird NM. A discordant-sibship test for disequilibrium and linkage: no need for parental data. Am J Hum Genet. 1998;63:1886-97.
105. Spielman RS, Ewens WJ. TDT clarification. Am J Hum Genet. 1999;64:668.
106. Knapp M. Using exact P values to compare the power between the reconstruction-combined transmission / disequilibrium test and the sib transmission/disequilibrium test. Am J Hum Genet. 1999;65:1208-10.
107. Knapp M. The transmission/disequilibrium test and parental-genotype reconstruction: the reconstruction-combined transmission/ disequilibrium test. Am J Hum Genet. 1999;64:861-70.
108. Martin ER, Monks SA, Warren LL, Kaplan NL. A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am J Hum Genet. 2000;67:146-54.
109. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol. 2000;19 Suppl 1:S36-42.
110. Rabinowitz D, Laird N. A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. Hum Hered. 2000;50:211-23.
111. Schulze TG, McMahon FJ. Genetic association mapping at the crossroads: which test and why? Overview and practical guidelines. Am J Med Genet. 2002;114:1-11.
112. Rocca WA, Hofman A, Brayne C, et al. The prevalence of vascular dementia in Europe: facts and fragments from 1980-1990 studies. EURODEM-Prevalence Research Group. Ann Neurol. 1991;30:817-24.
113. van Duijn CM, Clayton D, Chandra V et al. Familial aggregation of Alzheimer's disease and related disorders: a collaborative re-analysis of case-control studies. EURODEM Risk Factors Research Group. Int J Epidemiol. 1991;20 Suppl 2:S13-20.
114. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol (Berl). 1991;82:239-59.

115. Pedersen NL, Posner SF, Gatz M. Multiple-Threshold Models for Genetic Influences on Age of Onset for Alzheimer Disease: Findings in Swedish Twins. Am J Med Genet. 2001;105:724-28
116. van Duijn CM, Farrer LA, Cupples LA, Hofman A. Genetic transmission of Alzheimer's disease among families in a Dutch population-based study. J Med Genet. 1993;30:640-6.
117. St George-Hyslop PH, Tanzi RE, Polinsky RJ et al. The genetic defect causing familial Alzheimer's disease maps on chromosome 21. Science. 1987;235:885-90.
118. Goate A, Chartier-Harlin MC, Mullan M et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature. 1991;349:704-6.
119. Schellenberg GD, Bird TD, Wijsman EM et al. Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. Science. 1992;258:668-71.
120. Sherrington R, Rogaev EI, Liang Y et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature. 1995;375:754-60.
121. Rogaev EI, Sherrington R, Rogaeva EA et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. Nature. 1995;376:775-8.
122. Goddard KA, Olson JM, Payami H, Van Der Voet M, Kuivaniemi H, Tromp G. Evidence of linkage and association on chromosome 20 for late-onset Alzheimer disease. Neurogenetics. 2004;5:121-8.
123. Lendon C, Craddock N. Susceptibility gene(s) for Alzheimer's disease on chromosome 10. Trends Neurosci. 2001;24:557-9.
124. Pericak-Vance MA, Bass MP, Yamaoka LH, Gaskell PC, Scott WK et al. Complete genomic screen in late-onset familial Alzheimer disease. Evidence for a new locus on chromosome 12. JAMA. 1997 Oct 15;278:1237-41.
125. Mayeux R. Gene-environment interaction in late-onset Alzheimer disease: the role of apolipoprotein-epsilon4. Alzheimer Dis Assoc Disord. 1998;12 Suppl 3:S10-5.
126. Chandra V, Pandav R. Gene-environment interaction in Alzheimer's disease: a potential role for cholesterol. Neuroepidemiology. 1998;17:225-32.
127. Combarros O, Alvarez-Arcaya A, Oterino A et al. Polymorphisms in the presenilin 1 and presenilin 2 genes and risk for sporadic Alzheimer's disease. J Neurol Sci. 1999;171:88-91.
128. Gervil M, Ulrich V, Kaprio J et al. The relative role of genetic and environmental factors in migraine without aura. Neurology. 1999;53:995-9.
129. Ziegler DK, Hur YM, Bouchard TJ et al. Migraine in twins raised together and apart. Headache. 1998;38:417-22.
130. Verin M, Rolland Y, Landgraf F, et al. New phenotype of the cerebral autosomal dominant arteriopathy mapped to chromosome 19: migraine as the prominent clinical feature. J Neurol Neurosurg Psychiatry. 1995;59:579-85.
131. Dichgans M, Mayer M, Uttner I et al. The phenotypic spectrum of CADASIL: clinical findings in 102 cases. Ann Neurol. 1998; 44:731-9.
132. Terwindt GM, Haan J, Ophoff RA et al. Clinical and genetic analysis of a large Dutch family with autosomal dominant vascular retinopathy, migraine and Raynaud's phenomenon. Brain. 1998;121:303-16.
133. Lea RA, Shepherd AG, Curtain RP et al. A typical migraine susceptibility region localizes to chromosome 1q31. Neurogenetics 2002;4:1:17-22.
134. Wessman M, Kallela M, Kaunisto MA et al. A susceptibility locus for migraine with aura, on chromosome 4q24. Am J Hum Genet 2002;70:652-62.
135. Carlsson A, Forsgren L, Nylander PO et al. Identification of a susceptibility locus for migraine with and without aura on 6p12.2-p21.1.Neurology 2002;59:11:1804-7.

136. Cader ZM, Noble-Topham S, Dyment DA et al. Significant linkage to migraine with aura on chromosome 11q24. Hum Mol Genet. 2003;12:2511-7.
137. Soranga D, Vettori A, Carraro G et al. A locus for migraine without aura maps on chromosome 14q21.2-q22.3. Am J Hum Genet. 2003;72:161-7.
138. Nyholt DR, Dawkins JL, Brimage PJ et al. Evidence for an X-linked genetic component in familial typical migraine. Hum Mol Genet 1998;7:459-63.
139. Nyholt DR, Curtain RP, Griffiths LR. Familial typical migraine: significant linkage and localization of a gene to Xq24-28. Hum Genet 2000;107:18-23.
140. Bjornsson A, Gudmundsson G, Gudfinnsson E et al. Localization of a gene for migraine without aura to chromosome 4q21. Am J Hum Genet. 2003;73:986-93.
141. Tijink-Callenbach PMC. Clinical and genetic studies in heriditary epilepsy syndromes. Febodruk B.V. 2003. ISBN 9090172637.
142. Hauser WA. The prevalence and incidence of convulsive disorders in children. Epilepsia. 1994;35 Suppl 2:S1-6.
143. Wright J, Pickard N, Whitfield A, Hakin N. A population-based study of the prevalence, clinical characteristics and effect of ethnicity in epilepsy. Seizure. 2000;9:309-13.
144. McNamara JO. The neurbiological basis of epilepsy. Trends Neurosci. 1992;15:357-9.
145. Berkovic SF, Howell RA, Hay DA, Hopper JL. Epilepsies in twins: genetics of the major epilepsy syndromes. Ann Neurol. 1998;43:435-45.
146. Jain S, Padma MV, Puri A, Jyoti, Maheshwari MC. Occurrence of epilepsies in family members of Indian probands with different epileptic syndromes. Epilepsia. 1997;38:237-44.
147. Singh NA, Charlier C, Stauffer D et al. A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. Nat Genet. 1998;18:25-9.
148. Hirose S, Zenri F, Akiyoshi H et al. A novel mutation of KCNQ3 (c.925T-->C) in a Japanese family with benign familial neonatal convulsions. Ann Neurol. 2000;47:822-6.
149. Wallace RH, Wang DW, Singh R et al. Febrile seizures and generalized epilepsy associated with a mutation in the Na+-channel beta1 subunit gene SCN1B. Nat Genet. 1998;19:366-70.
150. Escayg A, MacDonald BT, Meisler MH et al. Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2. Nat Genet. 2000;24:343-5.
151. Heron SE, Crossland KM, Andermann E et al. Sodium-channel defects in benign familial neonatal-infantile seizures. Lancet. 2002;360:851-2.
152. Cormier-Daire V, Dagoneau N, Nabbout R et al. A gene for pyridoxine-dependent epilepsy maps to chromosome 5q31. Am J Hum Genet. 2000;67:991-3.
153. Guipponi M, Rivier F, Vigevano F et al. Linkage mapping of benign familial infantile convulsions (BFIC) to chromosome 19q. Hum Mol Genet. 1997;6:473-7.
154. Zara F, Gennaro E, Stabile M et al. Mapping of a locus for a familial autosomal recessive idiopathic myoclonic epilepsy of infancy to chromosome 16p13. Am J Hum Genet. 2000;66:1552-7.
155. Scheffer IE, Bhatia KP, Lopes-Cendes I et al. Autosomal dominant nocturnal frontal lobe epilepsy. A distinctive clinical disorder. Brain. 1995;118:61-73.
156. Scheffer IE, Phillips HA, O'Brien CE et al. Familial partial epilepsy with variable foci: a new partial epilepsy syndrome with suggestion of linkage to chromosome 2. Ann Neurol. 1998;44:890-9.

CHAPTER 2

Population stratification

JJ Hottenga[1,2], JJ Houwing-Duistermaat[2], LA Sandkuijl[1,2†], CM van Duijn[2]

1. Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.
2. Department of Epidemiology & Biostatistics, Erasmus Medical Center Rotterdam, the Netherlands.

Based on Nederlands Tijdschrift voor Geneeskunde 2002;146:17-22

† In memory of Lodewijk A. Sandkuijl

Abstract

Association studies have been criticized because of the failure to replicate results. Population stratification is often cited as being a major cause for the large number of false-positive findings. The aim of this study was to examine how much population diversity and stratification is required to cause spurious associations, and whether this is caused by genetic drift. To this end we simulated genetically isolated populations with various degrees of founder effects and genetic drift. Our study shows that in case one marker is tested the probability of finding a spurious association with an increased risk of 1.50 is less than 5%. Only when the genetic drift is very strong, or in case that the stratification of the two populations is extremely discordant, the risk for spurious association exceeds 5%. In case of testing multiple markers population stratification may become an issue. Methods that can be applied to test and correct for population stratification are discussed as well.

Keywords

Population stratification, population admixture, association, genetic drift.

An introduction to population stratification

The validity of genetic association or population-based case-control studies in genetic research remains subject of substantial debate[1-6]. In association studies differences in disease frequencies are correlated with differences in allele frequencies for a genetic marker. The association study is more cost effective compared to family-based designs and has a high power for finding genes in complex disorders. In addition, it is often an essential step for cloning disease genes[1-3]. However, the high rate of false-positive outcomes and the possibility of confounding have resulted in strict guidelines and even rejection of this study design in important journals[4,7]. A frequently mentioned cause for false-positive results in association studies is unknown or hidden population stratification[8].

Two well-known examples of population stratification are the so called 'chopsticks' gene in the population of San Francisco and the association of the *Gm3;5;13;14* marker with Diabetes Mellitus type II in Pima Indians[9,10]. Both will be discussed here because they present two different mechanisms; population stratification and population admixture. Regarding the first example, if one assumes that the association between the ability of eating with chopsticks and the HLA-10 allele is studied by randomly taking persons from the population of San Francisco, the result will be that eating with chopsticks is associated with an increased frequency of the HLA-10 allele[9]. The reason for this association is that the HLA-10 allele is more frequent in Asians as compared to Caucasians, and so is eating with chopsticks. Therefore, there will be more Asian people in the 'cases' as compared to the 'controls' and also the HLA-10 frequency will be increased in cases as compared to controls. In this example, population substructure or population stratification confounded the results. If a sample of only Asians, Caucasians or a population-matched sample had been studied, there would have been no association.

The second example is a study performed in Pima Indians, in which the prevalence of Diabetes type II was compared with the presence of the *Gm3;5;13;14* haplotype[10]. These Indians have lived for generations as a separate group. For this reason their genetic background is different in comparison to that of Caucasians (because of genetic drift, different founders). Under the influence of environmental as well as genetic factors the metabolism of the Indians has changed in a way that, given a western diet, they have an increased risk for type II diabetes. In recent generations the Indians mixed with the Caucasian population. This admixed population was studied in an association study and the results showed that a decreased frequency of the *Gm3;5;13;14* haplotypes was associated with an increased risk for Diabetes type II. However, when the amount of Indian ancestry was taken into account in cases and controls, the results did not show an association anymore. The specific *Gm* haplotype was a measure of Indian-Caucasian population admixture and not necessarily the causal or closely linked factor related to Diabetes type II.

Unknown population stratification and population admixture is difficult to account for in association studies and will occur whenever cases and controls are not matched for their population background[8]. Also in prospective cohort studies population stratification can confound results if the cohort is based on two or more different subpopulations that have different risks for the studied disorder. In the two examples discussed earlier there were two very different subpopulations, however there may also be several more similar subpopulations. If population admixture between the subgroups has taken place over a period of a few generations, the exact genetic background of persons may be impossible to define. Furthermore, the relevant allele frequency differences between population subgroups are often unknown.

Allele frequency differences between populations are caused by a number of factors such as founder effects, genetic drift, assortative mating, disease

bottlenecks and mutation rates[11]. Evaluation of empirical data on allele frequencies of genes across populations suggests that there will only be a relatively small bias in the risk estimate due to population stratification, except under extreme conditions[12-14]. How frequently these extreme conditions may occur within populations remains unknown. Here, two populations were stratified with various mixtures in cases and controls, and the effect of allele frequency differences on finding spurious associations was tested. Subsequently, the question was addressed how frequently genetic drift and founder effects will result in a spurious association caused by population stratification. To this end several mixed populations were simulated under various conditions. Finally, methods to control population stratification are discussed at the end of the chapter.

Methods

An important question is what differences in allele frequencies between populations will lead to substantial confounding. This effect was examined using an association study design with a two-allele polymorphism (alleles *A* and *B*) stratifying two different populations (1 and 2). The parameters that define the effect of confounding in this situation are the proportion of persons from population 2, *P* in cases and *Q* in controls, and the frequencies of allele *A* carriers (*AA* + *AB*) F_1 for population 1 and F_2 for population 2. With these parameters, the odds ratio comparing the allele *A* carrier frequency in cases vs. controls was calculated to quantify the direct effect of population stratification. We did not consider p-values or 95% confidence intervals as these are largely determined by sample size of an individual association study[15].

In order to study the effect of genetic drift, we considered a mixed population, which comprised a large general outbred population and a smaller isolated population. In the general population the genotype frequencies were assumed to be constant over generations and mating between individuals was considered to be random. Also no mutations occurred and no selection existed

(that is full Hardy-Weinberg equilibrium (HWE) applied). A single genetic marker not related to the disorder, with two alleles *A* and *B* was modeled. We restricted our study to the effect of a bi-allelic marker e.g. a single nucleotide polymorphism (SNP), because these markers will be most suitable for association studies[16]. The allele frequencies of the general population were set to constant values of 0.05, 0.10, 0.30 and 0.50. These values correspond to a proportion of subjects carrying the *A* allele (F_{gen}) of 0.10, 0.19, 0.51, and 0.75 respectively. These frequencies F_{gen} cover the typical range that is usually addressed in candidate gene studies.

The isolated inbred population was simulated under three different conditions generating 25 000 replicates for each condition. In each condition the founders for the isolated population were selected from the general population (founder effect). For *condition 1*, 3000 founders were sampled from the general population. For *conditions 2* and *3*, the number of founders sampled was 100 and 38, respectively. Random mating was assumed within the population isolate and generations were discrete. The number of children per couple ranged from 0 to maximally 20 and was assumed to follow a negative binomial distribution with parameters (p = 0.40, n = 20)[11]. This resulted in an average population growth of 1.16 in each generation. Migration between the isolate and the general population was not allowed, nor was there any new mutation or selection. The isolation progressed for 10, 40 and 52 generations for *conditions 1* to *3*, respectively. As a result of the growth over generations, the population was large enough to have a stable allele *A* carrier frequency F_{iso} for the marker in the last generation. Due to founder effects and genetic drift, differences in genetic make-up exist between the general population and the simulated isolates. This effect is mild for *condition 1*, severe for *condition 2,* and extreme for *condition 3*.

The simulated data of the isolated and general population were analyzed in a population-stratified case-control design. The exact stratification of the

populations in cases and controls will depend on differences in disease prevalence between the populations, selection bias and / or referral bias. Various mixtures of the general and isolate population between cases and controls were studied. The mixtures were again described by the parameters: P being the proportion of persons from the isolated population in cases, and Q being the proportion of persons from the isolated population in controls. Five combinations of P and Q were considered. The first is a totally unmatched combination $P = 0.95$ and $Q = 0.05$. The second $P = 0.75$ and $Q = 0.25$, an extremely unmatched combination. The third, a severely unmatched combination $P = 0.50$ and $Q = 0.25$. The fourth combination is mildly unmatched $P = 0.50$ and $Q = 0.30$ and the fifth combination is slightly unmatched $P = 0.50$ and $Q = 0.40$.

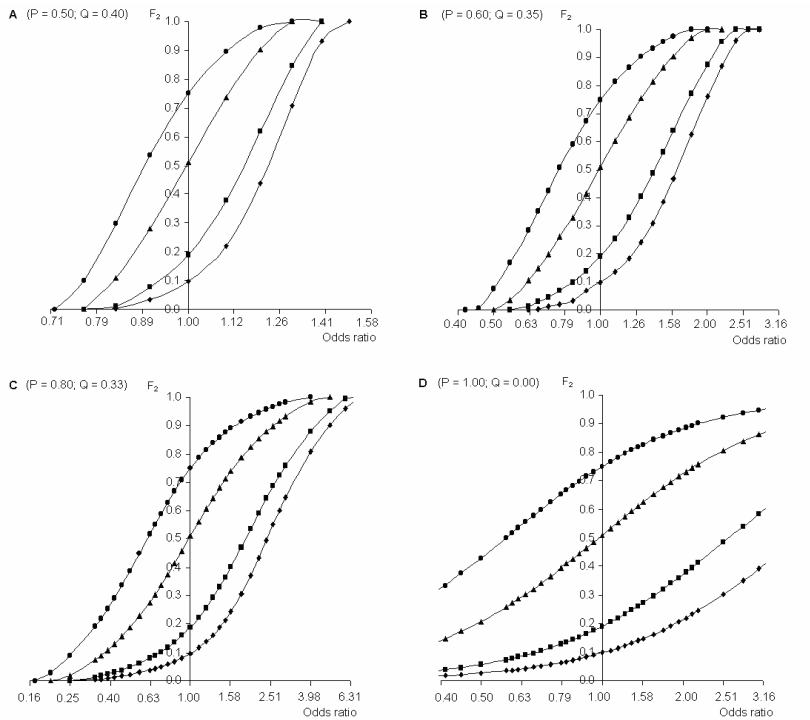
To analyze the effect of genetic drift and population stratification the odds ratio comparing the allele A carrier frequency in cases and controls, based on P, Q, F_{gen} and F_{iso} was calculated for all individual replicates. Since no association was simulated between the marker and the disease, the odds ratio is directly reflecting the effect of population stratification ($P \Leftrightarrow Q$) and genetic drift ($F_{gen} \Leftrightarrow F_{iso}$). From the individual odds ratios the cumulative probability distribution of the odds ratios was determined from the 25 000 replicates. This was repeated for the three separate conditions of genetic drift. These distributions were subsequently plotted in figures. For odds ratios smaller than one, the cumulative distribution was plotted, for odds ratios larger than one, 1-cumulative distribution was plotted. The figures directly illustrate the probability of confounding by population stratification and genetic drift. For the odds ratios of 1.50 and 1/1.50 = 0.67 the probabilities were summarized and for the odds ratios 1.20 and 0.83 as well.

Results

The results of the population stratification in relation to allele frequency differences, given various values of *P* and *Q* are presented in figure 1.

Figure 1

Extend of confounding due to allele frequency differences and stratification of two populations with various mixtures in cases and controls.



The effect of two-population stratification, given by the deviation of the odds ratio from one, is presented. *P* represents the proportion of cases and *Q* the proportion of controls from population 2. Four different mixtures of *P* and *Q* are presented in figures A to D. For a random SNP polymorphism in population 2, the carrier frequency F_2 of the first allele is plotted on the vertical axis. The four lines present the odds ratios given the allele carrier frequency F_1 of population 1 for the same polymorphism. The allele carrier frequency F_1 has the values (●) 0.75, (▲) 0.51, (■) 0.19 and (◆) 0.10.

Figure 1 shows that when both F_1 and F_2 differ, and *P* and *Q* as well, there always will be an effect of stratification i.e. the odds ratio is not equal to one.

In addition, it is important to note that population stratification does not only cause false-positive associations, but false-negative associations as well. What is considered to be substantial confounding is quite arbitrary and depends on the type of research and research question. Here, we use odds ratios of 1.50 and 0.67 as cut-off points for substantial confounding. If we assume that in a study 50% of the patients and 40% of the controls are from population 2 (figure 1A), then populations 1 and 2 need to differ strongly in allele frequencies in order to obtain an odds ratio of 1.50 (F_1 = 0.10, F_2 = 1.00). Obviously, the confounding of the relative risk becomes stronger when there is a larger difference in population selection between cases and controls (figure 1B – 1D). When P and Q differ considerably (P = 0.80, Q = 0.33) the frequency difference needs to be for example F_1 = 0.19 and F_2 = 0.37 to lead to an odds ratio of 1.50. In both cases, the allele frequency differences remain substantial and require a large diversity between the stratified populations.

In order to give an indication of the effects of genetic drift, the distribution of gene frequency differences between a general population and an isolated population was determined under various conditions. With these differences odds ratios were calculated to measure the effect of population stratification and genetic drift. The cumulative probability distributions of the odds ratios are given in figures 2 to 4 for mild - (figure 2), severe - (figure 3), and extreme genetic drift (figure 4). The odds ratio is plotted on the X-axis using a logarithmic scale. For odds ratios smaller than one, the Y-axis on the left part of the graphs shows the cumulative probability. For odds ratios larger than one, the Y-axis on the right part of the graphs shows the 1-cumulative probability. The total probability for finding a spurious association caused by stratification of the isolated and general population, can then be obtained by adding the two probabilities from the left and right graphs. Note that we only consider P larger than Q in these figures; the results for Q larger than P can be obtained by taking one divided by the odds ratio. This mirrors the left and right part of the graph.

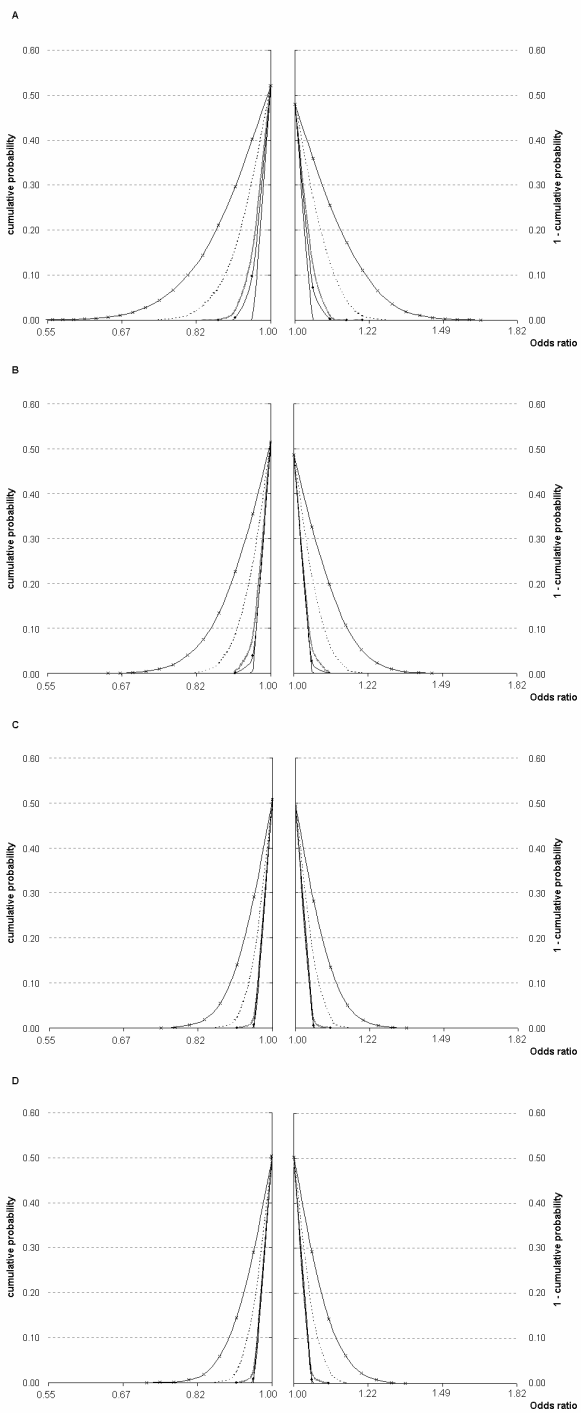
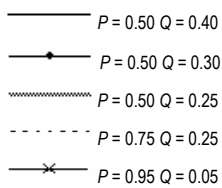


Figure 2

The effects of mild genetic drift (*condition 1* in text) and population stratification of an isolated and general population.

The figure is made out of four parts A to D that represent the four SNP allele *A* carrier frequencies F_{gen} of the general population (A) F_{gen} = 0.10, (B) F_{gen} = 0.19, (C) F_{gen} = 0.51, and (D) F_{gen} = 0.75. The frequency of the isolated population was subjected to genetic drift. With the frequencies of the population and the proportions *P* and *Q* of the isolated population in cases and controls, respectively, the odds ratio for an arbitrary case-control study was calculated. The deviation of the odds ratio from 1 represents the effect of population stratification. On the vertical axis the cumulative probability is plotted to find a given odds ratio in the simulations. This probability is divided in a part for odds ratios smaller than one, the left part of the graph, and also for odds ratios larger than one, the right part of the graph. For odds ratios larger than one, the 1 - cumulative probability represents the same probability for spurious association as the cumulative probability for odds ratios smaller than one. Different lines present the various mixtures of *P* and *Q*.



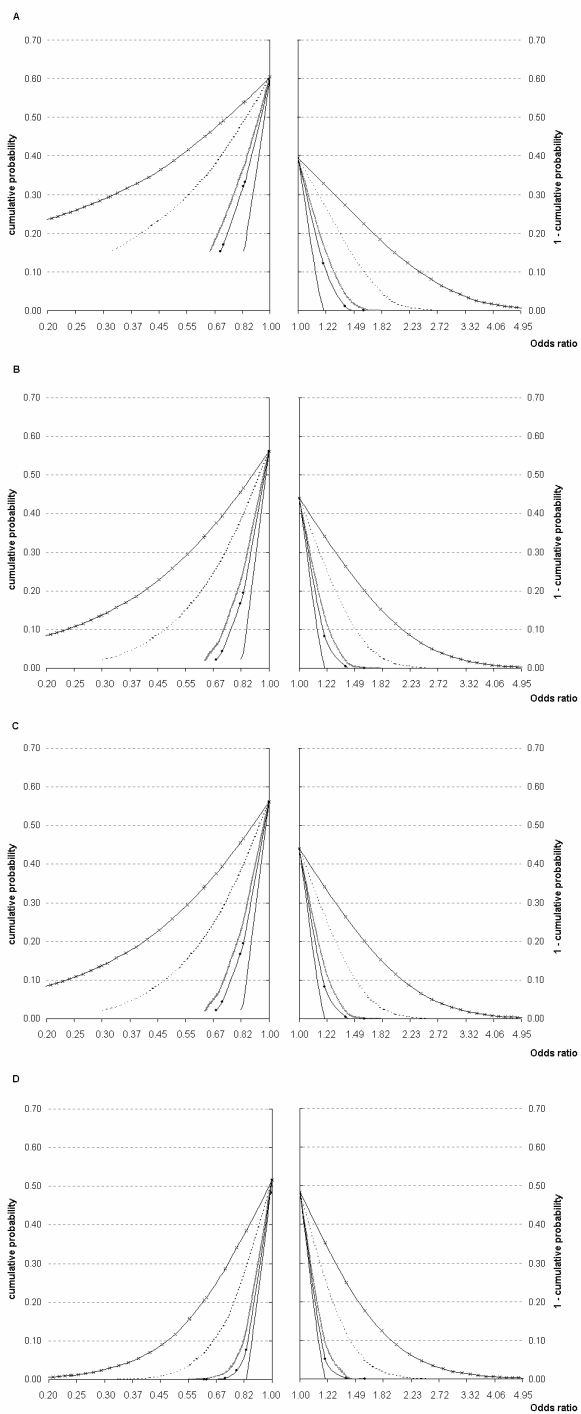
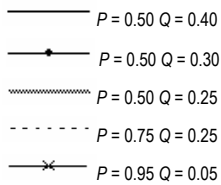


Figure 3

The effects of severe genetic drift (*condition 2* in text) and population stratification of an isolated and general population.

The figure is made out of four parts A to D that represent the four SNP allele *A* carrier frequencies F_{gen} of the general population (A) F_{gen} = 0.10, (B) F_{gen} = 0.19, (C) F_{gen} = 0.51, and (D) F_{gen} = 0.75. The frequency of the isolated population was subjected to genetic drift. With the frequencies of the population and the proportions *P* and *Q* of the isolated population in cases and controls, respectively, the odds ratio for an arbitrary case-control study was calculated. The deviation of the odds ratio from one represents the effect of population stratification. On the vertical axis the cumulative probability is plotted to find a given odds ratio in the simulations. This probability is divided in a part for odds ratios smaller than one, the left part of the graph, and also for odds ratios larger than one, the right part of the graph. For odds ratios larger than one, the 1 - cumulative probability represents the same probability for spurious association as the cumulative probability for odds ratios smaller than 1. Different lines present the various mixtures of *P* and *Q*.



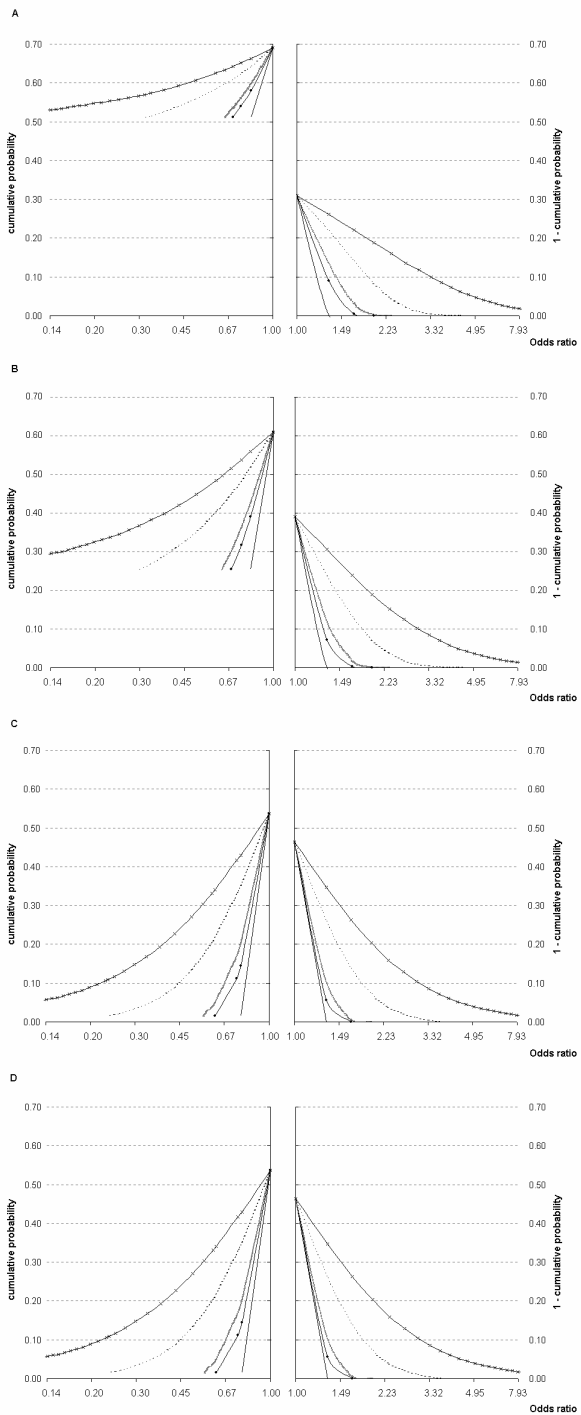
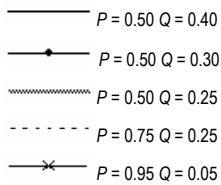


Figure 4

The effects of extreme genetic drift (*condition 3* in text) and population stratification of an isolated and general population.

The figure is made out of four parts A to D that represent the four SNP allele *A* carrier frequencies F_{gen} of the general population (A) F_{gen} = 0.10, (B) F_{gen} = 0.19, (C) F_{gen} = 0.51, and (D) F_{gen} = 0.75. The frequency of the isolated population was subjected to genetic drift. With the frequencies of the population and the proportions *P* and *Q* of the isolated population in cases and controls, respectively, the odds ratio for an arbitrary case-control study was calculated. The deviation of the odds ratio from one represents the effect of population stratification. On the vertical axis the cumulative probability is plotted to find a given odds ratio in the simulations. This probability is divided in a part for odds ratios smaller than one, the left part of the graph, and also for odds ratios larger than one, the right part of the graph. For odds ratios larger than one, the 1 - cumulative probability represents the same probability for spurious association as the cumulative probability for odds ratios smaller than one. Different lines present the various mixtures of *P* and *Q*.



The results in figure 2 show that under mild genetic drift (*condition 1*) the probabilities of finding odds ratios that are larger than 1.50 or lower than 0.67 are always less than 0.01. This holds for all allele *A* carrier frequencies 0.10, 0.19, 0.51 and 0.75 (figure 2A to D). Comparing figure 2A with 2D shows that the confounding effect of population stratification is smaller for common alleles as compared to rare alleles.

In figure 3 the genetic drift of the population isolate was much stronger (severe drift, *condition 2*). As a consequence the differences in allele frequencies between the isolate and the general population were often larger; this resulted in a larger effect of the population stratification as well. If we assume again odds ratios of 1.50 and 0.67 as cut-off points, then under mixtures of $P = 0.50$ $Q = 0.25$, $P = 0.50$ $Q = 0.30$ and $P = 0.50$ $Q = 0.40$ the probabilities of finding population stratification remain below 0.03. The exception is the rare allele *A* carrier frequency of 0.10, which has a higher probability of 0.18 (figure 3A). More extreme mixtures of $P = 0.75$ $Q = 0.25$ and $P = 0.95$ $Q = 0.05$ show that under these conditions of genetic drift the effects of population stratification and finding spurious associations can be substantial. In addition, the figure shows that for a low allele *A* carrier frequency in the general population a spurious inverse association (odds ratio < 1) may occur more frequently when $P > Q$.

In figure 4 the results are presented for an isolated population that is under extreme genetic drift (*condition 3*). In this situation the isolated population often reached the state that only one allele in the replicates was present. Therefore, the effect of population stratification is much stronger dependent on the mixtures of *P* and *Q* as the effect of the allele frequencies is maximal. The probabilities of finding spurious association are high for all situations, except for when the mixtures are not extremely different ($P = 0.50$, $Q = 0.40$). Again inverse spurious associations, in which the studied *A* allele protects for the

disorder, are more often present than associations in which the *A* allele is less frequent in the disorder given $P > Q$.

Table 1

Probabilities of finding spurious association in relation to simulated genetic drift and founder effects in an isolated population.

			P	0.95	0.75	0.50	0.50	0.50
			Q	0.05	0.25	0.25	0.30	0.40
F_{gen}	Isolated population	Odds ratio (OR)[a]						
0.100	Isolate 1	OR < 0.67 and OR > 1.50		0.014	0.000	0.000	0.000	0.000
		OR < 0.83 and OR > 1.20		0.256	0.043	0.000	0.000	0.000
	Isolate 2	OR < 0.67 and OR > 1.50		0.717	0.508	0.187	0.001	0.000
		OR < 0.83 and OR > 1.20		0.870	0.769	0.549	0.453	0.171
	Isolate 3	OR < 0.67 and OR > 1.50		0.878	0.780	0.591	0.027	0.000
		OR < 0.83 and OR > 1.20		0.945	0.900	0.802	0.754	0.553
0.190	Isolate 1	OR < 0.67 and OR > 1.50		0.001	0.000	0.000	0.000	0.000
		OR < 0.83 and OR > 1.20		0.128	0.006	0.000	0.000	0.000
	Isolate 2	OR < 0.67 and OR > 1.50		0.594	0.342	0.051	0.001	0.000
		OR < 0.83 and OR > 1.20		0.808	0.665	0.385	0.278	0.041
	Isolate 3	OR < 0.67 and OR > 1.50		0.779	0.615	0.336	0.018	0.000
		OR < 0.83 and OR > 1.20		0.902	0.823	0.648	0.571	0.306
0.510	Isolate 1	OR < 0.67 and OR > 1.50		0.000	0.000	0.000	0.000	0.000
		OR < 0.83 and OR > 1.20		0.035	0.000	0.000	0.000	0.000
	Isolate 2	OR < 0.67 and OR > 1.50		0.455	0.181	0.006	0.001	0.000
		OR < 0.83 and OR > 1.20		0.737	0.546	0.220	0.128	0.003
	Isolate 3	OR < 0.67 and OR > 1.50		0.670	0.447	0.114	0.049	0.000
		OR < 0.83 and OR > 1.20		0.846	0.729	0.487	0.385	0.086
0.750	Isolate 1	OR < 0.67 and OR > 1.50		0.000	0.000	0.000	0.000	0.000
		OR < 0.83 and OR > 1.20		0.042	0.001	0.000	0.000	0.000
	Isolate 2	OR < 0.67 and OR > 1.50		0.473	0.192	0.004	0.000	0.000
		OR < 0.83 and OR > 1.20		0.746	0.558	0.233	0.133	0.001
	Isolate 3	OR < 0.67 and OR > 1.50		0.673	0.445	0.096	0.024	0.000
		OR < 0.83 and OR > 1.20		0.845	0.732	0.487	0.381	0.066

P = the proportion of the isolated population in cases, *Q* = the proportion of the isolated population in controls, F_{gen} = the allele *A* carrier frequency of a random SNP in the general population. Isolate 1 has 3000 founders and was fully isolated for 10 generations. Isolate 2 has 100 founders and was fully isolated for 40 generations. Isolate 3 has 38 founders and was fully isolated for 52 generations. [a] The odds ratio reflecting the effect of population stratification, based on *P*, *Q*, F_{gen} and the allele *A* carrier frequency in the isolated population. Here the limits are shown for which the cumulative probability of finding the odds ratio was calculated in 25 000 replicates.

The current results are based on odds ratios, which are higher than 1.50 and lower than 0.67. For an *A* allele carrier frequency of 0.10, about 400 cases and controls would be needed to detect an odds ratio of 1.50 with a power of 80% using a significance level of 5%. Detailed probabilities of spurious associations given the conditions of drift, allele *A* frequencies and mixtures are presented in table 1. In the same table the probabilities for odds ratios 1.20 and 0.83 are given. It should be noted that when the odds ratio is chosen to be closer to one, for example 1.20, the probabilities of finding spurious association increase (table 1). Therefore, the effect size that is expected in the association between the disorder and the polymorphism is an important determinant for the importance of population stratification.

Discussion

Both the differences in population selection as well as allele frequencies of a given marker need to be large to cause considerable spurious associations (odds ratio = 1.50) as a result of population stratification. Large allele frequency differences can frequently be observed between strongly isolated populations subjected to a large amount of genetic drift. Together with a severely unmatched case-control sampling from a general and such isolated population false associations may frequently occur. However, for a mixture of two more similar populations, the probability of finding spurious association is much smaller even under extreme unmatched case-control sampling. Similarly, when cases and controls are more closely matched for their population background, the probability of spurious associations is small, even when genetic drift between the populations is strong.

If a studied allele of a polymorphism is rare, two mechanisms will result in an increased probability for spurious associations caused by stratification. The first mechanism is that a small change in allele frequency has a much larger effect on the odds ratio[17]. The second mechanism is the effect that genetic drift more frequently causes a substantial change in allele frequency, especially in

small populations. Furthermore, the rare allele more frequently disappears and the common allele becomes fixed[18]. This results in the situation that, when more cases are taken from the isolated population as compared to controls and the allele carrier frequency is below 0.75, spurious 'protective' associations will occur more frequently. This relation is opposite to the 'expected' relation in which an increased frequency of a rare genetic variant is resulting in a higher risk for disease. In addition, this mechanism may also counteract to a rare allele with an increased risk, and therefore, causing the loss of association[19].

In this study, we simulated one large unstructured population and three conditions for an isolated population subjected to genetic drift and founder effects. Note that our assumption, that the isolate is founded from the large population, is often made for founder effects[11,20]. In general, the situation may be more complex in that more subpopulations may exist that mix to some extend. Each of these populations may be subjected to genetic drift. However, it is shown that when stratifying multiple populations, spurious associations are less likely to occur[12]. The reason for this is that (random) differences in allele frequencies between several populations cause a regression to the mean effect of the overall frequency in cases and controls. In addition, the relative contribution of each subpopulation in both groups becomes less pronounced. Our approach of only two separate populations therefore seems a worst-case scenario. Furthermore, we assumed that there is no mixture and migration between the populations. In case of mixture between the populations, the effect of population stratification will be less pronounced as well, as the populations will be more alike[11].

Large differences in genetic make-up can be obtained by strong founder effects; isolation for a large number of generations and substantial genetic drift in a small population. It is unlikely that investigators do not detect the fact that cases and controls are extremely unmatched for their genetic background in

these cases. Our findings are in line with the empirical findings of Wacholder et al., who evaluated the differences in allele frequencies of the NAT2 gene across the world[12]. Wacholder et al., demonstrated that the risk ratio is expected to be biased by less than 10% in US studies. Other studies have also shown that for some other polymorphisms the differences between populations are not substantial as well[14,21]. Obviously, the effect of population stratification is dependent on the polymorphism that is tested. Current SNP genotyping efforts and comparison of allele frequencies over various populations may help to identify polymorphisms that have large frequency differences over subpopulations[22,23].

An important implication of the findings of Wacholder and the present study is that in the case of a single tested marker population stratification is not a major determinant of false-positive findings in association studies. If more polymorphisms, for example 10 000 or 50 000 SNPs, are tested the probability of finding spurious associations caused by population stratification increases. Usually a Bonferroni correction is applied to adjust the significance level (α) for multiple testing[24]. If a slight increase of the significance level is present because of population stratification, which is not taken into account in the correction, then a larger number of false-positive markers is expected when testing multiple markers. This effect will strongly increase if many markers are tested. Therefore, when testing multiple markers the Bonferroni correction or the general study design should take the population stratification effects into account.

As is the case with confounding in a non-genetic study, the easiest way to adjust the study data is to match cases and controls for their genetic background. The effect of confounding on the relative risk using this strategy is usually minimized[25]. An alternative for this approach is to collect data of the population background for cases and controls and adjust for this in the statistical analysis. A major problem is, that it is very tedious and expensive to

carefully collect this data for both patients and controls. The answers to questions about the ethnic background may be unreliable. Furthermore, because of mixture between populations, persons cannot be assigned to either population exclusively.

To avoid the problems with control selection, effective methods dealing with population stratification using family members of the patient have been developed[26,27]. An example of a different study method is the use of the patients' parental alleles as control persons[26]. In this test the comparison is made between the alleles that are, and are not transmitted to the patient. The non-transmitted alleles are used as the control. In this way the problem of unknown population stratification is solved elegantly because the artificial control with the two non-segregating alleles is from the same genetic background as the patient by definition. Similarly, the transmission disequilibrium test (TDT) can be used, where the transmission of the parental alleles is tested against the expected 50% Mendelian transmission[27]. A drawback of these methods is that the genotypes of the parents need to be known. For traits with a late age of onset the method will therefore be difficult to apply, as the parents usually are deceased[28]. In addition, the tested marker needs to be highly informative to obtain unambiguous information on the allele transmission. Other family-based tests were developed that use different relatives in order to obtain this information[29-35]. In addition, the family-based tests were extended to examine QTL loci and haplotypes[36,37]. Power studies indicate that, particularly when there is no strong population stratification and a small disease effect, family-based studies are less powerful than association studies[38,39]. In addition, the case-control study design is much more cost-effective as compared to the family-based tests[12,20].

Methods to test and correct for population stratification in case-control studies have been developed as well[20,40]. These tests are based on the principle that every marker in the genome will indicate the population diversity in case there

is population stratification. For example, not only the HLA-10 allele or "the chopsticks-gene" has different allele frequencies between Asians as compared to Caucasians, but markers related to eye development and appearance as well. By genotyping random markers it is therefore possible to examine how much cases and controls differ in genetic background. The testing of about 15-30 informative markers, not related to the disease and not related to each other, should be sufficient to detect population stratification[20]. The applied statistic is a summed χ^2 test for all the random marker associations tested between cases and controls. However, with this test it is not possible to correct the case-control sample for population stratification. Correcting for population stratification was initiated with the "GC" = Genomic Control test[40,41]. This test also uses unrelated markers scattered across the genome, but the statistic used to measure the association is weighted by a correction factor. The correction factor is based on the increase of variance in allele frequencies that is caused by population stratification (the Wahlund principle)[42]. The amount of population stratification is proportional to the increase of this factor, therefore, it can be used to correct for population stratification. The method requires around 50 markers to be tested and the estimates of association are generally conservative[8,43]. The method on the other hand is easy to implement in current statistical procedures. More complex latent-class models, or structure assessment tests in which the underlying population subgroups are defined based on the tested unlinked genetic markers have been developed recently [44-48]. The detected subgroups can accordingly be matched or weighted, and the association between the disorder and the marker of interest can be studied while correcting for population stratification. Practical applications of these tests are sparse and most results are based on simulated data, though some results seem promising[8,49,50]. Also there is no clear indication which test can be used optimally in certain case-control association studies[8].

Conclusions on population stratification

The relevance of population stratification for association studies is an ongoing point of discussion[8,13,51]. The effect of stratification depends on the population differences for the markers that are studied, the selection of cases and controls from the populations and the effects of the marker on the disorder as shown in this chapter. In association studies that are done with the often moderate size (N = 500-1000), searching for large effects (odds > 1.50), the stratification will be no serious threat if the studies are conducted following good epidemiological practice[12,52]. A recent summary of association studies shows that the expected effects of association are, however, generally smaller. Therefore, in larger sized studies (N = 10 000) searching for small effects, population stratification should preferably be taken into account by design or control. The stratification becomes more relevant when many markers are studied. In this case, as for example in an association genome scan, stratification can be tested and controlled using the summarized methods[20,40,44-48].

References

1. Horikawa Y, Oda N, Cox NJ et al. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet. 2000;26:163-75.
2. Cox NJ. Challenges in identifying genetic variation affecting susceptibility to type 2 diabetes: examples from studies of the calpain-10 gene. Hum Mol Genet. 2001;10:2301-5.
3. Baron M. The search for complex disease genes: fault by linkage or fault by association? Mol Psychiatry. 2001;6:143-9.
4. Little J, Bradley L, Bray MS et al. Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. Am J Epidemiol. 2002;15;156:300-10.
5. Ioannidis JP. Genetic associations: false or true? Trends Mol Med. 2003;9:135-8.
6. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. Lancet. 2003;8;361:865-72.
7. Montgomery H, Dansek AH. In search of genetic precision. Lancet. 2003;31;361:1909.
8. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet. 2003;15;361:598-604.
9. Lander ES, Schork NJ. Genetic dissection of complex traits. Science. 1994;265:2037-48.
10. Knowler WC, Williams RC, Pettitt DJ, Steinberg AG. Gm3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. Am J Hum Genet. 1988;43:520-6.

11. Cavalli-Sforza LL, Bodmer WF. The genetics of human populations. 1 ed. Books in biology. Ed. Kennedy D, Park RB. San Francisco: W.H. Freeman and company. 1971.
12. Wacholder S, Rothman N, Caporaso N. Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. J Natl Cancer Inst. 2000;19;92:1151-8.
13. Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. Cancer Epidemiol Biomarkers Prev. 2002;11:513-20.
14. Ardlie KG, Lunetta KL, Seielstad M. Testing for population subdivision and association in four case-control studies. Am J Hum Genet. 2002;71:304-11.
15. Rothman J, Greenland S. Modern epidemiology. 2nd edition. Lippincott Williams & Wilkins. Philadelphia. 1998. ISBN 0316757764.
16. Risch NJ. Searching for genetic determinants in the new millennium. Nature. 2000;405:847-56.
17. Neuhauser C. Mathematical models in population genetics. In Handbook of statistical genetics. Ed. Balding D, Bishop M, Cannings C. John Wiley (press). 2001.
18. Falconer DS, Trudt FC. Introduction to quantitative genetics – chapter 2. Addison-Wesley Pub Co; 4th edition. ISBN: 0582243025. 1996.
19. Deng HW. Population admixture may appear to mask, change or reverse genetic effects of genes underlying complex traits. Genetics. 2001;159:1319-23.
20. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet. 1999;65:220-8.
21. Hutchison KE, Stallings M, McGeary J, Bryan A. Population stratification in the candidate gene study: fatal threat or red herring? Psychol Bull. 2004;130:66-79.
22. Sachidanandam R, Weissman D, Schmidt SC et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature. 2001;15;409:928-33.
23. Collins FS, Guyer MS, Charkravarti A. Variations on a theme: cataloging human DNA sequence variation. Science. 1997;28;278:1580-1.
24. Bonferroni, CE. Teoria statistica delle classi e calcolo delle probabilit `a. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8, 3-62. 1936.
25. Vandenbroucke JP, Hofman A. Grondslagen der epidemiologie. 6 dr. Maarssen: Elsevier. p. 262. 1999.
26. Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. Ann Hum Genet. 1987;51:227-33.
27. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet. 1993;52:506-16.
28. Spielman RS, Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet. 1996;59:983-9.
29. Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. Am J Hum Genet. 1998;62:450-8.
30. Knapp M. The transmission/disequilibrium test and parental-genotype reconstruction: the reconstruction-combined transmission/ disequilibrium test. Am J Hum Genet. 1999;64:861-70.
31. Horvath S, Laird NM, Knapp M. The transmission/disequilibrium test and parental-genotype reconstruction for X-chromosomal markers. Am J Hum Genet. 2000;66:1161-7.

32. Monks SA, Kaplan NL. Removing the sampling restrictions from family-based tests of association for a quantitative-trait locus. Am J Hum Genet. 2000;66:576-92.
33. Lange C, DeMeo D, Silverman EK et al. Using the noninformative families in family-based association tests: a powerful new testing strategy. Am J Hum Genet. 2003;73:801-11.
34. Rabinowitz D, Laird N. A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. Hum Hered. 2002;50:211-23.
35. Cordell HJ, Barratt BJ, Clayton DG. Case/pseudocontrol analysis in genetic association studies: A unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. Genet Epidemiol. 2004;26:167-85.
36. Gauderman WJ. Candidate gene association analysis for a quantitative trait, using parent-offspring trios. Genet Epidemiol. 2003;25:327-38.
37. Allison DB. Transmission-disequilibrium tests for quantitative traits. Am J Hum Genet. 1997;60:676-90.
38. McGinnis R, Shifman S, Darvasi A. Power and efficiency of the TDT and case-control design for association scans. Behav Genet. 2002;32:135-44.
39. Jorde LB. Linkage disequilibrium and the search for complex disease genes. Genome Res. 2000;10:1435-44.
40. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999;55:997-1004.
41. Bacanu SA, Devlin B, Roeder K. Association studies for quantitative traits in structured populations. Genet Epidemiol. 2002;22:78-93.
42. Hedrick PW. Population genetics. The wahlund principle p.284-85. Jones and Bartlett publishers, Boston. 1985.
43. Bacanu SA, Devlin B, Roeder K. The power of genomic control. Am J Hum Genet. 2000;66:1933-44.
44. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945-59.
45. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. Am J Hum Genet. 2000;67:170-81.
46. Satten GA, Flanders WD, Yang Q. Accounting for unmeasured population substructure in case-control studies of genetic association using a novel latent-class model. Am J Hum Genet. 2001;68:466-77.
47. Ripatti S, Pitkaniemi J, Sillanpaa MJ. Joint modeling of genetic association and population stratification using latent-class models. Genet Epidemiol. 2001; 21 Suppl 1:S409-14.
48. Hoggart CJ, Parra EJ, Shriver MD et al. Control of confounding of genetic associations in stratified populations. Am J Hum Genet. 2003;72:1492-504.
49. Pritchard JK. Deconstructing maize population structure. Nat Genet. 2001;28:203-4.
50. Pfaff CL. Adjusting for population structure in admixed populations. Genetic Epidemiology. 2002;1. 22:196–201.
51. Thomas DC, Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene associations? Cancer Epidemiol Biomarkers Prev. 2002;11:505-12.
52. Ioannidis JP, Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG. Genetic associations in large versus small studies: an empirical assessment. Lancet. 2003;15;361:567-71.

CHAPTER 3

A straightforward approach to overcome false-positive associations in studies of gene-gene interaction

JJ Hottenga[1,2], G Roks[1], B Dermaut[3], C van Broeckhoven[3], CM van Duijn[1,3]

1. Department of Epidemiology & Biostatistics, Erasmus Medical Center Rotterdam, the Netherlands.
2. Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.
3. Department of Molecular Genetics, Flanders Interuniversity Institute for Biotechnology, University of Antwerp, Antwerp, Belgium.

For submission

Abstract

Research of gene-gene interactions will be important in unraveling genetic risk factors involved in complex traits. However, based on association studies, gene interactions are susceptible to false-positive as well as false-negative findings. One of the reasons may be stratification of a limited number of cases and controls, leading to a small number of subjects in each stratum. In particular the number of controls who carry the risk allele of both genes studied is often small, even for common polymorphisms. We present a straightforward approach to reveal spurious associations due to over-stratification that can be applied in gene-gene interaction studies. Basically, we propose to analyze cases and controls separately. In controls, association between two unlinked genes will indicate bias in the findings of the study. From this approach it also follows that one may improve the statistical power and reduce the probability of false-positive findings by genotyping controls for the second gene in the limiting stratum specifically, i.e. control carriers of the risk allele of the first gene studied. This approach may be useful in large-scale epidemiological studies, in which multiple genes often have been characterized. We will illustrate the approach using an example of a study of interaction of the Apolipoprotein E and Presenilin-1 gene in relation to Alzheimer's disease. In this study, a false-positive association was detected using this method.

Keywords

Gene-gene interaction, Apolipoproteins E, Presenilin-1, Alzheimer's disease, interaction test.

Introduction

Candidate gene research has shown to be prone to false-positive findings, despite its potential for studies of complex genetic disorders[1,2]. Typically these studies target polymorphisms in genes, which in effect may depend, for a large part, on other genes. Studies of gene-gene interactions are therefore common practice in genetic epidemiological research. To this end, the effect of a polymorphism is often studied in a stratified analysis based on the presence of a second allele of a polymorphism. In this paper we argue that as a result of the stratification, false-positive findings may occur because of too small numbers in the stratified groups. A simple straightforward approach is presented to evaluate this form of bias. Furthermore, the statistical power when using this approach was examined and the increase shows that it may also prevent false-negative studies. To illustrate the method, a study of interaction between the Apolipoprotein E (APOE) and Presenilin-1 (PSEN1) genes in Alzheimer's disease (AD) is presented. The APOE ε4 allele (APOE*4) has shown to be an established risk factor for AD[3]. The PSEN1 gene can be mutated in patients with early-onset familial AD[4]. Findings on polymorphisms in the PSEN1 gene in relation to late-onset sporadic AD have been inconsistent[5–8]. In this study an interaction was found between APOE*4, PSEN1 and AD. However, when evaluated with the approach, the result was shown to be a false-positive.

Methods

Approach

In studies of gene-gene interaction, polymorphisms of two genes are studied for a specific disorder. Data of the two polymorphisms can be analyzed under several interaction models[9,10]. Table 1A illustrates an often used approach, in which the data are stratified for the presence of allele *X* for gene 1 and allele *Y* for gene 2. If there is interaction present between the genes, the odds ratio for gene 2 will be different for the strata of gene 1. These odds ratios are given as (*A x D*) / (*B x C*) and (*E x H*) / (*G x F*). If the studied genes are unlinked, for

example because they are located on different chromosomes, then according to Mendelian laws the two genes will segregate independently from each other. For most combinations of genes that are studied in interaction studies this is true. Consequently, the alleles of the two genes, when rewriting table 1A into 1B stratifying by case-control status, should not show association in controls (table 1B)[11,12]. A significant deviation of the odds ratio from one in controls is not compatible with Mendelian segregation of alleles. There may be some explanations, for example mixture of genetically different populations. However, a more likely explanation to consider is that the finding is a false-positive result due to the stratification of the controls into small subgroups. *A priori* one expects to find a low number of controls in the individual strata, especially in the cell having the two risk alleles associated with the studied disorder (cell *G* in tables 1A and 1B). This problem can be overcome by increasing the number of controls in the study[13,14]. Controls can be added at random enlarging the total number of controls. However, in large-scale epidemiological studies in which several genes have already been examined, controls can also be added to specific strata for which there are specifically unstable low numbers. This is illustrated in table 1C, in which the control series in the stratum of allele *X* carriers of gene 1 is increased by a factor *K*, yielding an odds ratio of $(C \times KH) / (D \times KG) = (C \times H) / (D \times G)$ similar to the one found in table 1B. Therefore, with either method of adding controls the new odds ratio will remain unbiased, with the second approach being a more cost effective alternative. The use of this approach was illustrated in an example of AD.

Example

In this example, the empirical data comprise a series of patients with late-onset Alzheimer's disease, which were drawn from the Rotterdam study[15]. This is a population-based prospective study of over 8000 residents aged 55 years and older of a suburb in Rotterdam in the Netherlands. All participants in the study gave informed consent.

Table 1

Odds ratios in a case- control study of gene-gene interaction.

A

Gene 2		Gene 1 allele *X*+	allele *X*-	Odds ratio
allele *Y*-	cases	*A*	*B*	*(A x D) / (B x C)*
	controls	*C*	*D*	
allele *Y*+	cases	*E*	*F*	*(E x H) / (F x G)*
	controls	*G*	*H*	

B

	Gene 2	Gene 1 allele *X*+	allele *X*-	Odds ratio
cases	allele *Y*-	*A*	*B*	*(A x F) / (B x E)*
	allele *Y*+	*E*	*F*	
controls	allele *Y*-	*C*	*D*	*(C x H) / (D x G)*
	allele *Y*+	*G*	*H*	

C

	Gene 2	Gene 1 allele *X*+	allele *X*-	Odds ratio
cases	allele *Y*-	*A*	*B*	*(A x F) / (B x E)*
	allele *Y*+	*E*	*F*	
controls	allele *Y*-	*C*	*D*	*(C x KH) / (D x KG)*
	allele *Y*+	*KG*	*KH*	

Table 1A is stratified for the presence of allele *X*, then for cases and controls.
Table 1B is stratified for cases and controls, then for the presence of allele *X*.
Table 1C the number of controls in the allele *X*+ stratum is increased by factor *K*.

The –22 C/T PSEN1 polymorphism located in the promoter region of the gene was genotyped in a nested case- control study including 316 AD patients and 219 age-matched controls[16,17]. Because APOE is a pivotal gene involved in lipid metabolism and neurodegenerative disorders, all of the participants of the Rotterdam study have been genotyped for the APOE gene previously for other studies[18]. No linkage disequilibrium is expected between the two genes, since the PSEN1 gene is located on chromosome 14q24.2 and the APOE gene on chromosome 19q13.31 [4,19]. The number of controls carrying APOE*4 was subsequently increased by a factor K = 1.96 by genotyping 54 additional

controls for the PSEN1 gene following our approach. Odds ratios were calculated for the strata of the APOE*4 allele and PSEN1 genotypes and interactions were tested using binary logistic regression. For the statistical analysis, CT and TT genotypes of the PSEN1 gene were pooled because of the low number of persons carrying the TT genotype[20]. All statistics were tested with SPSS for Windows 10.0.

Power

Power calculations were performed to illustrate the effects of adding controls in a single stratum. Hence, the number of controls with genetic risk factor *X* was increased for values of *K* ranging from 1 to 4 (table 1C). Increasing *K* more than four times the number of controls usually does not increase the power further[13,14]. The power to detect association for allele *Y* given *X* was calculated for 250 and 500 cases and controls, respectively. Allele frequencies of *X* and *Y* were varied with values of 0.10, 0.25 and 0.50. The relative risk of the disorder given *X* was assumed to be 2.00 and the additional disorder risk for *Y* given presence of *X* was 2.00 as well. The significance level α was set to 0.05 and the prevalence of the disorder was 0.10. No genetic model was presumed for both polymorphisms and analysis was assumed to be done with allelic tests of association. The power was calculated with the web-based program PAWE 1.2[21,22].

Results

Gene-gene interaction was studied between the APOE ε4 allele and the –22 C/T PSEN1 promoter polymorphism in relation to Alzheimer's disease. For both genes, no significant deviations from Hardy-Weinberg equilibrium were observed in cases and controls. Table 2A shows the association between the PSEN1 genotype and AD, stratified by the presence of the APOE*4 allele. The results show that the number of CC carriers in AD patients is reduced as compared to the controls in the stratum of APOE*4 carriers. The odds ratio

(OR) for carriers of the CT/TT genotype equaled 0.24 with a 95% confidence interval (CI) of 0.07 to 0.84. In subjects who did not carry the APOE*4 there was no association between PSEN1 and AD. The OR, pooling the CT and TT genotype, equals 1.14 (95% CI = 0.66-1.98). Testing for interaction between PSEN1 and APOE showed evidence for interaction using a multiplicative model (p = 0.025).

Table 2

A gene-gene interaction study of the Apolipoprotein E ε4 allele (APOE*4) and the –22 C/T Presenilin-1 promoter polymorphism in relation to Alzheimer's disease (AD).

A

APOE		PSEN1 CC	CT+TT	Number	Odds ratio	95% CI
APOE*4-	AD	169 (84%)	32 (16%)	201	1.14	0.66-1.98
	controls	134 (82%)	29 (18%)	163		
APOE*4+	AD	93 (81%)	22 (19%)	115	0.24	0.07-0.84
	controls	53 (95%)	3 (5%)	56		

B

	APOE	PSEN1 CC	CT+TT	Number	Odds ratio	95% CI
AD	APOE*4-	169 (84%)	32 (16%)	201	1.25	0.69-2.27
	APOE*4+	93 (81%)	22 (19%)	115		
Controls	APOE*4-	134 (82%)	29 (18%)	163	0.26	0.08-0.90
	APOE*4+	53 (95%)	3 (5%)	56		

APOE*4+ and APOE*4- = presence and absence of the Apolipoprotein E ε4 allele,
PSEN1 = the Presenilin-1 genotype, 95% CI = The 95% confidence interval of the odds ratio.
Table 2A is stratified for the presence of the APOE*4 allele, then for cases and controls.
Table 2B is stratified for cases and controls, then for the presence of the APOE*4 allele.

When examining the association between APOE and PSEN1 for cases and controls separately, evidence for association between the two genes was found only in the control group (OR = 0.26; 95% CI = 0.08-0.90) (table 2B). No evidence for association was found in cases (OR = 1.25; 95% CI = 0.69-2.27). As the APOE and PSEN1 genes are located on different chromosomes, the

association in controls is not compatible with Mendelian segregation. Therefore, a true interaction effect of the two genes is unlikely. In addition, differential survival related to the genotypes is unlikely, because it is assumed to occur both in cases and age-matched controls. The interaction observed in table 2A is most likely the result of the low number of controls carrying both the APOE*4 allele and the T allele at the PSEN1 promoter.

Table 3

A reanalysis of the gene-gene interaction study of the Apolipoprotein E ε4 allele (APOE*4) and the –22 C/T Presenilin-1 promoter polymorphism in relation to Alzheimer's disease (AD) with added controls.

		PSEN1				
	APOE	CC	CT+TT	Number	Odds ratio	95% CI
AD	APOE*4-	169 (84%)	32 (16%)	201	1.25	0.69-2.27
	APOE*4+	93 (81%)	22 (19%)	115		
Controls	APOE*4-	134 (82%)	29 (18%)	163	0.62	0.31-1.25
	APOE*4+	97 (88%)	13 (12%)	110		

APOE*4+ and APOE*4- = presence and absence of the Apolipoprotein E ε4 allele, PSEN1 = the Presenilin-1 genotype. 95% CI = The 95% confidence interval of the odds ratio.

As all participants of the Rotterdam study were already genotyped for the APOE gene, we genotyped 54 extra controls for PSEN1 from the non-demented subjects carrying the APOE*4 allele, increasing the sample with factor $K = 1.96$. When adding these controls to the initial set, the frequency of the CC genotype decreased from 95% to 88%, while the number CT genotype carriers increased from 5% to 12% (table 3). The odds ratio analyzed in controls only, changed from 0.26 to 0.62 with the 95% confidence interval ranging from 0.31 to 1.25. This suggests that adding the controls overcame the problem presented in table 2B (table 3). As a result the PSEN1 genotype frequencies did not longer differ between cases and controls in the APOE*4 carrier stratum (OR = 0.57; 95% CI = 0.27-1.19). Testing for interaction using

logistic regression analysis showed no more evidence for interaction (p = 0.137).

Table 4

Power calculations for increasing the number of controls in a specific stratum of a gene-gene interaction case-control study.

F(*X*+)	N cases / controls[a]	N controls Random[b]	*K*	N cases *X*+[c]	N Controls *X*+[d]	Power F(*Y*+) = 0.10	Power F(*Y*+) = 0.25	Power F(*Y*+) = 0.50
0.10	250 / 250	250	1	45	23	0.286	0.507	0.527
	250 / 273	500	2	45	46	0.428	0.695	0.714
	250 / 296	750	3	45	69	0.518	0.784	0.785
	250 / 319	1250	4	45	92	0.583	0.816	0.823
	500 / 500	500	1	91	45	0.498	0.793	0.814
	500 / 545	1000	2	91	90	0.707	0.936	0.946
	500 / 590	1500	3	91	135	0.807	0.972	0.973
	500 / 635	2500	4	91	180	0.864	0.981	0.983
0.25	250 / 250	250	1	100	58	0.586	0.869	0.881
	250 / 308	500	2	100	116	0.788	0.968	0.972
	250 / 366	750	3	100	174	0.871	0.988	0.988
	250 / 424	1250	4	100	232	0.914	0.992	0.992
	500 / 500	500	1	200	117	0.871	0.992	0.994
	500 / 617	1000	2	200	234	0.974	1.000	1.000
	500 / 734	1500	3	200	351	0.992	1.000	1.000
	500 / 851	2500	4	200	468	0.997	1.000	1.000
0.50	250 / 250	250	1	167	120	0.861	0.990	0.991
	250 / 370	500	2	167	240	0.967	0.999	0.999
	250 / 490	750	3	167	360	0.988	1.000	1.000
	250 / 610	1250	4	167	480	0.994	1.000	1.000
	500 / 500	500	1	333	240	0.990	1.000	1.000
	500 / 740	1000	2	333	480	1.000	1.000	1.000
	500 / 980	1500	3	333	720	1.000	1.000	1.000
	500 / 1220	2500	4	333	960	1.000	1.000	1.000

K multiplication factor for the number of controls tested, F(*X*+) = frequency of allele *X*,

F(*Y*+) = frequency of allele *Y*.

[a] Total number of cases and controls in the study.

[b] The required number of random controls that need to be tested.

[c] The expected number of cases that have *X*+ in the study.

[d] The expected number of controls that have *X*+ in the study.

To examine the effect of adding controls in a specific stratum in gene-gene interaction studies power calculations were done (table 4). The results indicate that when the risk alleles *X* and *Y* are common (frequency > 0.25), typing additional controls increases the power. This is however not required, as the power without the additional controls is already substantial (>0.869). When one of the risk alleles, *X* and / or *Y*, is more rare (frequency = 0.10), a large increase in power can be obtained by genotyping only a limited number of additional controls. For most cases $K = 2$ is sufficient, unless both risk alleles have a frequency of 0.10. In this case *K* needs to be 3 or 4 and a larger sample of cases and controls needs to be analyzed.

When random controls are added instead of controls carrying one risk allele, the required number of additional tested samples is much higher (table 4). In this case the initial number of controls is multiplied by *K*. The genotyping of only controls which carry the risk allele of the first tested polymorphism (allele *X* in table 4) is consequently more cost efficient.

Discussion

Studying gene-gene interaction is likely the next step in unraveling complex genetic traits. However, very large population-based samples are required that are not always available[23,24]. Here, we show a simple and direct approach to evaluate false-positive findings in gene-gene interaction studies with an example of Alzheimer's disease, PSEN-1 and APOE[5-8]. The first step of the approach is to evaluate cases and controls separately to identify the group(s) in which an interaction is found. Basically, we assume that in gene-gene interaction studies the interaction is explained by association in cases and not in controls. Studying cases only to test for gene-gene interaction has been proposed but often controls are included to allow for estimations of risk[11,12]. If an association between two unlinked genes is found in controls, while cases do not show such a relation, it may be the result of over-stratification of the data.

In this case the second step of the approach, adding controls to the study can be applied.

It should be noted that other reasons for association in controls are not excluded in this manner. Of course, linkage disequilibrium, which may lead to association between the two genes, needs to be excluded *a priori*. Another reason for association may be differential mortality related to the risk alleles. Both mechanisms will have effect in cases and (age-matched) controls, leading to association in both series. No indication for population stratification was found in the Rotterdam study. However, when working in a more mixed population, for example the US population, one may want to consider testing for hidden population stratification first[25-28]. Adding controls may not overcome this problem, nor may it eliminate all variation in the risk estimate for controls, as it was the case in our example of AD.

Extra controls can be added genotyping a random control group, or samples of a specific stratum. In large epidemiological studies where specific genes have been tested already, the addition of controls in a stratum is a more cost-effective alternative. The efficiency of this approach is particularly high when rare risk alleles are studied. The addition of extra controls also has another advantage, namely that the detection power of interactions increases. Therefore, in addition to obtaining a more accurate risk estimate for the control group carrying both risk alleles, the approach also decreases the false-negative rate of the study. Future gene-gene interaction studies intending to find moderate risk factors for complex diseases that require large numbers of genotyped and phenotyped individuals may therefore benefit from this approach.

Acknowledgements

This project was supported by research grants from the Netherlands Organization of Scientific Research (NWO), the Special Research Fund of the

University of Antwerp, the Fund for Scientific Research Flanders (FWO). Bart Dermout is a PhD fellow of the FWO, the Medical Foundation Queen Elisabeth, the Interuniversity Attraction Poles (IUAP) program P5/19 of the Federal Office of Scientific, Technical and Cultural Affairs (OSTC), and the International Alzheimer Research Foundation (IARF), Belgium.

References

1. Baron M. The search for complex disease genes: fault by linkage or fault by association? Mol Psychiatry. 2001;6:143-9.
2. Cardon LR, Bell JI. Association study designs for complex diseases. Nat Rev Genet. 2001;2:91-9.
3. Farrer LA, Cupples LA, Haines JL et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer's disease. A meta-analysis. APOE and Alzheimer's Disease Meta Analysis Consortium. JAMA. 1997;278:1349-56.
4. Sherrington R, Rogaev EI, Liang Y et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature. 1995;375:754-60.
5. Wragg M, Hutton M, Talbot C. Genetic association between intronic polymorphism in presenilin-1 gene and late-onset Alzheimer's disease. Alzheimer's Disease Collaborative Group. Lancet. 1996;347:509-12.
6. van Duijn CM, Cruts M, Theuns J et al. Genetic association of the presenilin-1 regulatory region with early-onset Alzheimer's disease in a population-based sample. Eur J Hum Genet. 1999;7:801-6.
7. Taddei K, Yang D, Fisher C et al. No association of presenilin-1 intronic polymorphism and Alzheimer's disease in Australia. Neurosci Lett. 1998;246:178-80.
8. Bagli M, Papassotiropoulos A, Schwab SG et al. No association between an intronic polymorphism in the presenilin-1 gene and Alzheimer's disease in a German population. J Neurol Sci. 1999;167:34-6.
9. Goldstein AM, Falk RT, Korczak JF, Lubin JH. Detecting gene-environment interactions using a case-control design. Genet Epidemiol. 1997;14:1085-9.
10. Goldstein AM, Andrieu N. Detection of interaction involving identified genes: available study designs. J Natl Cancer Inst Monogr. 1999;26:49-54.
11. Yang Q, Khoury MJ, Sun F, Flanders WD. Case-only design to measure gene-gene interaction. Epidemiology. 1999;10:167-70.
12. Brennan P. Gene-environment interaction and aetiology of cancer: what does it mean and how can we measure it? Carcinogenesis. 2002;23:381-7.
13. Vandenbroucke JP, Hofman A. Grondslagen der epidemiologie. 6 dr. Maarssen: Elsevier. p. 131. 1999.
14. Miettinen OS. Individual matching with multiple controls in the case of all-or-none response. Biometrics. 1969;22:339-55.
15. Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol. 1991;7:403-22.
16. Dermaut B, Roks G, Theuns J et al. Variable expression of presenilin 1 is not a major determinant of risk for late-onset Alzheimer's disease. J Neurol. 2001;248:935-9.

17. Theuns J, Remacle J, Killick R et al. Alzheimer-associated C allele of the promoter polymorphism -22C>T causes a critical neuron-specific decrease of presenilin 1 expression. Hum Mol Genet. 2003;12:869-77.
18. Slooter AJ, Bots ML, Havekes LM et al. apolipoprotein E and carotid artery atherosclerosis: the Rotterdam study. Stroke. 2001;32:1947-52.
19. Das HK, McPherson J, Bruns GA et al. Isolation, characterization, and mapping to chromosome 19 of the human apolipoprotein E gene. J Biol Chem. 1985;260:6240-7.
20. Lambert JC, Mann DM, Harris JM et al. The -48 C/T polymorphism in the presenilin 1 promoter is associated with an increased risk of developing Alzheimer's disease and an increased Abeta load in brain. J Med Genet. 2001;38:353-5.
21. Gordon D, Finch SJ, Nothnagel M, Ott J. Power and sample size calculations for case-control genetic association tests when errors present: application to single nucleotide polymorphisms. Human Heredity 2002;54:22-33
22. Gordon D, Levenstien MA, Finch SJ, Ott J. Errors and linkage disequilibrium interact multiplicatively when computing sample sizes for genetic case-control association studies. Pacific Symposium on Biocomputing: 2003;490-501.
23. Gauderman WJ. Sample size requirements for association studies of gene-gene interaction. Am J Epidemiol. 2002;155:478-84.
24. Yang Q, Khoury MJ, Flanders WD. Sample size requirements in case-only designs to detect gene-environment interaction. Am J Epidemiol. 1997;146:713-20.
25. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet. 1999;65:220-8.
26. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. Am J Hum Genet. 2000;67:170-81.
27. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999;55:997-1004.
28. Hoggart CJ, Parra EJ, Shriver MD et al. Control of confounding of genetic associations in stratified populations. Am J Hum Genet. 2003;72:1492-504.

CHAPTER 4

The 3p21.1-p21.3 hereditary vascular retinopathy locus increases the risk for Raynaud phenomenon and migraine

JJ Hottenga[1,2], KRJ Vanmolkot[1], EE Kors[3], S Kheradmand Kia[1], PTVM de Jong[4,5,6] J Haan[3,7], GM Terwindt[3], RR Frants[1], MD Ferrari[3], AMJM van den Maagdenberg[1,3]

1. Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.
2. Department of Medical Statistics, Leiden University Medical Center, Leiden, the Netherlands.
3. Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands.
4. Netherlands Ophthalmic Research Institute, Amsterdam, the Netherlands.
5. Department of Ophthalmology, Academic Medical Center, Amsterdam, the Netherlands.
6. Department of Epidemiology and Biostatistics, Erasmus Medical Center, Rotterdam, the Netherlands.
7. Department of Neurology, Rijnland Hospital, Leiderdorp, the Netherlands.

Accepted by Cephalalgia

Abstract

Previously, we described a large Dutch family with hereditary vascular retinopathy (HVR), Raynaud phenomenon, and migraine. A locus for HVR was mapped on chromosome 3p21.1-p21.3 but the gene has not yet been identified. The fact that all three disorders share a vascular aetiology prompted us to study whether the HVR haplotype also contributed to Raynaud phenomenon and migraine in this family. Despite of the low-powered parent-child transmission disequilibrium tests showing no significant association, the sibling transmission disequilibrium tests revealed that the HVR haplotype harbors a susceptibility factor for Raynaud phenomenon and migraine. Identification of the HVR gene may therefore improve the understanding of the pathophysiology of HVR, Raynaud phenomenon and migraine.

Keywords

Hereditary vascular retinopathy, migraine, Raynaud phenomenon, locus.

Introduction

Migraine is a common neurovascular headache disorder affecting up to 18% of the general population[1,2]. The aetiology of migraine is complex, with probably various environmental and susceptibility genes involved[3,4]. Migraine can also be a part of autosomal dominant cerebrovascular syndromes, such as CADASIL and hereditary vascular retinopathy (HVR)[5-7]. Identification of genes involved in these cerebrovascular syndromes may therefore contribute to the understanding of migraine pathophysiology as well. Recently, HVR was mapped to chromosome 3p21.1-p21.3 in a large Dutch family[8]. Retinopathy in this family is characterized by microangiopathy of the retina, accompanied by micro-aneurysms and telangiectatic capillaries that appear preferentially around the macula and the posterior pole[9]. In later stages, capillary occlusions may develop leading to retinal ischemia and neovascularisation. Leukoencephalopathy was seen on MRI scans in some patients[7]. Genetic testing revealed that two additional families with autosomal dominant cerebroretinal vasculopathy (CRV) and hereditary endotheliopathy with retinopathy, nephropathy and stroke (HERNS), respectively, were also linked to this locus[10,11]. Despite the fact that all three families were linked to the locus, there was a considerable variation in clinical symptoms between the families[8,10,11]. The presence of different haplotypes suggests that the clinical variation might be related to different mutations in the same gene, although we cannot definitely exclude that different genes in the same chromosomal region may be involved in these three families.

Migraine was investigated and reported in the HERNS and HVR families[7,11]. Raynaud phenomenon, which is an episodic pathological vasomotor reaction of the digital vessels, was investigated and reported as a predominant feature in the HVR family[7]. Currently, no genes have been identified for Raynaud phenomenon, and only one linkage study for Raynaud phenomenon has been performed[12]. Several linkage and association studies have been done for

migraine but no involvement of the 3p21.1-p21.3 region has been reported[13-20]. Here, we tested whether the haplotype co-segregating with HVR in the Dutch family also contributed to increased susceptibility for migraine and Raynaud phenomenon using parent-child and sibling-based transmission disequilibrium tests[21-26].

Materials and methods

Diagnosis of patients and family members

Detailed clinical information on the extended Dutch HVR family was published previously[7,8]. In total, 198 of the 289 family members were personally interviewed and information of the other individuals was obtained indirectly through relatives. Retinopathy was diagnosed by ophthalmologic examination, supplemented with fluorescence angiography of both eyes[9]. Migraine diagnosis was made based on a standard questionnaire, using the criteria of the International Headache Society (IHS)[27]. The diagnosis of Raynaud phenomenon was made according to standardized criteria of Miller et al.[28] All persons gave written informed consent and medical ethical approval was obtained from Leiden University Medical Center (LUMC).

Genotyping

For genotyping, genomic DNA was extracted from peripheral blood using a standard salting out extraction method[29]. To determine the presence of the HVR haplotype, DNA samples of 254 family members and spouses were genotyped for three genetic markers of the 3p21.1-p21.3 region; D3S3564, D3S1581 and D3S1289 [8]. All primer sequences are available through The Genome Database (http://www.gdb.org/). PCRs were performed in 15 µl reaction volume, containing 1.5 µl 10x PCR Buffer II (Applied Biosystems, Foster City, CA), 1.5 µl $MgCl_2$ (25 mM), 1.5 µl dNTPs (2.5 mM), 1.5 µl primer mix (5 pmol/µl), 5.85 µl H_2O, 0.15 µl Amplitaq Gold (5U/µl) (Applied Biosystems, Foster City, CA), and 3.0 µl of genomic DNA (15 ng/µl). The

markers were amplified in two steps using 10 cycles of 30s at 94 °C, 30s at 55 °C, 30s at 72 °C followed by 25 cycles of 30s at 89 °C, 30s at 55 °C, 30s at 72 °C. PCRs were performed using a PTC-200 Thermocycler. Upon PCR, products were separated using an ABI 3700 DNA sequencer (Applied Biosystems, Foster City, CA). Genotypes were analyzed and independently scored by SKhK and KRJV using Genescan and Genotyper 2.1 software (Applied Biosystems, Foster City, CA). Haplotypes were constructed by inspection of allele segregation within the pedigree assuming a minimal number of recombinations.

Statistical analysis

A parent-child trio analysis was performed based on the principle of the transmission disequilibrium test (TDT)[21,23]. In TDT analysis the probability of transmission of the HVR haplotype is compared with the expected probability of 0.5. All trios in which a single parent is heterozygous for the HVR haplotype, and the child is affected with migraine and/or Raynaud phenomenon were selected from the HVR pedigree. Trios in which an affected child was also a heterozygous parent for a trio in a subsequent generation were excluded[30]. Accordingly, the TDT test is not biased by the fact that the trios are related. To evaluate significance, one-sided exact probabilities comparing the HVR haplotype transmission with the expected 0.50 *a priori* probability were calculated using the cumulative binomial distribution function of MS Excel 2000.

In addition to the parent-child TDT tests, sibling case-control associations were tested between the HVR haplotype and migraine or Raynaud phenomenon using the S-TDT design[22-26]. In this test the presence of the HVR haplotype is compared between sibling cases and controls from the nuclear families that make up the complete HVR pedigree. An overall statistic for the risk of carrying the HVR haplotype is subsequently calculated from the

individual results of the nuclear families. From the HVR pedigree, a maximal number of nuclear families were selected in which one parent was a carrier of the HVR haplotype. From these nuclear families all the siblings were selected for the association analysis. Risk estimates and significance were calculated with the Mantel-Haenszel extension (M-H) test, using nuclear family as stratification variable (SPSS for Windows 11.5)[25]. Furthermore, the Z' score approach implemented in the TDT/S-TDT 1.1 program was employed as a control statistic[24].

Results

Possible associations between Raynaud phenomenon, migraine and the HVR haplotype were tested in the large Dutch pedigree. Parent-affected child TDT analyses were used to test for deviations of the haplotype transmission probability, which *a priori* is 0.50 (table 1). For migraine, 23 trios and for Raynaud phenomenon, 26 trios were analyzed. Transmission probability of the HVR haplotype was only slightly increased for individuals with migraine and individuals with Raynaud phenomenon. However, the differences from the expected transmission did not reach significance.

Table 1

TDT test comparing the transmission of HVR haplotype from a heterozygous parent to offspring with Raynaud phenomenon or migraine.

Phenotype in child	HVR T+ N *(p)*	HVR T- N *(p)*	p value
Migraine (n = 23)	13 (0.57)	10 (0.44)	0.202
Raynaud phenomenon (n = 26)	16 (0.62)	10 (0.39)	0.084

HVR T+ = transmitted HVR haplotype, HVR T- = non-transmitted HVR haplotype, N Number of children, *p* = transmission probability.

Next, sibling-control TDT tests were used to compare the risk of migraine and Raynaud phenomenon in carriers and non-carriers of the HVR haplotype. In total, 71 siblings were available from 19 nuclear families in which the haplotype was segregating (table 2). The risk of migraine was increased in HVR carriers. Of the 30 migraineurs, 13 were diagnosed with migraine without aura, 3 with migraine with aura, and 14 with mixed migraine with and without aura. Association analysis for the migraine types separately was not meaningful due to small numbers in each group. The risk of being affected with Raynaud phenomenon was also increased in carriers of the HVR haplotype (table 2).

Table 2

Associations of Raynaud phenomenon and migraine in relation to the presence of the HVR haplotype in siblings of nuclear families.

Siblings in the nuclear families	HVR+	HVR-	Odds ratio (95% CI)[a]	χ^2	p value	Z'	p value
Migraine	19	11	5.87 (1.06 - 32.60)	4.07	0.04	2.02	0.022
No migraine	15	26					
Raynaud phenomenon	25	10	11.36 (2.10 - 61.28)	11.87	0.001	2.70	0.003
No Raynaud phenomenon	9	27					

HVR+ = Hereditary Vascular Retinopathy haplotype carriers, HVR- = non Hereditary Vascular Retinopathy haplotype carriers. χ^2 = The χ^2 test of the common odds ratio equaling 1. Z' = The Z' score test of the TDT/S-TDT 1.1 program. Tests for homogeneity of the odds ratios between nuclear families were not significant for both disorders ($p > 0.09$).

[a] The 95% confidence interval of the Mantel-Haenszel common odds ratio.

Although there was a strong increase in risk in HVR haplotype carriers for having migraine and/or Raynaud phenomenon, the frequency of both disorders was also high in non-carriers of the HVR haplotype; 11 out of 37 had migraine and 10 out of 37 had Raynaud phenomenon. This clearly indicates that other migraine and Raynaud factors must be present in the HVR family.

Discussion

We tested whether the HVR haplotype, harboring the retinopathy gene, contributed to an increased susceptibility to migraine and Raynaud phenomenon in the Dutch family. Siblings with the HVR haplotype did show a significant increased risk of migraine and Raynaud phenomenon compared to non-carrier siblings. We found no significant increase in transmission of the haplotype with the less powerful parent-affected child TDT tests. We have provided genetic evidence that the HVR haplotype harbors a factor that increases the susceptibility for both vascular disorders in the Dutch family. However, the high incidence of these vascular diseases in non-HVR carriers suggests the presence of additional causative factors in this family.

Since retinal cerebrovascular disorders are rare, a sufficiently large sample of unrelated individuals for association analysis is difficult to obtain[8]. Testing for associations between the HVR haplotype and migraine and Raynaud phenomenon gave us a unique opportunity using a within-family approach. The fact that the rare autosomal dominant HVR haplotype is present in only one of the parents of the nuclear families allowed us to study the transmission unambiguously. Since only a single family was studied, a potential problem for testing significance is that the observations may be related[23-26]. Because transmission of the HVR haplotype from parents to offspring is random (i.e. Hardy-Weinberg equilibrium) the use of the applied parent-affected child TDT tests circumvents this potential bias. Furthermore, the results of the TDT test are independent of the disorder prevalence in controls[23]. Unfortunately, parent-child TDT tests lack sufficient detection power because of the low number of tested individuals. Moreover, small changes in the number of transmitted haplotypes have large effects on the significance of the outcome and we, therefore, tend to give less weight to these results.

For the sibling-based TDT test, taking case-control siblings from the same family provides a perfect match for confounding risk factors like population stratification, but it may increase the risk of false-positive results[31,32]. Mantel-Haenszel statistics for related samples, and the Z' score approximation of Spielman and Ewens were used to adjust for the effects of related observations. Because the limited number of samples in each stratum may affect the p-values of the Mantel-Haenszel test, we used two tests[33].

In conclusion, we provided evidence that, within the HVR disease haplotype on chromosome 3p21.1-p21.3, a gene is present that enhances susceptibility for both Raynaud phenomenon and migraine. Future analysis will have to show whether it is the retinopathy gene itself that is associated with migraine and Raynaud phenomenon or whether it is a closely linked gene within the HVR haplotype.

Acknowledgements

We are indebted to the contribution of L.A. Sandkuijl (deceased in December 2002). We would like to thank the family for their co-operation. This work was supported by grants of the Netherlands Organization for Scientific Research (NWO) (903-52-291, M.D.F, R.R.F), The Migraine Trust, (R.R.F, M.D.F), and the European Community (EC-RTN1-1999-00168, R.R.F. and A.M.J.M.v.d.M).

References

1. Lipton RB, Stewart WF. Migraine headaches: epidemiology and comorbidity. Clin Neurosci 1998;5:2-9.
2. Launer LJ, Terwindt GM, Ferrari MD. The prevalence and characteristics of migraine in a population-based cohort: the GEM study. Neurology 1999;53:537-42.
3. Russell MB, Olesen J. Increased familial risk and evidence of genetic factor in migraine. BMJ 1995;311:541-4.
4. Russell MB, Iselius L, Olesen J. Inheritance of migraine investigated by complex segregation analysis. Hum Genet 1995;96:726-30.

5. Verin M, Rolland Y, Landgraf F et al. New phenotype of the cerebral autosomal dominant arteriopathy mapped to chromosome 19: migraine as the prominent clinical feature. J Neurol Neurosurg Psychiatry 1995;59:579-85.
6. Dichgans M, Mayer M, Uttner I et al. The phenotypic spectrum of CADASIL: clinical findings in 102 cases. Ann Neurol 1998;44:731-9.
7. Terwindt GM, Haan J, Ophoff RA, Groenen SM et al. Clinical and genetic analysis of a large Dutch family with autosomal dominant vascular retinopathy, migraine and Raynaud's phenomenon. Brain 1998;121:303-16.
8. Ophoff RA, DeYoung J, Service SK et al. Hereditary vascular retinopathy, cerebroretinal vasculopathy, and hereditary endotheliopathy with retinopathy, nephropathy, and stroke map to a single locus on chromosome 3p21.1-p21.3. Am J Hum Genet 2001;69:447-53.
9. Storimans CW, Van Schooneveld MJ, Oosterhuis JA, Bos PJ. A new autosomal dominant vascular retinopathy syndrome. Eur J Ophthalmol 1991;1:73-8.
10. Grand MG, Kaine J, Fulling K et al. Cerebroretinal vasculopathy. A new hereditary syndrome. Ophthalmology 1988;95:649-59.
11. Jen J, Cohen AH, Yue Q et al. Hereditary endotheliopathy with retinopathy, nephropathy, and stroke (HERNS). Neurology 1997;49:1322-30.
12. Susol E, MacGregor AJ, Barrett JH et al. A two-stage, genome-wide screen for susceptibility loci in primary Raynaud's phenomenon. Arthritis Rheum 2000;43:1641-6.
13. Lea RA, Shepherd AG, Curtain RP et al. A typical migraine susceptibility region localizes to chromosome 1q31. Neurogenetics 2002;4:1:17-22.
14. Nyholt DR, Dawkins JL, Brimage PJ et al. Evidence for an X-linked genetic component in familial typical migraine. Hum Mol Genet 1998;7:459-63.
15. Nyholt DR, Curtain RP, Griffiths LR. Familial typical migraine: significant linkage and localization of a gene to Xq24-28. Hum Genet 2000;107:18-23.
16. Cader ZM, Noble-Topham S, Dyment DA et al. Significant linkage to migraine with aura on chromosome 11q24. Hum Mol Genet 2003;12:2511-7.
17. Wessman M, Kallela M, Kaunisto MA et al. A susceptibility locus for migraine with aura, on chromosome 4q24. Am J Hum Genet 2002;70:652-62.
18. Bjornsson A, Gudmundsson G, Gudfinnsson E et al. Localization of a gene for migraine without aura to chromosome 4q21. Am J Hum Genet. 2003;73:986-93.
19. Carlsson A, Forsgren L, Nylander PO et al. Identification of a susceptibility locus for migraine with and without aura on 6p12.2-p21.1. Neurology 2002;59;11:1804-7.
20. Soranga D, Vettori A, Carraro G et al. A locus for migraine without aura maps on chromosome 14q21.2-q22.3. Am J Hum Genet. 2003;72:161-7.
21. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 1993;52:506-16.
22. Curtis D. Use of siblings as controls in case-control association studies. Ann Hum Genet 1997; 61:319-33.
23. Martin ER, Kaplan NL, Weir BS. Tests for linkage and association in nuclear families. Am J Hum Genet 1997;61:439-48.
24. Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. Am J Hum Genet 1998;62:450-8.
25. Laird NM, Blacker D, Wilcox M. The sib transmission/disequilibrium test is a Mantel-Haenszel test. Am J Hum Genet 1998;63:1915-6
26. Horvath S, Laird NM. A discordant-sibship test for disequilibrium and linkage: no need for parental data. Am J Hum Genet 1998;63:1886-97.

27. Headache Classification Committee of the International Headache Society. Classification and diagnostic criteria for headache disorders, cranial neuralgia, and facial pain. Cephalalgia. 1988;8 Suppl 7:1-96.
28. Miller D, Waters DD, Warnica W et al. Is variant angina the coronary manifestation of a generalized vasospastic disorder? N Engl J Med 1981;304:763-6.
29. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16:1215.
30. Martin ER, Bass MP, Kaplan NL. Correcting for a potential bias in the pedigree disequilibrium test. Am J Hum Genet 2001;68:1065-7.
31. Witte JS, Gauderman WJ, Thomas DC. Asymptotic bias and efficiency in case-control studies of candidate genes and gene-environment interactions: basic family designs. Am J Epidemiol 1999;149:693-705.
32. Weinberg C. Asymptotic bias and efficiency in case-control studies of candidate genes and gene-environment interactions: basic family designs. Am J Epidemiol 2000;152:689-91.
33. Knapp M. Using exact P values to compare the power between the reconstruction-combined transmission/disequilibrium test and the sib transmission/disequilibrium test. Am J Hum Genet 1999;65:1208-10.

CHAPTER 5

Segregation analysis in Dutch migraine families

JJ Hottenga[1,2], GM Terwindt[3], EE Kors[3], LA Sandkuijl[2†], J Haan[4], RR Frants[1], JC van Houwelingen[2], MD Ferrari[3], AMJM van den Maagdenberg[1,3]

1. Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.
2. Department of Medical Statistics, Leiden University Medical Center, Leiden, the Netherlands.
3. Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands.
4. Rijnland Hospital, Leiden, the Netherlands.

† In memory of Lodewijk A. Sandkuijl

Abstract

There is controversy whether migraine with aura (MA) – and without aura (MO) attack types represent separate disease entities. Therefore, it is uncertain whether both types should be analyzed as affected separately or combined in genetic linkage studies. Here, we performed segregation analyses of migraine types in 134 trios of 55 Dutch migraine families to determine the mode of inheritance. Results show that the transmission of migraine types is predominantly of the same type in parents and offspring. Interestingly, a considerable fraction of children with a mixed MA and MO (MA/MO) migraine type has unaffected parents or parents with MA or MO, suggesting a more complex genetic aetiology for this type. This study gives support to designs of separate linkage analysis in MA and MO. Inclusion of individuals with a mixed MA/MO phenotype in linkage analysis remains controversial, although segregation analysis revealed that this mixed phenotype only had minor consequences on the segregation of MO and related estimated parameters.

Keywords

Migraine, complex segregation analysis.

Introduction

Migraine is a common neurovascular disorder manifesting by attacks of severe disabling headache. The lifetime prevalence is up to 6% of men and 18% of women in the general population[1,2]. Peak incidence is during adolescence with a slightly later onset in females as compared to males[3]. Everyone can have a few migraine attacks in their life, but the frequent recurrence of attacks defines a migraine patient. The diagnosis is made on the basis of patient history and is categorized in attack types according to standardized diagnostic criteria as defined by the International Headache Society (IHS)[4]. The two most common attack types are migraine without aura (MO) and migraine with aura (MA). Attacks of MO are characterized by severe, often unilateral, throbbing headache that is aggravated by physical activity and is accompanied by other disabling neurological symptoms like vomiting, nausea, photophobia and/or phonophobia. In attacks of MA, the headache is preceded or accompanied by an aura phase. The aura symptoms are primarily visual symptoms and occur in about one-third of migraine patients.

Knowledge on the pathophysiology of migraine headache is increasing, but it is hardly known why migraine attacks begin. Genetic - as well as environmental factors are involved in the migraine pathophysiology[5-9]. Genes may predispose to the migraine recurrence by lowering the threshold for migraine attacks. This may render a migraine sufferer more sensitive to often-mentioned triggers like stress, menstrual cycle, alcohol use and lack of sleep.

Mapping of migraine genes was initiated in familial hemiplegic migraine (FHM), a rare autosomal dominant form of MA where patients develop one-sided hemiparesis during the attack[4]. Two FHM genes have been identified using parametric linkage approaches. The first FHM gene, CACNA1A is located on chromosome 19p13 and encodes the $Ca_v2.1$ (formerly known as alpha 1A) subunit of P/Q-type calcium channels[10]. Mutations in this gene were found in several FHM families, but a definite involvement of this gene in the

common types of migraine needs to be established[11-15]. The second FHM locus (FHM2), located on chromosome 1q23.2, encodes the Na^+,K^+-ATPase $\alpha 2$ subunit[16]. Also the involvement of this gene in the common types of migraine has not yet been established[15].

Gene identification for the common types of migraine has been challenging. Main reasons are the high prevalence of migraine, the large influence of environmental factors and genetic and clinical heterogeneity. A crucial issue is whether the MA and MO migraine types can be considered one disease entity or not. As a consequence, it remains a controversial issue whether both types should be analyzed as affected separately or combined in genetic linkage studies. Separation has been advocated based on clinical - and physiological characteristics, results of twin studies and differences in heritability[4,17-22]. An important observation is that single patients with both attack types are more frequently and more severely affected[21]. This may indicate that type-specific risk factors contribute to the severity of disease in these individuals. In contrast, one could argue that both migraine types are one disease entity because of clinical overlap in the attacks, frequent co-occurrence of attacks in a single patient, and the observation that family members have different migraine types[17,23,24]. Locations for migraine genes have been identified analyzing both types, uniquely or in combination. Two loci for MA were identified on chromosomes 4q24 and 11q24 [25,26]. Loci for MO were found on chromosomes 14q21.2-q22.3 in an Italian family and 4q21 in Icelandic families[27,28]. Additional migraine loci have been identified on chromosomes Xq24-q28 and 6p12.2-p21.1 using both MA and MO as affected status in Australian and Swedish families, respectively[29-31].

In the present study, the segregation of MA and MO was studied in 55 Dutch migraine families. The main focus was to study the segregation of the major migraine types MO and MA in these families, but the consequences of adding patients with both attack types to the analyses were investigated as well.

Segregation of MA and MO from parents to offspring was examined in 134 independent trios selected from these families. Complex segregation analysis was performed with POINTER, to study the segregation of MO and to evaluate if inclusion of individuals with a combined MA/MO phenotype affected the segregation findings[32]. Next, a separate segregation analysis was performed in a specific subset of seven families with a seemingly autosomal dominant transmission of MO and only a minimum of persons affected with MA. This subset was selected because it could serve as a homogenous sample for a genome wide linkage analysis to identify novel migraine loci (see *chapter six*).

Methods

Data collection

Patients of the Dutch outpatient headache clinic of the Leiden University Medical Center (LUMC) as well as patients responding to calls in local newspapers or in the periodical of the Dutch Migraine Patients Association were screened for a positive family history of migraine[11]. In case of a positive family history, all available family members of the proband were interviewed thus aiming to extend the number of family branches with affected individuals as much as possible. Probands and relatives that gave informed consent were personally examined and interviewed by experienced neurologists using a semi-structured questionnaire. Medical ethical approval for the study was obtained from the LUMC. Patients with migraine were diagnosed according to the International Headache Society classification criteria[4]. The main inclusion criterion for the ascertained families was at least two generations affected with migraine. Families with patients affected with FHM or prolonged aura were excluded from this study. The selection of families does *not* represent a random sample from the population. However, because no selection for the attack type of migraine was performed, the sample is useful to study the relation between MA and MO *within* this set of families. In total 55 Dutch migraine families were included. From the complete sample, a specific subset

of seven migraine families was selected based on the fact that migraine segregated as an apparent autosomal dominant trait and nearly all patients in these families suffered from MO (89%). Only a small fraction of the patients presented either MA (3%) or a mixed phenotype of MA and MO (8%).

Statistical analysis

Trio analysis

To examine the transmission of migraine types, parent-child trios were selected from the complete sample of 55 migraine families. The trios were chosen to be independent from each other; all parents were selected only once, and a single affected child with migraine was selected in case there were multiple affected siblings. The trios were then divided based on the migraine attack type of the children. Since patients can have attacks of both types of migraine, a third group named MA/MO was introduced, using a similar approach as Kallela et al[21]. The affected children were subsequently stratified for the migraine type(s) of the parents. A group of seven trios with two affected parents were not considered in the analysis mainly because of the small group size. Investigation of attack types in children and their parents provides a good indicator of how migraine types are transmitted. It also provides a measure for the probability that a child develops a specific migraine type given the parents' disease status. A Pearson χ^2 test was employed to test whether the migraine types were transmitted independently (SPSS for Windows 10.0).

Pointer analysis

Complex segregation analysis was carried out with the program POINTER, incorporating the unified mixed model as described by Lalouel and Morton[32-34]. The model assumes that the liability of the disorder can be described by an underlying continuous liability scale (x). The liability for individuals is modeled by the contribution of three independent factors, namely $x = g + c + e$. The first factor is a major single gene locus (g) that causes a displacement of

at least one standard deviation on the liability scale between the normal (*NN*) and abnormal (*AA*) genotypes. The second factor is a polygenic component (*c*) attributable to a number of additive genetic and environmental risk factors transmitted from parents to offspring. The third factor is a random non-transmitted factor (*e*). The polygenic - and environmental factors are considered to be normally distributed. The total variance *V* of the liability can be divided over the three factors *g*, *c*, and *e* similar to the liability scale. The variance then equals $V = G + C + E$ in which *G*, *C* and *E* represent variances of the factors *g*, *c* and *e*, respectively. The proportion of variance caused by the polygenic component is denoted as the heritability $H = C/V$. The major locus *g* is assumed to have two alleles *N* and *A* producing the genotypes *NN*, *NA* and *AA*. Equal transmission probability for both alleles from parents to offspring is assumed, i.e. there is Mendelian transmission and Hardy-Weinberg Equilibrium (HWE)[35-37]. The major locus is defined by three parameters: *q* the frequency of allele *A* in the population, *t* the distance measured in standard deviations on the liability scale between the *NN* and *AA* genotype carriers, and *d* the degree of dominance expressed as the position of the heterozygous class mean (*AN* carriers) in relation to the homozygous class means (*AA* and *NN* carriers). If $d = 0.0$ the gene is recessive, if $d = 0.5$ the gene is additive and if $d = 1.0$ the gene is dominant.

In the unified model, the affected state for a dichotomous trait such as migraine is defined by exceeding a threshold on the liability scale (*x*). In this case, the mean and variance *V* of the liability scale *x* are defined as 0 and 1. The threshold is determined from the prevalence of the disorder, here specified separately for males and females, because migraine prevalence shows gender differences[1,2,32]. The prevalences were obtained from the Dutch population and adjusted for the fact that only MO was studied by taking 70% of the total migraine prevalence[2]. The values were set to 0.160 for females and 0.048 for males.

Both data sets, the complete set of 55 migraine families and the subset of 7 MO families with an apparent autosomal dominant inheritance, were divided into nuclear families; the larger set was divided into 340 nuclear families with in total 995 children, whereas the subset of seven MO families was divided into 64 nuclear families with in total 200 children. Next, pointers were assigned to refer to the original probands following the methods of Lalouel and Morton[33]. Comments about the use of pointers were noted and applied to assign them correctly[38]. POINTER specifies the ascertainment probability π of an affected individual from the population becoming a sampled proband. All affected individuals are assumed to have the same ascertainment probability, and all probands are assumed independently ascertained. In the family samples presented here, these assumptions are reasonable. The ascertainment probability π was arbitrarily set to 0.001, approximating single selection of probands from the population[33]. POINTER handles the selection of additional familial cases as follows: if the case is a sibling, an approximate sampling correction consists of defining the proband as a pointer, whereas his siblings are treated under truncate selection. In other instances, conditioning on parents or pointers accounts for such mode of selection[33,39].

With POINTER the model parameters of the different segregation models are estimated on the basis of the maximum likelihood principle, using an iterative optimization procedure. The fit of the model is reported through the *deviance* $= -2\ln(L) + k$. The smaller the deviance, the better. Several starting values, a total of five very different combinations, were used for the parameters to ensure that the global optimum was found[35,40,41]. Initially, the following models were tested: the *Mixed general model* with all four parameters *d*, *t*, *q* and *H* free and the g*eneral single locus (GSL) model* in which the parameters *d*, *t* and *q* are free. The *GSL model* was then analyzed further for a specific mode of inheritance in which *d* was fixed to 1.0 (dominant), 0.5 (additive) or 0.0 (recessive), whereas parameters *t* and *q* were kept free. Subsequently, the *Polygenic model* was tested in which only parameter *H* was free and finally

the *Sporadic model* was tested with all parameters fixed, including $H = 0.0$ and $q = 0.0$.

All segregation models were tested, first with only the MO persons being labeled as affected and individuals with MA/MO set to unknown, and second with individuals suffering from MA/MO or MO being labeled as affected. Patients with only MA were set to unknown in all POINTER analyses. Comparison of nested models was performed with the likelihood ratio test based on Wilk's theorem. The Akaike information criterion (*AIC*) was used for other comparisons[42,43]. The likelihood ratio test is equal to the difference between the deviance values, a χ^2 distributed statistic, with the number of degrees of freedom (df) equal to the difference in dimension *dim* (= number of free parameters) between the tested models. The *AIC* was calculated for each model by *AIC = deviance + 2 dim*. The model with the lowest *AIC* has the best fit. The significance level was considered at $\alpha = 0.05$ and no correction was made for multiple testing.

Results

Trio analysis

Transmission of migraine attack types, MA, MO and mixed MA/MO was studied in a parent-child trio analysis. In total, 134 independent trios were selected from 55 migraine families. Results of this analysis are presented in table 1. Statistical analysis of the data with one affected parent strongly rejected the hypothesis that the transmission of migraine attack types from the parent to the offspring is independent ($\chi^2 = 23.6$, 4 df, $p < 0.0001$). A substantial proportion of children have the same migraine type as their parent; 58% have MA if one parent has MA, 73% have MO if one parent has MO. In the group of children with MA/MO most have an affected parent with either MA or MO, and only a minority (25%) has a parent with MA/MO. As only a few children had two affected parents with migraine, this group was therefore

not analyzed with further detail. However, an interesting observation in the MA/MO subgroup is that a large number of children have two unaffected parents (13 of 33 children).

Table 1

Segregation analysis of migraine with - and without aura in independent trios selected from 55 Dutch migraine families.

	Migraine type in child		
	MA (%)	MA/MO (%)	MO (%)
One parent MA - one parent unaffected	7 (58%)	9 (45%)	8 (11%)
One parent MA/MO - one parent unaffected	2 (17%)	5 (25%)	12 (16%)
One parent MO - one parent unaffected	3 (25%)	6 (30%)	53 (73%)
Sub total one parent affected	12 (100%)	20 (100%)	73 (100%)
Two parents affected	2	0	5
Two parents unaffected	5	13	4
Total	19	33	82

Pearson χ^2 test for independence of transmission in children with one affected parent:
$\chi^2 = 23.6$, 4 df, $p < 0.0001$

Complex segregation analysis 55 migraine families

Segregation analysis was first performed by testing only patients with MO as affected, while patients with a mixed MA/MO and MA phenotype were set to unknown. Secondly, both MA/MO and MO patients were tested as being affected, while MA remained unknown. The results of fitting the various segregation models are presented in table 2. Setting the MA/MO patients to unknown did not really influence the results as can be inferred from the fact that the estimated parameters for any of the models were similar: t (the difference in means of *NN* and *AA* measured in Sds), q (the frequency of allele *A*) and the heritability (H) for any of the models. Based on the order of the *AIC*, addition of MA/MO patients did not alter conclusions about the optimal segregation model of MO in the set of 55 migraine families.

Table 2

Estimated parameters and maximum likelihood of the tested segregation models in 55 Dutch migraine families.

Affected definition	Segregation model	*d*	*t*	*q*	*H*	*deviance*	*dim*	*AIC*
MO	Mixed general	1	2.503	0.098	0.0026	276.3	4	284.3
	GSL	1	2.419	0.096	0.0*	303.7	3	309.7
	GSL Dominant	1.0*	2.370	0.096	0.0*	303.7	2	307.7
	GSL Additive	0.5*	4.371	0.101	0.0*	310.6	2	314.6
	GSL Recessive	0.0*	2.326	0.438	0.0*	334.0	2	338.0
	Polygenic			0.0*	0.971	310.9	1	312.9
	Sporadic			0.0*	0.0*	589.0	0	589.0
MA/MO and MO	Mixed general	1	2.415	0.106	0.0039	481.6	4	489.6
	GSL	1	2.400	0.096	0.0*	514.2	3	520.2
	GSL Dominant	1.0*	2.379	0.096	0.0*	514.2	2	518.2
	GSL Additive	0.5*	4.338	0.102	0.0*	522.6	2	526.6
	GSL Recessive	0.0*	2.454	0.433	0.0*	550.4	2	554.4
	Polygenic			0.0*	0.969	521.7	1	523.7
	Sporadic			0.0*	0.0*	892.9	0	892.9

* Parameters fixed in the analysis, see the text for further explanation of the variables *d* = dominance, *t* = displacement, *q* = gene frequency, *H* = heritability, *deviance* = -2Ln(*L*) + *k*, *dim* = number of free parameters, *AIC* = Akaike information criterion, GSL = General single locus. MO = segregation analysis in which only patients with MO were considered affected and patients with MA/MO attacks were set to unknown. MA/MO and MO = patients with MA/MO and MO were considered being affected for the segregation analysis.

Given the similar results, comparison of segregation models was done only for the analysis in which only MO individuals were included as affected. The large difference in likelihood of the *Sporadic* and *Polygenic models* together with the high heritability value $H = 0.971$ shows that migraine is indeed genetic in these families ($\chi^2 = 278.10$, 1 df, $p < 0.00001$). Comparing the *AIC* of the *Polygenic model* with the *Dominant GSL model* shows that an autosomal dominant major single locus fits the data better. The recessive - and additive modes of segregation are very unlikely, since the *GSL model* has a much better fit (Recessive, $\chi^2 = 30.3$, 1 df, $p < 0.00001$) (Additive, $\chi^2 = 6.9$, 1 df, $p = 0.009$). A dominant mode of inheritance for migraine is also indicated in the *GSL model* as the dominance parameter also converges to one.

Compared to the *Dominant GSL model*, the *Mixed general model* with a major autosomal dominant factor and a very small residual heritability explains the data still significantly better (χ^2 = 27.4, 2 df, $p < 0.00001$). This result is unexpected since all estimated parameters remain very similar and is likely related to problems of fitting of the maximum likelihood with this model.

Complex segregation analysis in the subset of seven migraine families with an apparent autosomal dominant inheritance of MO

A subset of 7 families could be selected from the sample of 55 Dutch migraine families because of an apparent autosomal dominant segregation of MO. From a total of 111 migraine patients in these families, 99 suffered from only MO attacks, whereas 9 had MA/MO attacks and 3 had MA attacks exclusively. Segregation analyses for this specific subset of families were performed with or without MA/MO individuals regarded as being affected. In the analyses, the three MA patients' affected status was set to unknown. The *Mixed general model* with MA/MO patients as affected gave difficulties converging to the maximum likelihood. This could be a model with a small contribution of the major single locus and a large contribution of the polygenic component, or a model with a large contribution of the major single locus and a small contribution of the polygenic component. The current maximum likelihood was based on a grid search of the *H* values. The results of the segregation analyses are presented in table 3.

In the subset of MO-selected families, the estimated values of parameters *t*, *q* and *H* for the segregation models with MA/MO individuals being set to unknown or regarded as affected are again similar. An exception is the *Mixed general model* where the estimated *H* of the MO only analysis shows a large difference with the MA/MO plus MO analysis. Comparison of the different segregation models shows that the *Mixed general model* fits significantly better than the other models for the analyses with - and without MA/MO patients assigned affected (MA/MO unknown, *Mixed vs. Dominant GSL*, χ^2 =

9.28, 2 df, p = 0.0010) (MA/MO affected, *Mixed vs. Polygenic*, χ^2 = 24.70, 3 df, p = 0.00002). In both mixed models there is a major autosomal dominant genetic component and a smaller residual heritability component. Comparing the *GSL models* shows that the dominance *d* is optimal at 1.0, indicating an autosomal dominant mode of inheritance of the major factor. The *recessive GSL*, *additive GSL* and *sporadic models* can be excluded in the analyses both with - and without MA/MO patients considered affected. The *polygenic* and *GSL (dominant) model* fit equally well in the MA/MO unknown analysis, in the MA/MO affected analysis the *polygenic model* fits slightly better.

Table 3

Estimated parameters and maximum likelihood of the tested segregation models in the seven Dutch migraine without aura families.

Affected definition	Segregation model	*d*	*t*	*q*	*H*	*deviance*	*dim*	*AIC*
MO	Mixed general	1	2.568	0.067	0.2565**	188.4	4	196.4
	GSL	1	2.932	0.050	0.0*	197.6	3	203.6
	GSL Dominant	1.0*	2.979	0.050	0.0*	197.6	2	201.6
	GSL Additive	0.5*	3.233	0.231	0.0*	209.8	2	213.8
	GSL Recessive	0.0*	2.266	0.350	0.0*	220.9	2	224.9
	Polygenic			0.0*	0.997	198.2	1	200.2
	Sporadic			0.0*	0.0*	348.6	0	348.6
MA/MO and MO	Mixed general	1	2.639	0.083	0.057	180.7	4	188.7
	GSL	1	2.879	0.053	0.0*	207.6	3	213.6
	GSL Dominant	1.0*	2.876	0.053	0.0*	207.6	2	211.6
	GSL Additive	0.5*	3.218	0.234	0.0*	219.1	2	223.1
	GSL Recessive	0.0*	2.188	0.377	0.0*	228.0	2	232.0
	Polygenic			0.0*	0.997	205.4	1	207.4
	Sporadic			0.0*	0.0*	368.1	0	368.1

* Parameters fixed in the analysis, see the text for further explanation of the variables *d* = dominance, *t* = displacement, *q* = gene frequency, *H* = heritability, *deviance* = -2Ln(*L*) + *k*, *dim* = number of free parameters, *AIC* = Akaike information criterion, GSL = General single locus. ** Value of H was based on a grid search. MO = segregation analysis in which only patients with MO were considered affected and patients with MA/MO attacks were set to unknown. MA/MO and MO = patients with MA/MO and MO were considered being affected for the segregation analysis.

When comparing the analyses of the 55 migraine families with those of the subset of seven MO families, results indicated that the estimated heritability values (*H*) are very similar (0.969 vs. 0.997). The gene frequency *q* is generally estimated lower in the seven MO families, except for the *additive GSL model*. In all segregation analyses, the *Mixed general model* having a major dominant component and some residual heritability fits optimally. The *GSL dominant model* is the second best model in the analyses of the 55 families, however in the subset of 7 MO families the *polygenic model* was second best. For both samples, the effect of adding persons that have both MA and MO attacks as affected on the estimated parameter values is small.

Discussion

Segregation of migraine types MA and MO and the mixed MA/MO phenotype was studied in 55 Dutch migraine families. Results of the present study show that segregation of migraine attack types MA and MO from parent to child is dependent to a large extent on the type in the parent. This strongly suggests that different risk factors are transmitted for the MA and MO migraine types in this sample of families. However, the results also indicate that some common risk factors for the different migraine types could be present (MO $\Rightarrow$ MA 11%, MA $\Rightarrow$ MO 25%). Whether MA and MO are genetically distinct disease entities could not fully be determined using this approach. Research of identified migraine loci and future genes needs to be performed in families with mixed migraine types in order to provide more detailed information. Recently, separation of MA and MO has shown to be successful in migraine linkage studies[25,28]. Independent of whether MA and MO are separate genetic entities, separation of migraine attack types apparently increases the phenotype homogeneity, which probably results in an enhanced probability of finding migraine loci, and ultimately genes[44].

In the present complex segregation analyses it was shown that an autosomal dominant mode of inheritance of MO is most likely in the 55 migraine families and that migraine is highly heritable in this sample of families. Best fitting models are the *Mixed general (dominant) model* and the *Dominant GSL model*. Parameters estimates for these models are very similar, where the *Mixed general model* can be considered equal to the *Dominant GSL model* with a very small residual heritability of $H < 0.0036$. Probably there was an artifact in the likelihood maximization for the *mixed general model*, as the near zero heritability contributed to an extremely large likelihood difference.

For the analyses of the seven selected migraine families with an apparent autosomal dominant inheritance of MO, results show that MO seems fully inherited in these families ($H = 99.7\%$). Given the strong selection criteria that were applied for this sample, this finding is not surprising. The optimal mode of inheritance is a *Mixed general model* with an autosomal dominant component and residual heritability. Again, there is a substantial increase in likelihood as compared to the alternative models. Whether this is related to difficulties of fitting the maximum likelihood is difficult to determine, as the imputed H in the *Mixed general model* also shows a likelihood increase and the parameters estimates differ more between the models. Other models that showed a good fit and fitted equally well were the *Dominant GSL model* and the *Polygenic model*.

How should we consider patients that are affected with both MA and MO? The trio analysis shows that the parents of these individuals are either unaffected, or if affected then in most cases exclusively with either MA or MO. The high percentage of unaffected parents of MA/MO patients may indicate that non-penetrant risk factors or non-transmitted risk factors play an important role in this particular phenotype. Adding the MA/MO affected individuals in the complex segregation analysis of MO, did not change the results for the tested segregation models or estimated parameters. This was

observed in the sample of 55 families, as well as the sample of selected MO families. There are several possible explanations for this finding. First, the number of MA/MO affected individuals in a single family is too small and therefore does not influence the segregation calculations strongly. Second, MA/MO children have obtained only the MO risk factors from their parents. The prevalence of migraine, especially of MO is high, therefore, a combination of the two migraine types related to a different cause in a single patient is likely to occur[2,21]. Based on the segregation analyses, both inclusion and exclusion of the MA/MO individuals in a MO linkage analysis can be applied. Inclusion could increase the probability of heterogeneity and phenocopies, exclusion could reduce linkage detection power and loss of possible valuable recombinant information.

A limitation of this study is that the migraine families were not randomly selected from the population. In the trio analysis studying the segregation of MA, MO and MA/MO the problem was circumvented using the specific trio analysis. Regarding complex segregation analysis, conclusions about the parameter estimates and the optimal model are valid only for these family samples. Furthermore, the selection of two generations with affected members in the families likely excluded the *Recessive GSL model a priori*, and may have had a substantial contribution on the poor fit of the *Additive GSL model*. The finding of an autosomal dominant mode of inheritance was therefore expected. Furthermore, the genetic component of migraine in the families has probably been overestimated and parameter estimates are likely not useful for the Dutch population. The selection did not influence the effect of testing the addition of MA/MO affected individuals in the segregation analysis. It should be noted that, even with a random confirmed single selection of probands from the population, the results of segregation analysis can be misleading in complex disorders and even simple traits[41,45-47]. In complex disorders like migraine segregation analysis may be an oversimplification, i.e. a found dominant major single locus may still be on different chromosomes within

different migraine families and several genes may contribute to the migraine liability.

In conclusion, it seems reasonable to analyze linkage for MO without including MA patients as being affected, given the high probabilities of transmission of the same migraine types and the outcome of previous linkage and segregation studies. The inclusion of MA/MO individuals does not strongly alter parameter estimates or the optimal segregation model, therefore, inclusion or exclusion of MA/MO individuals for MO linkage analysis remains inconclusive. In case a model-based (parametric) linkage approach will be applied, the use of an autosomal dominant mode of inheritance seems optimal for the collected families.

Acknowledgements

We would like to thank the families for their co-operation. Many grateful thanks also go to Lodewijk A. Sandkuijl who passed away in December 2002. The work was supported by grants of the Netherlands Organization for Scientific Research (NWO), The Migraine Trust, and the European Community (EC-RTN1-1999-00168).

References

1. Lipton RB and Stewart WF. Migraine headaches: epidemiology and comorbidity. Clin Neurosci. 1998;5:2-9.
2. Launer LJ, Terwindt GM, Ferrari MD. The prevalence and characteristics of migraine in a population-based cohort: the GEM study. Neurology. 1999;53:537-42.
3. Stewart WF, Shechter A, Lipton RB. Migraine heterogeneity. Disability, pain intensity, and attack frequency and duration. Neurology. 1994;44(6 Suppl 4):S24-39. Review.
4. "International Headache Society". Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. Cephalalgia 8 1988;(supplement 7):1-96.
5. Russell MB, Olesen J. Increased familial risk and evidence of genetic factor in migraine. BMJ. 1995;311:541-44.
6. Russell MB, Iselius L, Olesen J. Migraine without aura and migraine with aura are inherited disorders. Cephalalgia. 1996;16:305-9.
7. Gervil M, Ulrich V, Kaprio J, Olesen J, Russell MB. The relative role of genetic and environmental factors in migraine without aura. Neurology. 1999;53:995-9.

8. Ziegler DK, Hur YM, Bouchard TJ et al. Migraine in twins raised together and apart. Headache. 1998;38:417-22.
9. Mulder EJ, Van Baal C, Gaist D et al. Genetic and environmental influences on migraine: a twin study across six countries. Twin Res. 2003;6:422-31.
10. Ophoff RA, Terwindt GM, Vergouwe MN et al. Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. Cell. 1996;87:543-52.
11. Terwindt GM, Ophoff RA, van Eijk R et al. Involvement of the CACNA1A gene containing region on 19p13 in migraine with and without aura. Neurology. 2001;56:8:1028-32.
12. Lea RA, Curtain RP, Hutchins C et al. Investigation of the CACNA1A gene as a candidate for typical migraine susceptibility. Am J Med Genet. 2001;105:707-12.
13. McCarthy LC, Hosford DA, Riley JH et al. Single-nucleotide polymorphism alleles in the insulin receptor gene are associated with typical migraine. Genomics. 2001; 78:135-49.
14. Kaunisto MA, Tikka PJ, Kallela M et al. Chromosome 19p13 loci in Finnish migraine with aura families. Am J Med Genet. 2004. Sep 24 Epub.
15. Jen JC, Kim GW, Dudding KA, Baloh RW. CACNA1A and ATP1A2 in probands with common types of migraine. Arch Neurol. 2004;61:926-8.
16. De Fusco M, Marconi R, Silvestri L et al. Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. Nat Genet. 2003;33:192-6.
17. Kallela M, Wessman M, Farkkila M et al. Clinical characteristics of migraine in a population-based twin sample: similarities and differences between migraine with and without aura. Cephalalgia. 1999;19:151-58.
18. Olesen J, Tfelt-Hansen P, Henriksen L, Larsen B. The common migraine attack may not be initiated by cerebral ischaemia. Lancet. 1981;29:438-40.
19. Olesen J, Larsen B, Lauritzen M. Focal hyperemia followed by spreading oligemia and impaired activation of rCBF in classic migraine. Ann Neurol. 1981;9:344-52.
20. Russell MB, Rasmussen BK, Fenger K, Olesen J. Migraine without aura and migraine with aura are distinct clinical entities: a study of four hundred and eighty-four male and female migraineurs from the general population. Cephalalgia. 1996;16:239-45.
21. Kallela M, Wessman M, Havanka H et al. Familial migraine with and without aura: clinical characteristics and co-occurrence. Eur J Neurol. 2001;8:5:441-49.
22. Rasmussen BK, Olesen J. Migraine with aura and migraine without aura: an epidemiological study. Cephalalgia. 1992;12:221-28.
23. Rapoport MA. MO and MA: distinct clinical entities? Cephalalgia. 1996;16:215.
24. Blau JN. Migraine with aura and migraine without aura are not different entities. Cephalalgia. 1995;15:186-90.
25. Wessman M, Kallela M, Kaunisto MA et al. A susceptibility locus for migraine with aura, on chromosome 4q24. Am J Hum Genet. 2002;70:652-62.
26. Cader ZM, Noble-Topham S et al. Significant Linkage to Migraine with Aura on Chromosome 11q24. Hum Mol Genet. 2003;12:2511-7.
27. Soranga D, Vettori A, Carraro G et al. A locus for migraine without aura maps on chromosome 14q21.2-q22.3. Am J Hum Genet. 2003;72:161-167.
28. Bjornsson A, Gudmundsson G, Gudfinnsson E et al. Localization of a gene for migraine without aura to chromosome 4q21. Am J Hum Genet. 2003;73:986-93.
29. Nyholt DR, Dawkins JL, Brimage PJ et al. Evidence for an X-linked genetic component in familial typical migraine. Hum Mol Genet. 1998;7:459-63.

30. Nyholt DR, Curtain RP, Griffiths LR. Familial typical migraine: significant linkage and localization of a gene to Xq24-28. Hum Genet. 2000;107:18-23.
31. Carlsson A, Forsgren L, Nylander PO et al. Identification of a susceptibility locus for migraine with and without aura on 6p12.2-p21.1. Neurology. 2002;59:11:1804-7.
32. Morton NE, Rao DC, Lalouel JM. Methods in genetic epidemiology. S. Karger Publishing. New York. 1983. ISBN 3805536682.
33. Lalouel JM, Morton NE. Complex segregation analysis with pointers. Hum Hered. 1981;31:312-21.
34. Lalouel JM, Rao DC, Morton NE, Elston RC. A unified model for complex segregation analysis. Am J Hum Genet. 1983;35:816-26.
35. Iselius L, Morton NE. Transmission probabilities are not correctly implemented in the computer program POINTER. Am J Hum Genet. 1991;49:459.
36. Dizier MH, Bonaiti-Pellie C, Clerget-Darpoux F. POINTER correctly estimates the transmission probabilities under the general transmission model in the case of incomplete selection. Am J Hum Genet. 1992;51:1449-50.
37. Demenais F, Martinez M, Andrieu N. The transmission probability model is useful to prevent false inference. Am J Hum Genet. 1993; 52:441-2.
38. Shields DC, Collins A, Marlow A. Coding of pointers in the segregation analysis program POINTER. Genet Epidemiol 1994;11:385-7.
39. Moilanen JS, Autere JM, Myllyla VV, Majamaa K. Complex segregation analysis of Parkinson's disease in the Finnish population. Hum Genet. 2001;108:184-9.
40. Walkup JT, LaBuda MC, Singer HS et al. Family study and segregation analysis of Tourette syndrome: evidence for a mixed model of inheritance. Am J Hum Genet. 1996;59:684-93.
41. McGuffin P, Huckle P. Simulation of Mendelism revisited: the recessive gene for attending medical school. Am J Hum Genet. 1990;46:994-9.
42. Wilk SS. Certain generalizations in the analysis of variance. Biometrica 1932;24:471-94.
43. Akaike H. A new look at the statistical model identification. IEEE Trans Automatic Control. 1974;19:716-23.
44. Lander ES, Schork NJ. Genetic dissection of complex traits. Science. 1994;265:2037-48.
45. Wang TN, Ko YC, Wang TH et al. Segregation analysis of asthma: recessive major gene component for asthma in relation to history of atopic diseases. Am J Med Genet. 2000;93:373-80.
46. Axenovich TI, Borodin PM. Some pitfalls of segregation analysis of complex traits. Am J Med Genet. 2002;111:228-9.
47. Cheng LS, Wang TN, Ko YC. Reply to correspondence from Axenovich and Borodin--"Some pitfalls of segregation analysis of complex traits". Am J Med Genet. 2003;117A:101-2.

CHAPTER 6

Involvement of the 4q21-q24 migraine locus in Dutch migraine without aura families

JJ Hottenga[1,2], EE Kors[3], GM Terwindt[3], KRJ Vanmolkot[1], S Kheradmand Kia[1], J Haan[3,4], LA Sandkuijl[2†], MD Ferrari[3], RR Frants[1], AMJM van den Maagdenberg[1,3]

1. Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.
2. Department of Medical Statistics, Leiden University Medical Center, Leiden, the Netherlands.
3. Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands.
4. Department of Neurology, Rijnland Hospital, Leiderdorp, the Netherlands.

† In memory of Lodewijk A. Sandkuijl

Abstract

Migraine is a common neurological disorder with a prevalence of about 12% in the general population. Several loci and genes have been identified for familial hemiplegic migraine (FHM), a Mendelian type of migraine with aura. Gene identification for the common types of migraine is more challenging, mainly because of the complex genetics and high prevalence of the disorder. Still, several loci have been identified for migraine with aura (MA) and migraine without aura (MO). In this study, seven Dutch families with apparent dominantly inherited MO were selected for a genome-wide scan (at 9 cM marker interval). In total, 392 markers were tested, and suggestive evidence for linkage was found for chromosomal region 4q21-q24. For marker D4S2361, the maximum multipoint LOD score of 1.98 was observed when analyzing all families combined. A LOD score of 2.59 was found combining only the four 4q-positive families. This study provides an independent replication of two previous studies showing linkage to the same region in MA and MO families, respectively.

Keywords

Migraine, genome scan, linkage replication.

Introduction

Migraine is a paroxysmal neurovascular disorder affecting up to 6% of males and 18% of females in the general population[1,2]. Diagnosis is made on the basis of patient history and is categorized in attack types according to standardized diagnostic criteria as defined by the International Headache Society (IHS)[3]. Attacks of migraine without aura (MO) are characterized by severe, often unilateral, throbbing headache that is aggravated by physical activity and is accompanied by other disabling neurological symptoms like vomiting, nausea, photophobia and/or phonophobia. In one-third of the migraine patients the headache phase is preceded or accompanied by, primarily visual, aura-symptoms; migraine with aura (MA).

Family and twin studies have clearly indicated that migraine is a complex disorder with involvement of both genetic and environmental factors[4-7]. Gene identification in the common types of migraine has been difficult, mainly because of the high prevalence and variable expression of the disorder. Furthermore, the lack of biochemical markers for unequivocal diagnosis makes the definition of the affected status for genetic studies difficult[8].

Gene identification studies in migraine were initiated in a rare, autosomal dominant type of migraine with aura; familial hemiplegic migraine (FHM). In FHM, the aura is accompanied by hemiparesis. The first FHM gene (FHM1), CACNA1A, encodes the alpha 1A ($Ca_v2.1$) subunit of P/Q-type calcium channels, indicating that FHM, at least in part, is a channelopathy[9]. About 50% of FHM families are linked to the CACNA1A locus on chromosome 19p13[10]. Sib-pair studies have shown that the CACNA1A region is also involved in the common types of migraine, especially MA, but CACNA1A gene mutations have not been identified so far[11,12]. It has been suggested that the insulin-receptor (INSR) gene could be the causative gene on 19p13 [13]. A second FHM locus (FHM2) was identified on chromosome 1q23.2. Recently, the causative gene, ATP1A2, has been identified that encodes the Na^+/K^+ ATPase $\alpha 2$

subunit[14,15]. At least a third FHM locus must exist since some families are linked to either chromosome 19p13 or 1q23.2 [16,17]. Genome scans have also revealed several loci for the common types of migraine. Loci have been reported on chromosomes 1q31 and Xq24-q28 in Australian families[18-20]. A locus on chromosome 11q24 was identified in Canadian MA families[21]. In addition, migraine loci have been mapped to chromosomes 6p12.2-p21.1 and 14q21.2-q22.3 in a Swedish MO/MA family and an Italian MO family, respectively[22,23]. Recently, a locus for MA has also been found in Finnish families on chromosome 4q24 [24]. Intriguingly, this locus has been replicated in migraine families from Iceland, however, mainly in women with MO[25]. For none of the loci involved in the common types of migraine the causative gene has been identified yet.

Here, we performed a genome-wide scan to identify migraine loci in seven Dutch MO families, which were selected because MO segregated as an apparent autosomal dominant trait. Suggestive evidence for linkage was found in the chromosome 4q21-q24 area. This study provides a new independent replication of the Finnish and Icelandic studies. It thereby supports evidence that one or more migraine loci are present in the chromosome 4q21-q24 region.

Materials and methods

Patients and diagnosis

From our set of 55 well-defined multigenerational Dutch families with common types of migraine, 7 families were selected in which nearly all patients had attacks of MO only (table 1). This selection was made based on the assumption that a more homogeneous clinical phenotype increases the power to detect linkage[26]. A very large *family 1* and smaller *families 2-7* were selected. In these families MO was inherited as an apparent autosomal dominant trait. To avoid bilinear transmission of migraine and accordingly additional genetic heterogeneity, nine family branches from *families 1*, *2*, and

5, in which one or more of the spouses had migraine, were excluded from the linkage analysis. After this exclusion, the families comprised a total of 204 individuals (table 1, figure 1). All individuals were carefully examined personally by experienced neurologists from the outpatient Headache Clinic of the Leiden University Medical Center, using a semi-structured validated migraine questionnaire according to the IHS criteria[3]. The project has been approved by the Committee of Medical Ethics of the Leiden University Medical Center.

Table 1

Description of the seven Dutch MO families that were included in the genome-wide scan.

Family	Total number of persons	Number of generations in the family	Patients with only MO attacks	Patients with MA and MO attacks	Patients with only MA attacks
1	77	4	38	5	1
2	20	3	11	0	0
3	14	3	7	1	0
4	19	4	7	1	1
5	34	4	15	2	0
6	17	3	9	0	0
7	23	4	12	0	1
Total	204		99	9	3

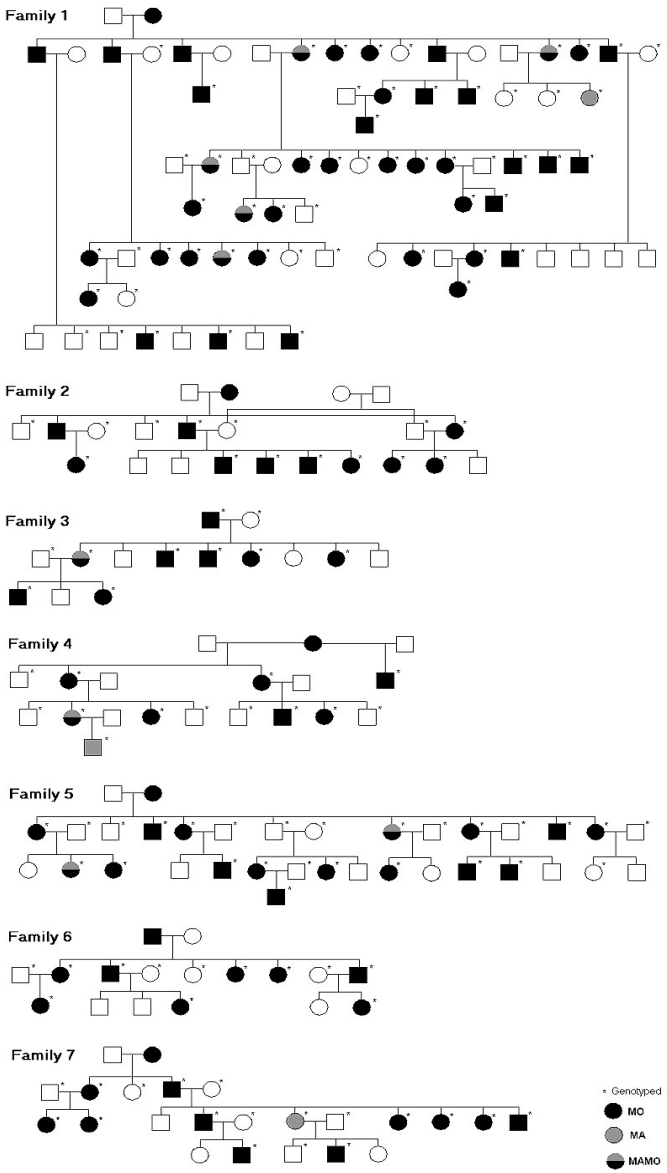
Patients with only migraine with aura (MA) attacks were set to unknown in the linkage analysis. Patients with migraine without aura (MO) and patients with both MA and MO attacks were analyzed as affected in the linkage analysis.

There is a continuing debate whether MO and MA are separate genetic entities. Russell et al. demonstrated that the co-occurrence of MA and MO is not significantly increased in monozygotic compared to dyzygotic twins[27]. This, and the fact that there are some clinical and physiological differences between both types of migraine besides the presence of the aura, indicates that there may be separate genetic factors for MA and MO[28-30]. It was proposed that MO and MA are separate entities, but both migraine types can co-exist in patients[31]. However, a common genetic background is supported by the fact that there is a large clinical overlap between MA and MO[28,32,33]. Furthermore,

in some linkage studies, linkage was observed with patients having either MA or MO, as well as patients that have both types of migraine[19,25].

Figure 1

Pedigrees of MO families that were used for linkage analysis.



Individuals indicated in black are affected with MO. Presence of mixed MO and MA attacks are indicated by half grey–black circles. Persons with MA are presented in grey. Genotyped persons are indicated by an asterix.

In our study the persons having only MO attacks, or co-existing MO and MA attacks, were labeled as "affected" for linkage analysis. In total there were 99 MO patients and 9 patients with mixed MA and MO attacks in the families. Three patients having only MA attacks were labeled unknown for the linkage analysis (table 1).

DNA Analysis

For linkage analysis, the most informative DNA samples from each family were selected for genotyping in the genome scan. Unaffected siblings, not required for reconstruction of untyped individuals, were excluded from the families. From 204 individuals a total of 149 persons were selected for genotyping (figure 1). Genomic DNA was extracted from peripheral blood lymphocytes[34]. Two additional CEPH samples (133101, 133102) were tested for standardization of allele labeling. All samples were genotyped by the Marshfield Mammalian Genotyping Service (Dr. J.L. Weber, WI, USA). Detailed information on genotyping methods, instruments, and software that was used can be found at the website of the Center for Medical Genetics, Marshfield Medical Research Foundation (http://research.marshfieldclinic .org/genetics). In total, 392 highly polymorphic repeat markers, 375 autosomal and 17 X-chromosomal, were genotyped. The markers had an average spacing of ~ 9 cM (Kosambi). The sex-averaged Marshfield '98 Kosambi map was used throughout the study[35]. Allele frequencies for the markers were calculated from the genotypes of all individuals in the sample set.

Statistical Analysis

Prior to linkage analysis, pedigree genotype data were checked for possible sample switches using the graphical relationship representation (GRR) program[36]. The loop in *family 2* was broken between the married-in brother and sister in all analyses. Errors of Mendelian inheritance were tested with the UNKNOWN program of the LINKAGE 5.1 software package[37]. Sample

switches were excluded and in the few cases with Mendelian inheritance problems, the individuals' genotypes were set to unknown.

Given the apparent dominant inheritance pattern of MO in our families, the data was analyzed using model-based linkage analysis. Patients with MO attacks, or co-existing MO and MA attacks, were labelled as "affected", as stated previously. For the linkage model, used in all analyses, the disease frequency based on current data of MO prevalence was set at 0.05 [2]. Furthermore, the penetrance was set at 0.10 for non-carriers and 0.70 for carriers of one or two copies of the disease allele. Linkage analysis was performed with the MLINK program of the LINKAGE 5.1 software package using an affected-only approach[37,38]. Two-point LOD scores, single marker *vs.* disease, were calculated for every marker for the individual families, and for the data of all families combined. The linkage findings were evaluated with simulation studies using the SLINK and MSIM programs[38-40]. For every MO family, 20 000 replicates were generated and analyzed using the same linkage model with an 'average' genetic marker having five alleles, unlinked and linked at a recombination rate of 5%. The maximum two-point LOD scores and LOD score distributions obtained from the simulations were used for comparison with the observed LOD scores. Furthermore, HLOD scores were calculated for every marker with the combined data of all families under the assumption of two-locus heterogeneity using the program HOMOG[38,41]. All (H)LOD scores higher than 1.00 were reported in table 2.

For the regions that showed increased two-point LOD scores, multipoint LOD scores were calculated with the program VITESSE 2.0 [42,43]. Evidence for linkage under all conditions was assumed, following the criteria of Lander and Kruglyak[44]. For LOD scores detected under homogeneity the thresholds are: a LOD score of 3.30 ($p = 0.000049$) for significant linkage, a LOD score of 1.90 for suggestive linkage ($p = 0.0017$) and a LOD score of 0.59 ($p = 0.05$) for nominal linkage. For the heterogeneity analysis the same p-values were

considered for significant (HLOD = 3.60), suggestive (HLOD = 2.20) and nominal (HLOD = 0.84) linkage [38,45]. Nominal p-values for the LOD scores were calculated and reported, following the methods as described by Nyholt[45]. The nominal p-values agreed extremely well with the empirical p-values found with the simulated data.

Results

Seven multigenerational Dutch migraine families were selected for linkage analysis, in which MO was the predominant migraine type (figure 1). In these families migraine was transmitted in an apparent autosomal dominant fashion. Of all 204 family members that were interviewed for this study, 111 suffered from migraine. By far the majority suffered from MO attacks exclusively, 9 patients suffered also from MA attacks and in only 3 patients migraine attacks were diagnosed as MA exclusively (table 1). DNA samples of 149 family members were used for a 9 cM genome-wide scan with 392 polymorphic genetic markers in order to identify migraine loci.

Two-point linkage analyses individual families

First, each family was analyzed for linkage separately. The two-point model-based linkage analysis results of markers showing a LOD score higher than 1.00 are summarized in table 2. No locations were found with significant or suggestive evidence for linkage. In *family 1* the highest LOD score of 1.55 was found for marker D1S1728 on chromosome 1p22. In the same family, the marker D4S2361 gave a LOD score of 1.23 on chromosome 4q21, in the region of the Icelandic MO locus. None of the markers in this family reached the maximum simulated LOD score of 3.33 assuming a single gene. In *family 2* the highest LOD score of 1.25 (identical to the maximum expected LOD score) was found on chromosome Xq21-q22 for markers DXS6789 and ATA31E12. No positive LOD scores were observed for the autosomal chromosomes in this family. However, given the fact that male-to-male transmission of migraine was observed three times in one branch of the family,

the results should be taken with extreme caution. Still, the involvement of an X-linked gene may still be possible because of the complexity of the disorder. In *families 3* and *4* new locations with markers nearly reaching the maximum expected LOD scores were found on chromosomes 2q14, 3q13 and 5p13 (table 2). In *families 6* and *7* no markers with LOD scores higher than 1.00 were identified.

Table 2

Results of two-point linkage analysis for the individual families.

Family	Chromosome	Marker	Position in cM[a]	LOD score[b]	p value	θ max[c]	Expected max LOD[d]
1	1	D1S1665	102	1.08	0.013	0.00	3.33
1	1	D1S1728	109	1.55	0.004	0.00	
1	3	D3S2418	216	1.31	0.070	0.00	
1	4	D4S2361	93	1.23	0.009	0.00	
1	4	D4S2368	168	1.11	0.012	0.00	
1	6	D6S1040	129	1.09	0.013	0.00	
2	X	DXS6789	63	1.25	0.008	0.00	1.25
2	X	ATA31E12	67	1.25	0.008	0.00	
3	2	D2S1328	133	1.03	0.015	0.00	1.04
3	5	D5S1457	59	1.03	0.015	0.00	
4	3	D3S3045	124	1.07	0.013	0.00	1.19
5	12	GATA49D12	18	1.10	0.012	0.00	2.09

Only LOD scores of 1.00 and above are shown.

[a] Sex-averaged map position in cM (Kosambi) based on the Marshfield '98 marker map.

[b] Two-point linkage maximum LOD score.

[c] Optimal recombination fraction θ for which the LOD score was found.

[d] Maximum expected LOD score assuming linkage to a single gene at θ = 0.05.

Two-point linkage analysis with all families combined

Two-point LOD scores were calculated for the data of all families combined. Heterogeneity analysis, taking into account that not all families need to be linked to the same locus, was performed as well. The markers of both analyses with (H)LOD scores higher than 1.00 are presented in table 3. None of the resulting markers showed significant or even suggestive evidence for linkage.

The highest LOD score of 1.64 was again found for marker D4S2361 at 93 cM on chromosome 4q21, which was previously implicated in migraine. Two additional markers, D4S2367 (LOD = 1.13) at 78 cM and D4S1647 (LOD = 0.88) at 105 cM, in the 4q13-q28 region also showed positive LOD score values. In addition, markers with LOD scores above 1.00 were found on chromosomes 1p31, 2q14 and 11p15 (table 3). The only marker that showed an increase in LOD score using heterogeneity analysis was marker D2S1328 (LOD = 0.88; HLOD = 1.00). The estimated fraction of contributing pedigrees, α, for this marker was 0.69.

Table 3

Results of two-point linkage analysis for the combined dataset of all families analyzed under homogeneity and heterogeneity.

Chromosome	Marker	Position in cM[a]	LOD score	p value	θ max[b]	HLOD score	p value	α[c]	θ max[b]
1	D1S1665	102	1.34	0.006	0.00	1.34	0.013	1.00	0.00
2	D2S1328	133	0.88	0.022	0.00	1.00	0.032	0.69	0.00
4	D4S2367	78	1.13	0.011	0.00	1.13	0.023	1.00	0.00
4	D4S2361	93	1.64	0.003	0.03	1.64	0.006	1.00	0.03
11	D11S1981	21	1.16	0.010	0.00	1.17	0.021	1.00	0.00

Only (H)LOD results of 1.00 and above are shown.

[a] Sex-averaged map position in cM (Kosambi) based on the Marshfield '98 marker map.

[b] The optimal recombination fraction θ for which the maximum LOD score was found.

[c] Estimated proportion of families adding to the HLOD score.

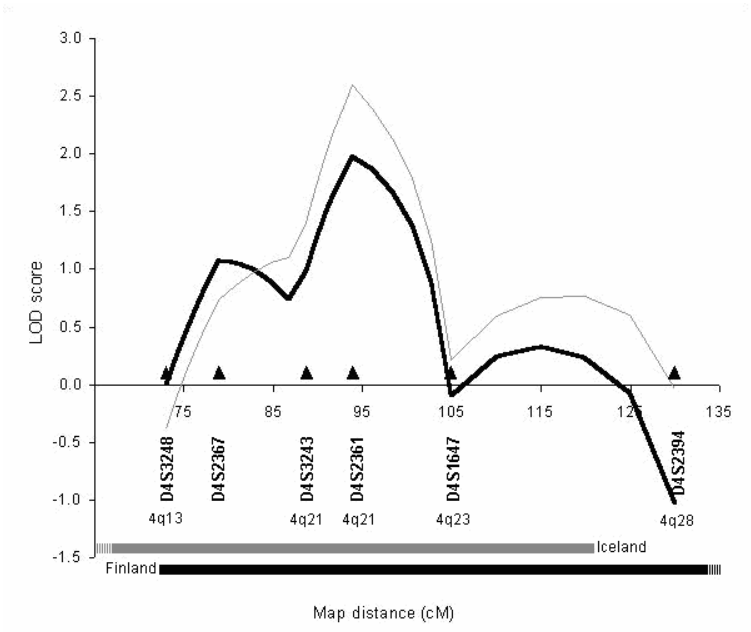
Multipoint analysis of all families

The positive results in chromosomal region 4q13-q28 between 73 and 105 cM prompted us to perform a multipoint analysis for markers D4S3248, D4S2367, D4S3243, D4S2361, D4S1647 and D4S2394 on the data of all families. The results of the multipoint analysis are presented in figure 2 (black line). The maximum LOD score of marker D4S2361 at 93 cM increased from 1.64 in the two-point analysis, to 1.98 (p = 0.0013) in the multipoint analysis. This multipoint LOD score indicates suggestive evidence for linkage, and replicates the Finnish and Icelandic chromosomal 4q21-q24 locus[24,25,44]. Examining the

multipoint LOD scores for marker D4S2361 of the individual families showed that *families 1*, *3*, *5*, and *7* contributed to this positive result. *Families 2*, *4*, and *6* all gave small negative two-point LOD score value. Combining only 4q-positive *families 1, 3, 5,* and *7* increased the multipoint LOD score to 2.59 (p = 0.00027) (figure 2, grey line).

Figure 2

Results of multipoint linkage analysis of our Dutch MO families with 4q13-q28 markers illustrating the region of linkage.



Position of genetic markers is indicated by black triangles. Genetic distances are given in centimorgans (Kosambi). The black line indicates the multipoint analyses of all MO *families 1-7,* whereas the grey line shows multipoint values for 4q-positive *families 1, 3, 5* and *7* combined. The thick horizontal dark grey and black bars below the markers indicate the positive regions of the Icelandic and Finnish loci, respectively.

Multipoint heterogeneity analysis did not indicate locus heterogeneity for the MO families for marker D4S2361 (HLOD = LOD = 2.00). Therefore, the

proportion of families that contributed to the multipoint HLOD score was 100% (estimated as α = 1.00, with a wide 95% confidence interval ranging from 0.05 to 1.00). The discrepancy between the heterogeneity analysis and the higher LOD score for *families 1, 3, 5* and *7* combined can be explained by the fact that the LOD scores of *families 2, 4* and *6* were only slightly negative (-0.36, -0.10, and -0.14, respectively) and could not be definitely excluded for linkage. Furthermore, haplotyping and comparison of the obtained and maximum LOD scores indicated that linkage to the chromosome 4 locus was incomplete (data not shown). In several small branches the shared haplotype was not transmitted. Therefore, there may also be allelic and / or locus heterogeneity within the individual MO families.

Discussion

In this study a genome-wide scan was performed to identify migraine loci in seven Dutch families with an apparent autosomal dominant inheritance pattern of MO. Suggestive evidence for linkage was found in the region 4q21-q24 with multipoint linkage analysis, providing an independent replication of the Finnish MA and Icelandic MO loci[24,25,44]. The Finnish locus spans a large region of approximately 59 cM on chromosome 4q13-q31. The highest LOD score was observed with marker D4S1647 (105 cM) at 4q24. In the Icelandic locus, the highest LOD score was observed at 4q21 with markers D4S1534 (95 cM) and D4S2909 (102 cM). In our study, marker D4S2361 (93 cM) at 4q21 showed the highest LOD score of 1.98 (p = 0.0013) when all families were analyzed, and of 2.59 (p = 0.00027) when only the four 4q-positive families were tested.

The observation that the Finnish, Icelandic and Dutch 4q loci are overlapping, but give highest LOD scores at slightly different locations, is intriguing. An important question therefore is, whether there is one or more than one migraine locus present in this chromosomal region. The fact that in the Finnish study MA patients were studied whereas in the Icelandic and Dutch study MO

was studied supports the idea that there may be different underlying genes. The fact that the inclusion of MO patients in the analysis of the Finnish data set reduced the LOD scores is in favor of this concept. The observation that the Dutch MO locus seems located near the Icelandic MO locus and not so much the Finnish MA locus also suggests that multiple, migraine type-specific, genes might be present on 4q. However, one has to keep in mind that the linkage peaks for all three studies are overlapping and are still very broad.

Future gene identification studies may show that there is only one migraine gene present at chromosomal region 4q21-q24. This may be explained because the causative gene could play a role in a common pathway of MO and MA pathogenesis. Different gene variants in a single gene may lead to different migraine phenotypes. Identification of the causative gene(s) is a huge endeavor but several candidate genes, including a glutamate receptor (GRID2), serine / threonine protein phosphatases (PPP3CA and PPEF2) and CGMP-dependent protein kinase 2 (PRKG2), are located in the 4q21-q24 region.

A second locus that might be worth further investigation was suggested on chromosome X in *family 2*. The maximum possible LOD score of 1.25 was observed for markers DXS6789 (Xq21) and ATA31E12 (Xq22) but provides no significant evidence for linkage. Still, this finding might have some significance because the entire Xq12-q26 shows positive LOD scores and overlaps with the region on the X chromosome that has previously been implicated in migraine[19,46,47]. However, three cases of male-to-male transmission of migraine are present in a branch of *family 2*. Although unlikely, X-chromosomal involvement in *family 2* may still be possible because of the genetic complexity of migraine. The relevance of other chromosomal areas that were identified with nominal linkage should be taken with great caution because they have a high probability of being false-positives. None of these loci coincided with previously reported migraine loci.

Linkage analysis in this study was performed with a conservative model-based approach. Recent linkage reports in selected migraine families have yielded four novel loci for common types of migraine, indicating that the approach is valid[21-25]. Alternative model-free linkage methods were not considered mainly because of the clear inheritance pattern and size of the Dutch MO families. No gain in locus detection power was expected using such an approach[48,49]. Our results have shown that reducing heterogeneity can lead to positive linkage results for a common disease like MO. Using a selection of large MO families with a dominant inheritance pattern, we found suggestive evidence for linkage to a locus on chromosome 4q21-q24, providing an independent replication of the Icelandic and Finnish linkage results. Future studies aiming at identifying underlying causative gene variants may provide further insight in the pathophysiological pathways of migraine.

Acknowledgements

We would like to thank the families for their co-operation. Many grateful thanks also go to Lodewijk A. Sandkuijl who deceased in December 2002. Genotyping was performed at the NHLBI Mammalian Genotyping Service (Dr. J.L. Weber, Center for Medical Genetics, Marshfield, WI, USA) with a grant of National Institute of Health (NIH). This work was also supported by grants of the Netherlands Organization for Scientific Research (NWO), The Migraine Trust, and the European Community (EC-RTN1-1999-00168).

References

1. Lipton RB, Stewart WF. Migraine headaches: epidemiology and co morbidity. Clin Neurosci 1998;5:2-9.
2. Launer LJ, Terwindt GM, Ferrari MD. The prevalence and characteristics of migraine in a population-based cohort: the GEM study. Neurology 1999;53:3:537-42.
3. "International Headache Society". Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. Cephalalgia 1988;(supplement 7) 8:1-96.
4. Russell MB, Olesen J. Increased familial risk and evidence of genetic factor in migraine. BMJ 1995;311:541-44.
5. Russell MB, Iselius L, Olesen J. Migraine without aura and migraine with aura are inherited disorders. Cephalalgia 1996;16:305-9.

6. Ziegler DK, Hur YM, Bouchard TJ et al. Migraine in twins raised together and apart. Headache 1998;38:417-22.
7. Gervil M, Ulrich V, Kaprio J et al. The relative role of genetic and environmental factors in migraine without aura. Neurology 1999;53:5:995-9.
8. Rasmussen BK, Olesen J. Migraine with aura and migraine without aura: an epidemiological study. Cephalalgia 1992;12:221-8.
9. Ophoff RA, Terwindt GM, Vergouwe MN et al. Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. Cell 1996;87:543-52.
10. Joutel A, Ducros A, Vahedi K et al. Genetic heterogeneity of familial hemiplegic migraine. Am J Hum Genet. 1994;55:1166-72.
11. Terwindt GM, Ophoff RA, van Eijk R et al. Involvement of the CACNA1A gene containing region on 19p13 in migraine with and without aura. Neurology 2001;56:8:1028-32.
12. Lea RA, Curtain RP, Hutchins C et al. Investigation of the CACNA1A gene as a candidate for typical migraine susceptibility. Am J Med Genet 2001;105:707-12.
13. McCarthy LC, Hosford DA, Riley JH et al. Single-nucleotide polymorphism alleles in the insulin receptor gene are associated with typical migraine. Genomics 2001;78:135-49.
14. De Fusco M, Marconi R, Silvestri L et al. Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. Nat Genet. 2003;33:192-96.
15. Vanmolkot KR, Kors EE, Hottenga JJ et al. Novel mutations in the Na+, K+-ATPase pump gene ATP1A2 associated with familial hemiplegic migraine and benign familial infantile convulsions. Ann Neurol. 2003;54:360-66.
16. Ducros A, Joutel A, Vahedi K et al. Mapping of a second locus for familial hemiplegic migraine to 1q21-q23 and evidence of further heterogeneity. Ann Neurol 1997;42:885-90.
17. Gardner K, Barmada MM, Ptacek LJ, Hoffman EP. A new locus for hemiplegic migraine maps to chromosome 1q31. Neurology 1997;49:5:1231-8.
18. Lea RA, Shepherd AG, Curtain RP et al. A typical migraine susceptibility region localizes to chromosome 1q31. Neurogenetics 2002;4:1:17-22.
19. Nyholt DR, Dawkins JL, Brimage PJ et al. Evidence for an X-linked genetic component in familial typical migraine. Hum Mol Genet 1998;7:459-463.
20. Nyholt DR, Curtain RP, Griffiths LR. Familial typical migraine: significant linkage and localization of a gene to Xq24-28. Hum Genet 2000;107:18-23.
21. Cader ZM, Noble-Topham S, Dyment DA et al. Significant linkage to migraine with aura on chromosome 11q24. Hum Mol Genet. 2003;12:2511-7.
22. Carlsson A, Forsgren L, Nylander PO et al. Identification of a susceptibility locus for migraine with and without aura on 6p12.2-p21.1.Neurology 2002;59:11:1804-7.
23. Soranga D, Vettori A, Carraro G et al. A locus for migraine without aura maps on chromosome 14q21.2-q22.3. Am J Hum Genet. 2003;72:161-7.
24. Wessman M, Kallela M, Kaunisto MA et al. A susceptibility locus for migraine with aura, on chromosome 4q24. Am J Hum Genet 2002;70:652-62.
25. Bjornsson A, Gudmundsson G, Gudfinnsson E et al. Localization of a gene for migraine without aura to chromosome 4q21. Am J Hum Genet. 2003;73:986-93.
26. Brzustowicz LM, Hodgkinson KA, Chow EW et al. Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22. Science 2000;288:678-82.
27. Russell MB, Ulrich V, Gervil M, Olesen J. Migraine without aura and migraine with aura are distinct disorders. A population-based twin survey. Headache 2002;42:332-6.

28. Kallela M, Wessman M, Farkkila M et al. Clinical characteristics of migraine in a population-based twin sample: similarities and differences between migraine with and without aura. Cephalalgia 1999;19:151-8.
29. Olesen J, Tfelt-Hansen P, Henriksen L, Larsen B. The common migraine attack may not be initiated by cerebral ischaemia. Lancet 1981;29:438-40.
30. Olesen J, Larsen B, Lauritzen M. Focal hyperemia followed by spreading oligemia and impaired activation of rCBF in classic migraine. Ann Neurol 1981;9:344-52.
31. Kallela M, Wessman M, Havanka H et al. Familial migraine with and without aura: clinical characteristics and co-occurrence. Eur J Neurol 2001;8:5:441-9.
32. Blau JN. Migraine with aura and migraine without aura are not different entities. Cephalalgia 1995;5:186-90.
33. Rapoport MA. MO and MA: distinct clinical entities? Cephalalgia 1996;16:215.
34. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16:1215.
35. Broman KW, Murray JC, Sheffield VC et al. Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet. 1998;63:861-9.
36. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. Bioinformatics 2001;17:742-3.
37. Lathrop GM, Lalouel JM. Easy calculations of lod scores and genetic risks on small computers. Am J Hum Genet 1984;36:460-5.
38. Ott J, Terwilliger JD. Handbook of human genetic linkage. Baltimore, MD: Johns Hopkins University Press. 1994
39. Weeks DE, Ott J, Lathrop GM. SLINK: a general simulation program for linkage analysis. Am J Hum Genet 1990;47:A204 Supplement.
40. Ott J. Computer-simulation methods in human linkage analysis. Proc Natl Acad Sci USA 1989;86:4175-8.
41. Bhat A, Heath SC, Ott J. Heterogeneity for multiple disease loci in linkage analysis. Hum Hered 1999;49:229-31.
42. O'Connell JR, Weeks DE. The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. Nature Genetics 1995;11:402-8.
43. O'Connell JR. Rapid Multipoint Linkage analysis via inheritance vectors in the Elston-Stewart algorithm. Hum Hered 2001;51:226-40.
44. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 1995;11:241-247.
45. Nyholt DR. All LODs are not created equal. Am J Hum Genet. 2000;67:282-8.
46. Oterino A, Monton F, Cid C et al. A new locus for migraine with aura on Xq13. Cephalalgia 2001;21:346.
47. Wieser T, Pascual J, Barmada MM et al. Genetic analysis of the X-chromosome in migraine families. Cephalalgia 2001;21:280
48. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet. 1996;58:1347-63.
49. Hodge SE. Model-free vs. model-based linkage analysis: a false dichotomy? Am J Med Genet 2001;105:62-4.

CHAPTER 7

A Dutch family with 'familial cortical tremor with epilepsy': clinical characteristics and exclusion of linkage to chromosome 8q23.3-q24.1

F van Rootselaar[1], PMC Callenbach[2], JJ Hottenga[3], FLMG Vermeulen[3], HD Speelman[1], OF Brouwer[2,4], MAJ Tijssen[1]

1. Department of Neurology, Academic Medical Center, Amsterdam, the Netherlands.
2. Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands.
3. Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.
4. Department of Neurology, University Hospital Groningen, Groningen, the Netherlands.

J Neurol 2002 249:829-34.

Abstract

Purpose: To describe the clinical characteristics of a large Dutch family with cortical tremor with epilepsy (FCTE) and to test for genetic linkage of FCTE to chromosome 8q23.3-q24.1. *Background:* FCTE is an idiopathic generalized epilepsy of adult onset with autosomal dominant inheritance. It is characterized by kinesigenic tremor and myoclonus of the limbs, generalized seizures, and electrophysiological findings consistent with cortical reflex myoclonus. Genetic analysis has been performed in five Japanese families. In all families, linkage was shown to chromosome 8q23.3-q24.1. *Methods:* Clinical and electrophysiological data of a four-generation family, suspected of autosomal dominantly inherited FCTE, were collected and linkage analysis was performed. *Results:* Clinical and electrophysiological findings were consistent with a diagnosis of FCTE. Of 41 relatives examined, 13 subjects were considered to be definitely affected, 3 were probably affected and 10 were unaffected. In 15 relatives, the diagnosis could not be established. Linkage to chromosome 8q23.3-q24.1 was excluded. *Conclusions:* In this family with autosomal dominant FCTE, specific clinical and electrophysiological features were identified. Exclusion of linkage to chromosome 8q23.3-q24.1 indicates that genetic heterogeneity exists for FCTE.

Keywords

Familial cortical tremor, myoclonus, epilepsy, linkage analysis.

Introduction

'Cortical tremor' was first described in 1990 by Ikeda et al. in two patients with a fine action tremor resembling essential tremor that was unresponsive to β-blockers[1]. Both patients suffered from occasional epileptic seizures. Electrophysiological studies revealed features of cortical reflex myoclonus, such as giant somatosensory evoked potentials (g-SEPs), enhanced long loop reflexes (C-reflexes), and premovement cortical spikes[1]. It was concluded that the tremor originated from the cerebral cortex and should be designated as a variant of cortical reflex myoclonus[1,2]. Subsequently, cortical tremor has been described in both sporadic and familial cases with autosomal dominant inheritance[3-7]. Five Japanese and one European family have been described so far with this syndrome, named 'familial cortical tremor with epilepsy' (FCTE), 'familial adult myoclonic epilepsy' (FAME) or 'benign adult familial myoclonic epilepsy' (BAFME)[5-7]. In these pedigrees, the disorder was characterized by a non-progressive 'essential-tremor-like' tremor, infrequent seizures and in some cases myoclonus. Results of electrophysiological studies were consistent with cortical reflex myoclonus, and electroencephalograms (EEG) showed spikes, spike-wave complexes, and polyspike-wave complexes[1,2,5-7]. In the European family, mental retardation was an additional finding[4]. Genetic analysis in five Japanese families showed linkage to chromosome 8q23.3-q24.1 [6,7].

Recently, we have identified a Dutch family with cortical tremor and epilepsy. As far as we know, this is the largest pedigree described until now. The objectives of the present study were (1) to describe the clinical and electrophysiological characteristics of this family; and (2) to examine whether linkage to chromosome 8q23.3-q24.1 could be established. Knowledge of the genetic basis of the syndrome might give insight into the pathogenesis of FCTE and could be a first step towards a specific treatment.

Methods

Patients

After having obtained written informed consent from 26 relatives and spouses (figure 1, pedigree), medical and family histories were taken and venous blood samples for genetic testing were drawn (by FvR and MT). In all participating relatives and spouses, special attention was given to tremor and epilepsy. If relatives had ever visited a neurologist before, existing clinical data were obtained (II:3, 7, 11, 13; III:3, 5, 10, 19 and 20). If possible, adult relatives with a tremor participated in electrophysiological studies (III:1, 5, 10; IV:1 and 2). Other diseases with tremor and epilepsy (progressive myoclonus epilepsies, MERRF, and spinocerebellar ataxias) were excluded or made unlikely with magnetic resonance imaging of the brain in III:3 and 10, normal lactate and pyruvate levels and exclusion of spinocerebellar ataxia (SCA) types 1, 2, 3, 6 and 7 in patient III:10, and no ragged-red-fibres in a muscle biopsy in patient III:3.

Electrophysiology

Electrophysiological measurements included surface electromyography (EMG) with tremor registration and long-latency reflex recording (C-reflex), somatosensory evoked potentials (SEP), electroencephalography (EEG), and jerk-locked averaging of cortical potentials. Electrophysiological findings were considered positive if a giant SEP (our laboratory standard: P27-N35 > 4μV) in combination with a C-reflex were found. If not, they were considered inconclusive.

Diagnostic criteria

(based on disease characteristics as described by Elia, Okuma, Mikami, and Plaster)[4-7]

Definitely affected: (1) a history of tremor and myoclonus, and myoclonic or epileptic seizures, and on neurological examination characteristic signs: kinesigenic tremor resembling essential tremor and distal action myoclonus of

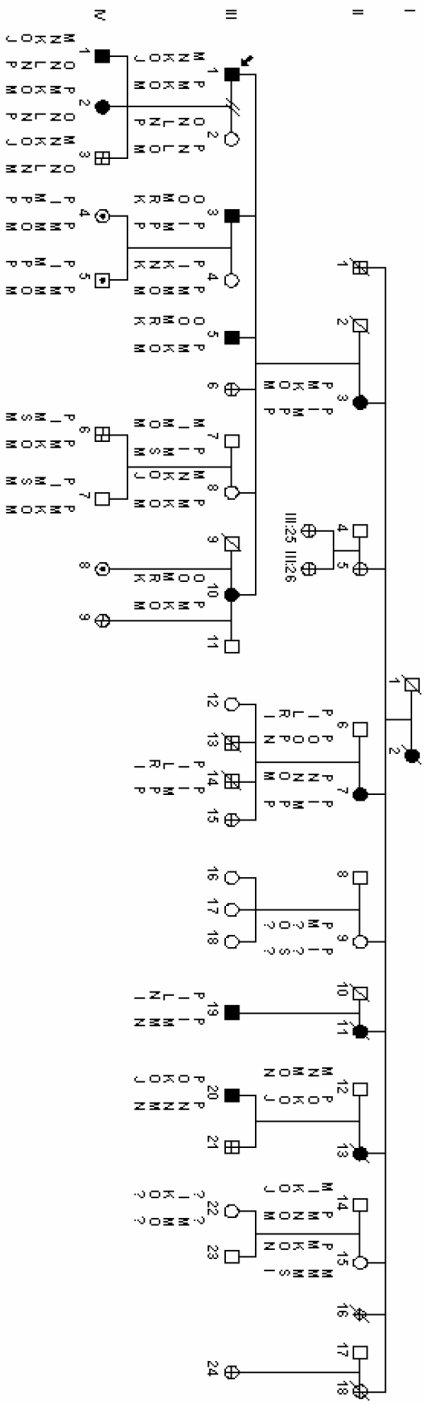


Figure 1

Pedigree of our Dutch family with familial cortical tremor with epilepsy (FCTE). Black symbols are definitely affected persons, and persons with a black dot are probably affected. Diagnosis could not be established in persons with a plus. Regarding the linkage analysis: diagnosis is considered to be unknown in all but the definitely affected persons. The proband is indicated by an arrow.

all limbs; or (2) a history of tremor without myoclonic or epileptic seizures but on examination a characteristic tremor and myoclonus plus positive electrophysiological findings (giant SEP and C-reflex).

Probably affected: a history of tremor, with or without use of anti-epileptic drugs. On examination a characteristic tremor and myoclonus but inconclusive or absent electrophysiological data.

Possibly affected: a history of tremor, no anti-epileptic drug treatment and on examination no characteristic tremor. Inconclusive or no electrophysiological data.

Unaffected: 50 years of age or older, without anti-epileptic drugs. No characteristic signs on examination.

No diagnosis: inconclusive or incomplete data.

Genotyping

Venous blood samples were taken from 10 definitely affected relatives and 16 other relatives and spouses. Genomic DNA was extracted from peripheral lymphocytes using standard methods[8]. Microsatellite markers D8S1784, D8S1779, D8S1694, D8S514, and D8S1720, all from the chromosome 8q23.3-q24.1 region, were tested by polymerase chain reaction (PCR). Oligonucleotide sequences are available through the Human Genome Database (GDB). PCRs for all markers were performed using standard conditions (www.gdb.org). PCR products for each template were pooled and an aliquot was loaded onto a 6% standard denaturing polyacrylamide gel and run in an Applied Biosystems (ABI) 377 automated DNA sequencer. Allele sizes were determined on the basis of an internal standard size marker, using GeneScan 3.1 and Genotyper 2.5 ABI software. Genotypes were determined by two individuals, and checked for Mendelian segregation using a standard program.

Linkage analysis

Single- and multipoint LOD score analysis were performed using the Linkage program, version 5.1, with an affected-only model. In the linkage analysis, only definitely affected relatives were regarded as affected. Diagnosis in all other relatives - probably affected, possibly affected and unaffected - was considered unknown[9]. The syndrome was considered to be autosomal dominantly inherited with 80% penetrance, no phenocopies, a gene frequency of 0.001, and equal allele frequencies for each individual marker. Multipoint analysis was performed with the following genetic positions of the microsatellite markers (in cM), according to the database of Généthon: D8S1784: 116.8; D8S1779: 122.6; D8S1694: 124.2; D8S514: 128.9; D8S1720: 139.7.

Results

Clinical and electrophysiological findings

Of the 41 relatives examined, 13 relatives were considered definitely affected, 3 were probably affected, and 10 relatives were unaffected (figure 1). In 15 relatives, diagnosis could not be established owing to the lack of data. The pedigree showed an autosomal dominant inheritance pattern (figure 1).

The 55-year-old proband (III:1) suffered from tremor, finger twitching and trembling of his legs from the age of 45 years. The involuntary movements were provoked by action and emotional stress, and were most intense after awaking. Propranolol was not effective but primidon reduced the tremor. At the age of 52 years, he had his first generalized seizure. On neurological examination, a kinesigenic tremor was observed with superimposed myoclonus of the fingers and toes, which made the tremor irregular. There were no signs of ataxia, nystagmus, or dysarthria. In a clinical setting, the primidon was slowly reduced. The tremor and myoclonus increased markedly and the patient suffered from a tonic-clonic seizure at night. EEG showed left parietal spike-wave complexes. Interestingly, intravenous clonazepam both

ceased the tremulous movements and normalized the EEG. Electrophysiological measurements were performed after initiated treatment with sodium valproate (2000 mg daily dose) and clobazam (10 mg daily dose). Tremor-registration showed an irregular 12-16 Hz tremor. C-reflex was not detected. Giant SEPs were measured on both arms (right P27-N30 = 6.5μV; left = 6.3μV). Jerk-locked averaging was not possible owing to the infrequency of myoclonic jerks and the interference of the tremor.

In general, tremulous movements started between the ages of 12 and 45 years (patients I:2; II:3, 7, 11, 13; III:1, 3, 5, 10, 19, 20; IV:1, 2, 4, 5, and 8, table 1). Patients could not differentiate between tremor and fine myoclonic movements. Generalized tonic-clonic seizures and myoclonic seizures started between the ages of 20 and 63 years, 1 to 33 years after the tremor. Severity of tremor, myoclonus, and seizures varied between individuals. The symptoms tended to be slightly to moderately progressive, leading to mild (III:1) to severe (II:7) handicap. Several affected relatives with seizures had complaints of memory deterioration (II:3, 7; III:3, 10).

On examination (patients II:3, 7; III:1, 3, 5, 10, 19, 20; IV:1, 2, 4, 5, and 8, table 1), a characteristic kinesigenic tremor, resembling essential tremor, and myoclonus was seen in fingers, arms, feet, and even in legs. The tremulous movements were provoked by exercise and emotional stress. Severity varied between patients and during the day, usually being worse in the morning. In addition, signs of slight cognitive impairment, such as short-term memory and attention deficits, were noticed in four patients with epilepsy (II:3, 7; III:3, 10) and reported in the medical notes of two of the deceased patients (II:11, 13). All these patients were on anti-epileptic medication: II:3, sodium valproate 1200 mg and phenytoin 800 mg dd; II:7, carbamazepine 1000 mg, phenytoin 1200 mg, phenobarbital 100 mg, clonazepam 2 mg; III:3, clonazepam 1 mg, sodium valproate 900 mg; III:10, sodium valproate 1300 mg, clonazepam 1 mg; II:11 unknown; II:13, phenobarbital 150 mg, sodium valproate 1250 mg,

clobazam 50 mg. Detailed neuropsychological testing has not been performed. There were no other neurological signs or symptoms in the three probably and ten definitely and living affected members, especially no signs of cerebellar dysfunction. MRI of the brain of two patients (III:3, 10) showed slight cerebellar atrophy but no atrophy of brainstem or basal ganglia. Many anti-epileptic drug regimens had been tried, with clonazepam and sodium valproate being the most successful in diminishing the tremulous movements and the frequency of myoclonic and tonic-clonic seizures.

Table 1

Clinical and electrophysiological features of definitely and probably affected relatives.

Patient	FCTE	Age	Sex	T	M	MS	TCS	EEG	SEP, right/left (μV)		C-reflex
									P14-N20	P27-N30	
I:2	D	†73	F	30	+	?	63	n.d.	n.d.	n.d.	n.d.
II:3	D	76	F	40	+	44?	44	n.d.	n.d.	n.d.	n.d.
II:7	D	72	F	+	+	37?	37	n.d.	n.d.	n.d.	n.d.
II:11	D	†67	F	+	+	?	+	n.d.	n.d.	n.d.	n.d.
II:13	D	†51	F	+	?	?	43	n.d.	n.d.	n.d.	n.d.
III:1	D	55	M	45	+	-	52	spike-wave, [irregular]	[0.4/0.1]	[6.5/6.3]	[-]
III:3	D	54	M	25	+	36	44	n.d.	n.d.	n.d.	n.d.
III:5	D	52	M	30	+	30	43	[spike-wave]	[1.2/0.8]	[3.8/3.9]	[+]
III:10	D	42	F	38	+	-	42	irregular	1.1/0.6	13.9/7.2	+
III:19	D	40	M	19	+	20	20	n.d.	n.d.	n.d.	n.d.
III:20	D	43	M	12	+	13?	31	n.d.	n.d.	n.d.	n.d.
IV:1	D	29	M	22	+	-	-	normal	0.6/0.6	5.4/4.7	+
IV:2	D	25	F	20	+/-	-	-	normal	0.5/0.6	9.4/12.5	+
IV:4	P	28	F	+	+	-	-	n.d.	n.d.	n.d.	n.d.
IV:5	P	24	M	20	+	-	-	n.d.	n.d.	n.d.	n.d.
IV:8	P	16	F	12	+	-	-	n.d.	n.d.	n.d.	n.d.

FCTE = familial cortical tremor with epilepsy, D = definite, P = probable, T = tremor, M = myoclonus, MS = myoclonic seizure, TCS = tonic-clonic seizure, EEG = electroencephalogram, SEP = sensory evoked potential, Pxx-Nxx = amplitude difference between Pxx-peak and Nxx-peak, C-reflex = cortical reflex,† = age of death, F = female, M = male, 30 = 30 years of age at onset, + = present, - = not observed, ? = unknown, ± = subtle, n.d. = not done, sw = spike-wave complexes, irr = irregular, n = normal, […] = on anti-epileptic drugs, P27-N30>4.0μV: giant potential.

Tremor-recording showed an irregular 10-16 Hz tremor (III:1, 5, 10; IV:1, 2). C-reflexes and SEPs were studied in five persons (III:1, 5, 10; IV:1, 2, table 1). In two definitely affected members on medication, a C-reflex (III:1), or a giant potential (III:5) could not be detected. EEG examination showed individual differences between patients. Parietal spike-wave complexes but no photosensitivity was found in patient III:1; frontotemporal spike-wave complexes and epileptic changes during photo stimulation in patient III:5; and no abnormalities in patients IV:1 and 2. Back averaging showed no premovement cortical spikes.

Diagnostic classification was difficult in two family members. Patient IV:4 had suffered from infrequent complex partial seizures at the age of nine years, sometimes evolving into generalized seizures. Carbamazepine was effective and was discontinued after a seizure-free period of four years. There was no history of tremor or myoclonus. However, neurological examination at the age of 28 years showed a characteristic tremor and myoclonus. The epilepsy seemed not to be related to FCTE, but based on the clinical symptoms she was considered 'probably affected'. The other relative, a 26-year-old male (IV:6) suffered from a tremor. Neurological examination showed a fine postural tremor of both hands without myoclonus. The EEG showed focal changes but no distinct epileptiform discharges. C-reflex and giant SEPs could not be registered. His 50-year-old mother (III:8) was clinically unaffected but she may represent a non-penetrant carrier of the gene. He was therefore classified as possibly affected.

Linkage analysis

Single-point LOD scores between the selected markers on chromosome 8q23.3-q24.1 and FCTE are shown in table 2. All LOD scores were negative, suggesting absence of linkage. The multipoint LOD scores for the markers on chromosome 8q23.3-q24.1 showed exclusion of almost the entire region between markers D8S1784 and D8S1720 (LOD score < -2.0, figure 2). Only a

region of approximately 5 cM between marker D8S514 and D8S1720 did not fulfill the formal criterion for exclusion (maximum LOD score -1.91). This region was, however, outside the FCTE locus, as described in the other FCTE families[6,7].

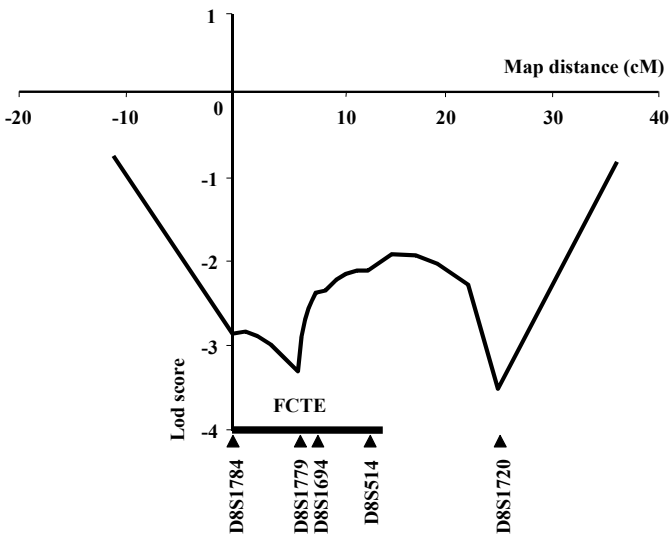
Table 2

Single-point LOD scores between FCTE and markers on chromosome 8q.

	Recombination fraction (θ)					
Marker	0.00	0.01	0.05	0.10	0.15	0.190
D8S1784	-1.715	-0.948	-0.372	-0.157	-0.061	-0.021
D8S1779	-3.326	-2.518	-1.589	-0.980	-0.611	-0.412
D8S1694	-1.463	-1.196	-0.740	-0.484	-0.329	-0.241
D8S514	-1.524	-1.486	-1.261	-0.879	-0.573	-0.394
D8S1720	-3.573	-2.428	-1.348	-0.837	-0.548	-0.391

Figure 2

Results of multipoint linkage analysis to exclude the known FCTE locus in our family with microsatellite markers from chromosome 8q23.3-q24.1, illustrating the region of exclusion.



Discussion

Our clinical findings are in line with previous studies and suggest that this epileptic syndrome can be distinguished from the other epileptic syndromes[1,4-7]. It is characterized by (1) autosomal dominant inheritance, (2) intention-like tremor and distal myoclonus, (3) myoclonic and generalized seizures, (4) late onset, (5) moderate progressive course, (6) no pyramidal or cerebellar signs, or features of a neurodegenerative disorder, (7) electrophysiological features of cortical hyperexcitability, (8) good response to anti-epileptic drugs, especially clonazepam and valproate. The clinical criteria based on the previous reports on FCTE could be confirmed in the Dutch patients.

As it has been described before, FCTE can be differentiated from other inherited cerebellar and epileptic syndromes[4-7]. Patients with autosomal dominant cerebellar ataxias can present with a tremor but the lack of cerebellar signs, absence of known SCA-mutations and only slight cerebellar atrophy on MRI-scans make such a diagnosis unlikely. Seizures in juvenile myoclonic epilepsy have an earlier onset and occur predominantly shortly after awakening[10,11]. Finally, the progressive myoclonus epilepsies with neurological deterioration consisting of cerebellar impairment and higher neurological dysfunction should be considered[10,11]. The clinical course and the autosomal dominant inheritance pattern of the family presented make this diagnosis unlikely. The familial transmission pattern is not consistent with mitochondrial encephalomyopathy with ragged-red-fibres (MERRF). This was further excluded by histochemical examination of a muscle biopsy, as typical ragged-red-fibres were not observed.

Linkage to chromosome 8q23.3-q24.1 described in Japanese pedigrees with FCTE was excluded in the Dutch pedigree[6,7]. The slightly different phenotype to previous descriptions of FCTE might explain this[1,4-7]. The clinical picture of the Dutch patients closely resembles the features described in other known FCTE pedigrees. However, some clinical differences between pedigrees are

noticed and have been previously reported[4]. In contrast to our family, where patients always presented with tremor, patients reported by Okuma et al. (three families, seven affected members), presented either with tremor or epilepsy[5]. In addition, cognitive deterioration seems to be a symptom of Dutch patients in a more progressed stage. The negative effect of anti-epileptic drugs cannot be ruled out and further neuropsychological investigations are necessary. Interestingly, mental retardation was considered a disease feature in another European family (one family, seven affected members)[4]. This was only described in the third generation in combination with a more severe clinical picture suggesting anticipation[4].

The observed typical tremulous myoclonus on EMG, the presence of C-reflexes and giant SEPs, and the immediate suppression of both tremulous movements and left parietal spike-wave complexes by clonazepam are in line with the diagnosis FCTE[1,4-7]. The amplitudes of the SEP were less than 4 μV in one definitely affected patient, due to the use of anticonvulsants[1,5]. In contrast to other studies, EEG back-averaging was not a diagnostic tool in our family[1,3-6]. In four Japanese families, back-averaging was not discussed[7]. The lack of a cortical correlate in the Dutch family is most likely due to the infrequency of myoclonic jerks and interference of the tremor. The tremor on which the myoclonic jerks are superimposed appears to have no cortical correlate, which makes the cortical origin of the tremor doubtful. Further detailed electrophysiological studies are needed to clear up the origin of the tremor. If the tremor does not originate in the cortex the name ‘familial cortical tremor and epilepsy’ should be reconsidered and changed in ‘familial myoclonus and epilepsy’.

The electrophysiological studies in FCTE are consistent with a lack of cortical inhibition. The main inhibitory neurotransmitter in the central nervous system, including the cerebellum, is γ-aminobutyric acid (GABA). Cortical myoclonus is associated with cerebellar changes, and changes in cerebellar inhibitory

function could be due to a changed GABA receptor function[12]. Supportive for this hypothesis is that anti-epileptic drugs that affect the GABA receptor function were most effective and normalized the electrophysiological findings at the same time in FCTE. Genes encoding GABA receptor subunits are, therefore, good candidate genes for FCTE. The GABA receptor is a chloride channel receptor.

In idiopathic epilepsy syndromes and other paroxysmal neurological disorders, several mutations in genes encoding ion channels have been described: mutations in genes encoding subunits of the neuronal nicotinic acetylcholine receptor in patients suffering from autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE) mutations in potassium channel genes in benign familial neonatal convulsions (BFNC) and mutations in sodium channel genes in generalized epilepsy with febrile seizures plus ($GEFS^+$)[13-20]. A channelopathy is most likely the cause of the epileptic syndrome FCTE. In conclusion, we describe a large Dutch pedigree with the typical clinical and electrophysiological features of FCTE. The lack of linkage to chromosome 8q23.3-q24.1 proves genetic heterogeneity of FCTE. This family gives a unique opportunity to further elucidate the molecular pathways leading to this syndrome.

Acknowledgements

The authors thank F. Baas, H. Koelman, R.A. Ophoff, B. Schmand, and L.A. Sandkuijl for their help. We thank C.T.E. Beljaars, E.N.H. Jansen Steur, S.A.J. de Froe, P.J. Koehler, J.F. Mirandolle, A. van Spreeken, R.T.M. Starrenburg, J.W. Vredeveld, and M.J. Wennekes for referring their patients to our clinic. P.M.C. Callenbach is a research fellow supported by a grant from the Netherlands Organisation for Scientific Research (NWO, 940-33-030) and the National Epilepsy Fund (98-14).

References

1. Ikeda A, Kakigi R, Funai N, et al. Cortical tremor: a variant of cortical reflex myoclonus. Neurology 1990;40:1561-5.
2. Hallett M, Chadwick D, Marsden CD. Cortical reflex myoclonus. Neurology 1979;29:1107-25.
3. Toro C, Pascual-Leone A, Deuschl G, et al. Cortical tremor. A common manifestation of cortical myoclonus. Neurology 1993;43:2346-53.
4. Elia M, Musumeci SA, Ferri R, et al. Familial cortical tremor, epilepsy, and mental retardation: a distinct clinical entity? Arch Neurol 1998;55:1569-73.
5. Okuma Y, Shimo Y, Shimura H, et al. Familial cortical tremor with epilepsy: an under-recognized familial tremor. Clin Neurol Neurosurg 1998;100:75-78.
6. Mikami M, Yasuda T, Terao A, et al. Localization of a gene for benign adult familial myoclonic epilepsy to chromosome 8q23.3-q24.1. Am J Hum Genet 1999;65:745-51.
7. Plaster NM, Uyama E, Uchino M, et al. Genetic localization of the familial adult myoclonic epilepsy (FAME) gene to chromosome 8q24. Neurology 1999;53:1180-3.
8. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16:1215.
9. Lathrop GM, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci U S A 1984;81:3443-6.
10. Roger J, Bureau M, Dravet C, et al. Epileptic syndromes in infancy, childhood and adolescence. London: John Libbey & Company Ltd, 1992.
11. Callenbach PMC, Brouwer OF. Hereditary epilepsy syndromes. Clin Neurol Neurosurg 1997;99:159-71.
12. Tijssen MA, Thom M, Ellison DW, et al. Cortical myoclonus and cerebellar pathology. Neurology 2000;54:1350-6.
13. Steinlein OK, Mulley JC, Propping P, et al. A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. Nat Genet 1995;11:201-3.
14. De Fusco M, Becchetti A, Patrignani A, et al. The nicotinic receptor beta2 subunit is mutant in nocturnal frontal lobe epilepsy. Nat Genet 2000;26:275-6.
15. Phillips HA, Favre I, Kirkpatrick M, et al. CHRNB2 Is the Second Acetylcholine Receptor Subunit Associated with Autosomal Dominant Nocturnal Frontal Lobe Epilepsy. Am J Hum Genet 2001;68:225-31.
16. Biervert C, Schroeder BC, Kubisch C, et al. A potassium channel mutation in neonatal human epilepsy. Science 1998;279:403-6.
17. Singh NA, Charlier C, Stauffer D, et al. A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. Nat Genet 1998;18:25-9.
18. Charlier C, Singh NA, Ryan SG, et al. A pore mutation in a novel KQT-like potassium channel gene in an idiopathic epilepsy family. Nat Genet 1998;18:53-5.
19. Wallace RH, Wang DW, Singh R, et al. Febrile seizures and generalized epilepsy associated with a mutation in the Na+-channel beta1 subunit gene SCN1B. Nat Genet 1998;19:366-70.
20. Escayg A, MacDonald BT, Meisler MH, et al. Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2. Nat Genet 2000;24:343-5.

CHAPTER 8

Familial partial epilepsy with variable foci in a Dutch family: clinical characteristics and confirmation of linkage to chromosome 22q

PMC Callenbach[1], AMJM van den Maagdenberg[1,2], JJ Hottenga[2], EH van den Boogerd[2], RFM de Coo[3], D Lindhout[4], RR Frants[2], LA Sandkuijl[5†], OF Brouwer[1,6].

1. Department of Neurology, Leiden University Medical Center, Leiden, the Netherlands.
2. Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.
3. Department of Pediatric Neurology, Erasmus Medical Center Rotterdam, Rotterdam, the Netherlands.
4. Department of Medical Genetics, University Medical Center Utrecht, Utrecht, the Netherlands.
5. Department of Medical Statistics, Leiden University Medical Center, Leiden, the Netherlands.
6. Department of Neurology, University Hospital Groningen, Groningen, the Netherlands.

† In memory of Lodewijk A. Sandkuijl

Epilepsia 2003 44:1298-305.

Abstract

Purpose: Three forms of idiopathic partial epilepsy with autosomal dominant inheritance have been described: (1) autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE); (2) autosomal dominant lateral temporal epilepsy (ADLTE) or partial epilepsy with auditory features (ADPEAF); and (3) familial partial epilepsy with variable foci (FPEVF). Here, we describe linkage analysis in a Dutch four-generation family with epilepsy fulfilling criteria of both ADNFLE and FPEVF. *Methods:* Clinical characteristics and results of EEG, CT and MRI were evaluated in a family with autosomal dominantly inherited partial epilepsy with apparent incomplete penetrance. Linkage analysis was performed with markers of the ADNFLE (1p21, 15q24, 20q13.3) and FPEVF (2q36, 22q11-q12) loci. *Results:* Epilepsy was diagnosed in ten relatives. Age at onset ranged from three months to 24 years. Seizures were mostly tonic, tonic-clonic, or hyperkinetic with a wide variety in symptoms and severity. Most interictal EEGs showed no abnormalities but some showed frontal, central, and/or temporal spikes and spike-wave complexes. From two patients, an ictal EEG was available, showing frontotemporal abnormalities in one and frontal and central abnormalities in the other. Linkage analysis with the known loci for ADNFLE and FPEVF revealed linkage to chromosome 22q11-q12 in this family. *Conclusions:* The clinical characteristics of this family fulfilled criteria of both ADNFLE and FPEVF. The frequent occurrence of seizures during daytime and the observation of interictal EEG abnormalities originating from different cortical areas were more in agreement with FPEVF. The observed linkage to chromosome 22q11-q12 supported the diagnosis of FPEVF and confirmed that this locus is responsible for this syndrome.

Keywords

FPEVF, ADNFLE, clinical characteristics, genetics, chromosome 22

Introduction

Three forms of idiopathic partial epilepsy syndromes with autosomal dominant inheritance have been described: (1) autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE)[1,2]; (2) autosomal dominant lateral temporal epilepsy (ADLTE) or partial epilepsy with auditory features (ADPEAF)[3,4]; and (3) familial partial epilepsy with variable foci (FPEVF)[5,6] .

ADNFLE is characterized by clusters of brief tonic and hyperkinetic motor seizures occurring mostly during sleep[1,2]. Onset is usually in childhood and seizures often persist throughout adult life with considerable intra-familial variation in severity. The predominant finding of the ictal electroencephalogram (EEG) is bilaterally sharp wave activity localized in the anterior quadrants, whereas interictal EEGs usually do not show diagnostic epileptiform abnormalities. In a number of families and in one isolated patient, three different mutations in the gene encoding the α4-subunit of a neuronal nicotinic acetylcholine receptor (CHRNA4) on chromosome 20q13.3 have been described[7-13]. All these mutations are situated in the second transmembrane region[8,9,11]. Considerable locus heterogeneity was documented for ADNFLE. Two different mutations in the gene encoding the β2-subunit of the neuronal nicotinic acetylcholine receptor (CHRNB2) on chromosome 1p21 were found responsible for the epilepsy in two families with ADNFLE[14-16]. In 1998, another locus was found on chromosome 15q24 but the responsible gene has not yet been identified[17].

ADLTE or ADPEAF is characterized by simple partial seizures with auditory symptoms and secondary generalization[3,18-20]. Sensory and psychic symptoms may also occur. Age at onset is usually in the first two decades of life. In one family, the symptoms were accompanied by aphasia[21]. Interictal EEGs sometimes show temporo-occipital sharp wave activity. Because the seizure semiology strongly suggests a seizure origin in the lateral temporal lobe, the

syndrome has been described as 'autosomal dominant lateral temporal epilepsy' (ADLTE)[4]. The syndrome was linked to chromosome 10q22-q24 [3,4,19-21]. Recently, 11 different mutations in the leucine-rich glioma-inactivated 1 gene (LGI1) on chromosome 10q24 have been found responsible for this syndrome in 11 families, including the one with aphasic seizures[22-27]. The loss of both copies of this gene promotes glial tumor progression, leading to the assumption that this gene might function as a tumor-suppressor gene[28]. The role of LGI1 in the pathogenesis of epilepsy is still unknown. It is the first gene not apparently encoding an ion channel or neurotransmitter receptor that has been identified for a human idiopathic epilepsy syndrome.

Seizures in FPEVF have a wide range of age at onset and are often heterogeneous within families, both clinically and neurophysiological[5,6]. They can be nocturnal or diurnal and may be simple or complex partial, originating from temporal, frontal, occipital or centroparietal areas, with sometimes secondary generalization. Until now, one Australian family and two French-Canadian families with FPEVF have been described[5,6]. In the French-Canadian families seizures were predominantly nocturnal and interictal EEGs were often normal, while most affected relatives of the Australian family had diurnal seizures and abnormalities in the interictal EEG. The frequent occurrence of daytime seizures and the observation of interictal EEG abnormalities originating from different cortical areas distinguish FPEVF from ADNFLE. FPEVF is autosomal dominantly inherited with incomplete penetrance. Suggestive linkage of the syndrome to chromosome 2q36 was observed in the Australian family by Scheffer et al.[5] The French-Canadian families with FPEVF were linked to chromosome 22q11-q12 [6].

We describe a Dutch four-generation family with autosomal dominant partial epilepsy, fulfilling criteria of both ADNFLE and FPEVF. We tested linkage to the known ADNFLE loci at chromosome 1p21 (CHRNB2), 15q24 and

20q13.3 (CHRNA4) and the known FPEVF loci at chromosome 2q36 and 22q11-q12.

Patients and Methods

Patients

This family is included in a nationwide study of the genetics of idiopathic epilepsies in the Netherlands. After obtaining written informed consent, clinical characteristics of the seizures of all participating affected members and, if performed, results of EEG, computed tomography (CT) and magnetic resonance imaging (MRI) were obtained from the treating physician. Furthermore, all participating relatives and married-in spouses (n = 42) were personally interviewed by PC and OB about their medical history and the medical history of their children. Based on the clinical characteristics of the seizures and the results of EEG, CT, and MRI, seizures were classified according to the criteria of the International League Against Epilepsy[29]. This study has been approved by the medical ethical committee of the Leiden University Medical Center.

Genotyping

Venous blood samples were taken from affected relatives and some of the healthy relatives and spouses. Genomic DNA was extracted from peripheral lymphocytes using standard methods[30]. The loci of ADNFLE and FPEVF were tested by polymerase chain reaction (PCR) with the following microsatellite markers: D1S498, D1S305 and D1S2635 for the chromosome 1p21 region; markers D15S211, D15S1041, and D15S979 for the chromosome 15q24 region; markers D20S100, D20S443 and D20S171 for the chromosome 20q13.3 region; markers D2S130, D2S133, and D2S2228 for the chromosome 2q36 region; and markers D22S310, D22S1167, D22S1144, D22S1163, D22S275, D22S1176, D22S273, D22S280, D22S1686, and D22S1162 for the chromosome 22q11-q12 region. Oligonucleotide sequences were available through the Human Genome Database (GDB). PCRs for all markers were

performed using the same protocol. The reaction was performed in a 15-µl reaction volume, containing 7.5 pmol of each primer, 1x superTaq PCR Buffer I (Enzyme Technologies Ltd, UK), 1.3 M betaine (ICN Biomedical Inc, Ohio, USA), 3.00 mM of dNTPs, 0.25 U Silverstar (Eurogentech, Liège, BE) and 45 ng of genomic DNA. This mixture was subjected to ten cycles of 30 seconds at 94°C, 30 seconds at 55°C, and 40 seconds at 72°C, followed by 25 cycles of 30 seconds at 89°C, 30 seconds at 55°C, and 40 seconds at 72°C. The PCR was preceded by an initial denaturation step of ten minutes at 94°C and was ended with an extension step of ten minutes at 72°C. PCR products for each template were pooled and run in an Applied Biosystems (ABI) 377 or 3700 automated DNA sequencer. Allele sizes were determined on the basis of an internal standard size marker (Genescan 400 HD [rox] size standard), using GeneScan 3.5 and Genotyper 3.6 ABI software. Genotypes were determined by two individuals, and checked for Mendelian segregation using UNKNOWN version 5.03.

Linkage analysis

Two- and multipoint LOD score analysis was performed using the Linkage program, version 5.1 [31]. In the linkage analysis, only definitely affected relatives were considered as affected. Diagnosis in all other relatives was considered unknown. Based on the model of Xiong et al., the mode of inheritance was assumed to be autosomal dominant with 50% penetrance[6]. Furthermore, we used a phenocopy rate of 0.01 and a gene frequency of 0.001. Allele frequencies for each individual marker were calculated with ILINK. Multipoint analysis was performed with inter-marker distances according to the database of the Marshfield Center for Medical Genetics for the markers in all five regions (www.marshfieldclinic.org/research/ genetics).

Results

Case history

The proband of this family (IV:9), a 19-year-old male, experienced his first seizure at the age of three months. After breastfeeding, he became apneic and cyanotic for two minutes with generalized hypertonia and staring. The following months, the same occurred several times. The interictal EEG showed right temporal epileptic discharges. The patient received valproic acid and became seizure free.

At the age of three years, he developed short lasting (30-60 seconds) complex partial seizures with tonic-clonic movements of the left arm and leg, accompanied by deviation of the eyes and unconsciousness occurring several times daily, predominantly late in the evening and during the night. Valproic acid was restarted but seizures did not remit. Furthermore, the patient displayed autistic-like behavior with aggressiveness. An initial interictal EEG at that time showed no abnormalities but in a second, long-lasting, EEG intermittently occurring sharp waves were observed in the right centroparietotemporal region without clinical manifestations. Medication was switched to carbamazepine but secondary generalized nocturnal seizures continued. At the end of the seizures he sighed and continued sleeping. In addition, diurnal seizures occurred with staring, unresponsiveness, and motor automatisms for 30 seconds. Medication was changed to phenytoin but without any improvement. The EEG showed right temporal and occipital sharp waves and (poly) spike-wave complexes. Brain MRI was normal.

At the age of six years, the seizure pattern changed into clusters of approximately 40 short seizures (20-30 seconds) in the morning with flushing and distortion of the left corner of his mouth, during which he remained conscious. He used phenobarbital and valproic acid at that time. The ictal EEG showed short series of right frontotemporal 7-8 Hz paroxysmal activity. After increasing the valproic acid dosage, he became seizure free. One year later, the

phenobarbital was stopped and the dose of the valproic acid decreased. The EEG showed no abnormalities at that time.

At the age of 15 years, all medication was stopped. Six months later, seizures re-occurred but at a much lower frequency. At the age of 17 years, seizure frequency increased again up to 2-3 seizures per night. The EEG showed a slow background rhythm with right temporal sharp waves and spikes, after which the dose of the valproic acid was increased. The behavioral problems became worse. A second brain MRI showed no abnormalities.

At the time of last evaluation (19 years), he experienced seizures almost every night during which he often fell out of bed. He also had tonic-clonic seizures during daytime. He was treated with valproic acid and lamotrigine.

Description of the family

In 12 persons of this family epilepsy had been diagnosed; two of them were deceased (figure 1). The family showed autosomal dominant inheritance with incomplete penetrance. There were six obligate carriers (II:3, II:6, III:11, III:17, III:21, and III:27). All patients had normal intelligence and no known history of any condition that could have caused seizures. Computed tomography was performed in five patients and did not show any abnormalities. Magnetic resonance imaging was performed in four patients and showed infratentorial and occipital atrophy in one patient (III:14), which, however, was not associated with clinical symptoms. Three patients had psychiatric problems such as autistic behavior in two (IV:7 and IV:9) and an obsessive-compulsive disorder in one (II:2).

Age at onset of seizures ranged from three months to 24 years (median 7.3 years, table 1). Eight patients had nocturnal seizures with a wide variety of symptoms and severity, and nine patients suffered from diurnal seizures, in one of them occurring shortly after awakening.

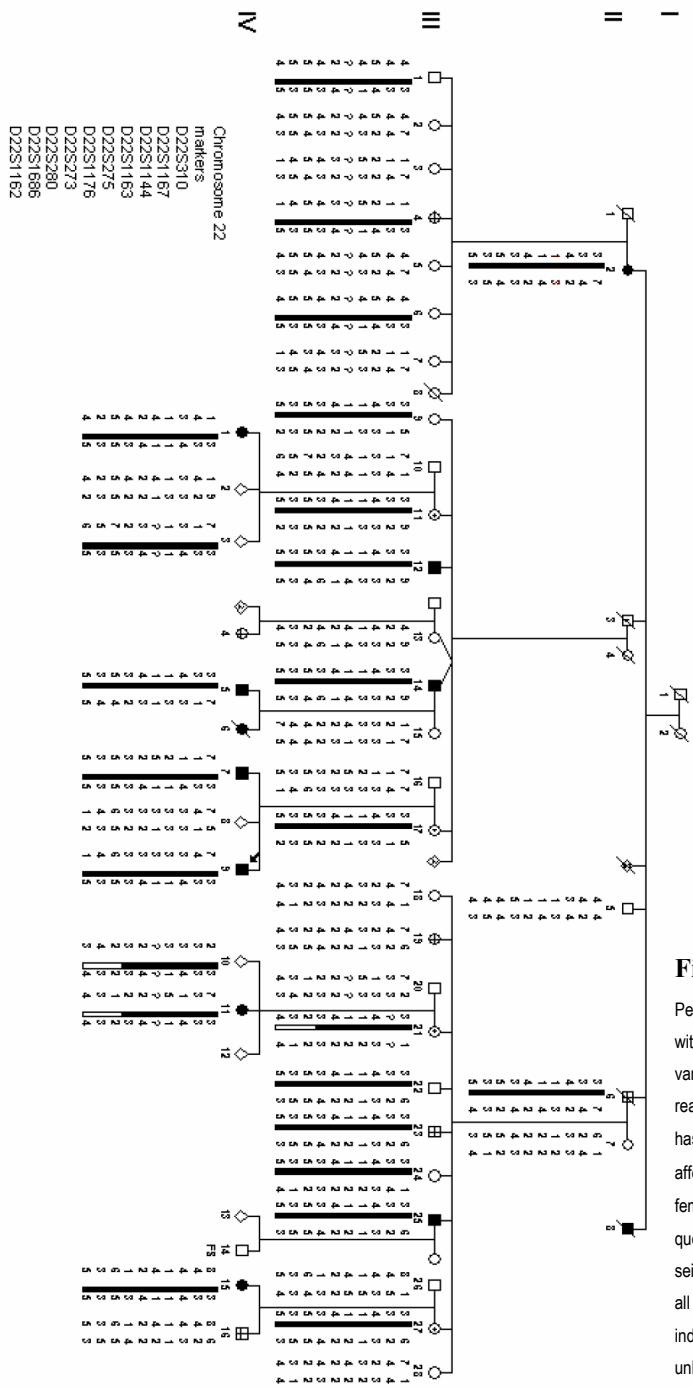


Figure 1

Pedigree of our Dutch family with familial partial epilepsy with variable foci. For privacy reasons, the order of individuals has been changed. ■ = affected male, ● = affected female, • = obligate carrier, + = questionable case, FS = febrile seizures. For linkage analysis, all but the definitely affected individuals were classified as unknown. The proband is indicated by an arrow.

Duration of seizures ranged from 20 seconds to approximately 15 minutes. Seizures were mostly tonic, tonic-clonic or hyperkinetic. They were preceded by autonomic, somatosensory or specific sensory auras in seven patients and accompanied by automatisms in five. None of the patients had auditory symptoms. Seizures occurred in clusters in at least four patients, and could be triggered by stress and sleep deprivation. Intra-individual variation in severity was also observed, with periods of seizures alternating seizure-free periods. Furthermore, seizures were often frequent during childhood and adolescence and tended to decrease in severity and frequency during adulthood although they rarely disappeared completely.

Of one person (III:25), the interictal EEG never showed epileptiform abnormalities; only one of several interictal EEGs of another person (IV:5) showed frontotemporal and frontocentral abnormalities; and half of the interictal EEGs of III:12, III:14, IV:9, IV:11, and IV:15 showed no abnormalities (table 1). In the other interictal EEGs, frontal, central and / or temporal spikes, sharp waves, and spike-wave complexes were observed. An ictal EEG was recorded in two patients, showing frontal and central sharp waves in one (IV:7), and frontotemporal abnormalities in the other (IV:9).

At the time of evaluation, seven patients (aged 12-55 years) still suffered from seizures, of whom six sporadically (< one / month). Nine patients used anti-epileptic drugs, of whom eight were well controlled: one patient had valproic acid monotherapy, one phenobarbital monotherapy, two carbamazepine monotherapy, and five had polytherapy (of whom three had polytherapy with carbamazepine and one with valproic acid).

Table 1

Clinical characteristics of affected members of the Dutch FPEVF family

Patient	Sex	Age (yrs)	Onset (yrs)	Seizure frequency	Current AED	Time of seizure	Seizure classification	EEG findings[a]
II:2	F	88	24	SF 62 yrs	PHB	nocturnal, diurnal	CPS + automatisms, phonatory	frontotemp. slow waves 3-5 Hz
III:12	M	52	24	1-2/year	VPA	nocturnal, diurnal	CPS, SPS speech arrest	50% n.a.; bifrontal epileptic abnormalities (no details)
III:14	M	55	4	1/1-2 mth	CBZ + CLB	nocturnal, diurnal	tonic, CPS + automatisms, SPS	50% n.a.; left frontocentr. waves + spike-wave complexes
III:25	M	47	19	sporadic	CBZ + LTG	shortly after awakening	SPS somatosensory, CPS speech arrest + automatisms	n.a.
IV:1	F	35	3	SF 7 yrs	none	diurnal	CPS, sec. gener. TCS	epileptic abnormalities (no details)
IV:5	M	27	10	sporadic	CBZ + LTG	nocturnal, diurnal	CPS + automatisms	n.a.; 1x frontotemp., frontocentr. parox. slow sharp wave act.
IV:7	M	23	1.8	1/1-3 mth	OXC + LTG	nocturnal, diurnal	SPS posturing of arm(s), CPS TCS	right front., centr., frontotemp., frontocentropar. spikes, spike-wave complexes ictal: right centr., front. sharp waves
IV:9	M	19	0.3	1/night	VPA + LTG	nocturnal, diurnal	CPS +/- automatisms, SPS	50% n.a.; right temp., occ. spikes, waves, polyspike-waves ictal: right frontotemp. 7-8 Hz parox. act.
IV:11	F	12	5	sporadic	CBZ	nocturnal	CPS posturing, SPS somatosensory	50% n.a.; centrotemp. spikes
IV:15	F	22	9.5	SF 22 yrs	CBZ	nocturnal, diurnal	SPS vertiginous, CPS	50% n.a.; prefrontotemp. rare spikes

M = male, F = female, yrs = years, mth = months, SF = seizure free since the age of, AED = anti-epileptic drugs, PHB = phenobarbital, VPA = valproic acid, CBZ = carbamazepine, CLB = clobazam, LTG = lamotrigine, OXC = oxcarbazepine, CPS = complex partial seizures, SPS = simple partial seizures, TCS = tonic-clonic seizures, sec. gener. = secondary generalized, [a] = Except for two EEGs of IV:7 and IV:9, all EEG findings are from interictal EEGs. n.a. = no abnormalities (50% = 50% of the EEGs recorded in this patient showed no abnormalities), temp. = temporal, centr. = central, front. = frontal, par. = parietal, occ. = occipital, parox. act. = paroxysmal activity.

Since no additional EEG studies were performed in these persons, it is unknown whether these periods had an epileptic origin. One person (IV:14) had two febrile seizures at the age of six months, and two others had each a single seizure-like episode of which the epileptic origin could not be confirmed (IV:4, IV:16). For the linkage analysis, the affection status of all these clinically questionable cases was, therefore, regarded as 'unknown'.

Since none of the patients reported auditory or visual symptoms during the seizures, the diagnosis ADLTE was unlikely, despite the fact that in some relatives, including the proband, temporal abnormalities were observed in the interictal EEG. We, therefore, focused our genetic studies on the loci of ADNFLE and FPEVF.

Linkage analysis

Linkage analysis was performed with the three known loci for ADNFLE on chromosome 1p21, 15q24 and 20q13.3, and the two known loci for FPEVF on chromosome 2q36 and 22q11-q12 using several microsatellite markers for each region (D1S498, D1S305 and D1S2635; D15S211, D15S1041, and D15S979; D20S100, D20S443 and D20S171; D2S130, D2S133, and D2S2228; D22S1163 and D22S275). Significantly negative LOD scores (< -2) were found for chromosome 1p21, 2q36, 15q24 and 20q13.3 (data not shown), whereas preliminary evidence for linkage was obtained with the two markers on chromosome 22q11-q12 (multipoint LOD score 2.7). To explore this region further, additional markers were selected from the Marshfield map around D22S1163 and D22S275. The results of haplotype analysis for these markers are shown in figure 1. All affected relatives, ten clinically unaffected relatives and three relatives with vivid dreams and sleepwalking (II:6), nocturnal frightening episodes (III:4), and nightmares (III:23), respectively, carried the disease haplotype (black bar). Of these 13 relatives, five were obligate carriers. Two-point LOD scores between the disease phenotype and each marker are given in table 2.

Table 2

Two-point LOD scores between our family and markers on chromosome 22q11-q12*

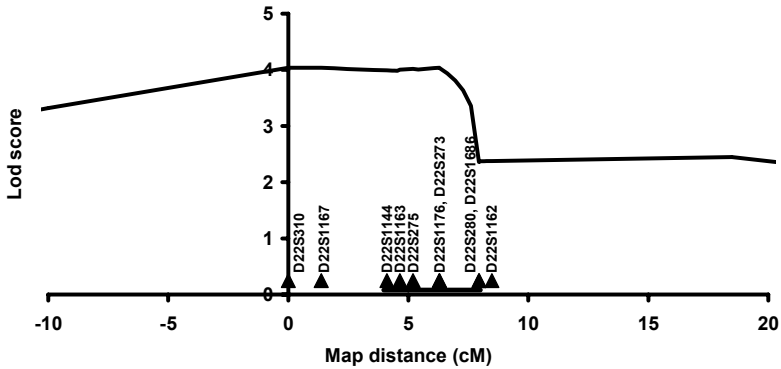
Marker	Location (cM)	Recombination fraction (θ)						Z_{max}	θ_{max}
		0.00	0.01	0.05	0.10	0.15	0.20		
D22S310	18.00	4.037	3.968	3.687	3.319	2.932	2.525	4.037	0.00
D22S1167	19.37	3.822	3.753	3.472	3.104	2.717	2.312	3.822	0.00
D22S1144	22.12	3.276	3.213	2.959	2.630	2.289	1.936	3.276	0.00
D22S1163	22.67	2.262	2.211	2.003	1.738	1.468	1.196	2.262	0.00
D22S275	23.22	1.939	1.897	1.728	1.511	1.291	1.067	1.939	0.00
D22S1176	24.31	3.686	3.617	3.336	2.968	2.582	2.176	3.686	0.00
D22S273	24.31	3.164	3.104	2.860	2.543	2.215	1.874	3.164	0.00
D22S280	25.96	0.916	0.961	1.024	0.981	0.872	0.726	1.024	0.05
D22S1686	25.96	0.952	0.932	0.849	0.745	0.643	0.542	0.952	0.00
D22S1162	26.51	1.898	1.933	1.953	1.855	1.689	1.479	1.960	0.03

Linkage analysis was performed under the assumption of an autosomal dominant mode of inheritance with 50% penetrance, a phenocopy rate of 0.01, and a gene frequency of 0.001. Locations of markers are according to Marshfield. Z_{max} = maximum LOD score for this marker, θ_{max} = recombination fraction at which the maximum LOD score was observed.

Multipoint analysis gave the maximum LOD score of 4.04 at D22S310, D22S1167, and D22S1176 (figure 2). The LOD score dropped to 2.37 at marker D22S280. In person III:21, we observed a haplotype with a recombination between markers D22S273 and D22S280, which was transmitted to her affected daughter (IV:11). On the basis of the LOD scores and the haplotypes, the locus in our family is at least between markers D22S310 and D22S280, a region of 7.93 cM. Since our locus shows complete overlap on the centromeric side with the locus published by Xiong et al. (figure 2, black bar), no additional markers were tested in this region. Therefore, we have no further knowledge of the exact boundary of the disease locus on the centromeric side in our family.

Figure 2

Results of multipoint linkage analysis of our family with chromosome 22q11-q12 markers illustrating the region of linkage. Genetic distances from D22S310 are given in centimorgans. The black bar indicates the region of linkage as described by Xiong et al.[6]



Discussion

We describe a Dutch four-generation family with autosomal dominantly inherited epilepsy with apparent incomplete penetrance. The clinical characteristics of the epilepsy fulfill criteria of both nocturnal frontal lobe epilepsy (ADNFLE) and familial partial epilepsy with variable foci (FPEVF)[2,5,6,10,12,14,32-36]. These syndromes are phenotypically overlapping and, therefore, possibly difficult to differentiate. The most important difference is that patients with FPEVF suffer more frequently from diurnal seizures than patients with ADNFLE, and that EEGs from patients with FPEVF show variable abnormalities, whereas EEGs from patients with ADNFLE predominantly show abnormalities originating from the anterior quadrants[2,5]. Diurnal seizures were, however, also described in families with ADNFLE[36], and EEG abnormalities might also originate from other regions in families with ADNFLE[2,32,35-37]. The frequent occurrence of seizures during daytime, with one of the affected family members (IV:1) suffering from diurnal seizures only, and the observation of interictal EEG abnormalities originating from

different cortical areas are more in agreement with the diagnosis FPEVF in our family. The question why patients with FPEVF have a much more heterogeneous phenotype than patients with ADNFLE has yet to be determined.

Three families with FPEVF have been described until now, one Australian family and two French-Canadian families[5,6]. The French-Canadian families shared an identical linked haplotype and can, therefore, be regarded as one large extended family[6]. Clinical features of the described families were virtually similar to our family: the patients had partial seizures originating from different cortical areas and with variable age at onset. The epileptic focus was frontal or temporal in most patients. Most patients in our family had nocturnal seizures but diurnal seizures were also observed. Since seizures were predominantly nocturnal in the French-Canadian family and mostly diurnal in the Australian family, the clinical characteristics of our family resemble those of the French-Canadian family more closely. The Australian family had suggestive linkage to chromosome 2q36 [5], while the French-Canadian family was linked to chromosome 22q11-q12 [6].

Three affected relatives in our family had psychiatric problems, whereas none of the non-epileptic relatives did. In the Australian family, behavioral problems were described in one affected relative. In the French-Canadian family, four persons with paranoid schizophrenia were identified, but they did not have epilepsy. It was not stated whether these persons had the disease haplotype. Because psychiatric problems were reported in only four out of almost 60 patients in the three families with FPEVF, it seems unlikely that psychiatric disorders are a part of the FPEVF phenotype.

In our family, linkage analysis was performed with the three known ADNFLE loci on chromosome 1p21, 15q24 and 20p13.3, and the two known FPEVF loci on chromosome 2q36 and 22q11-q12. Linkage was observed with

chromosome 22q11-q12, supporting the diagnosis FPEVF. Besides the affected relatives, 13 relatives carried the disease haplotype, including five obligate carriers and three of the four persons with abnormal phenomena during sleep (one of whom, II:6, was obligate carrier). The question, therefore, arises whether these phenomena might be of epileptic origin. Three of the other clinically unaffected relatives that carried the haplotype (III:22, IV:3, IV:10) were younger than 45 years and may still be at risk of developing epilepsy at a later age.

Our family is the first family that confirms linkage of FPEVF to chromosome 22q11-q12, previously reported by Xiong et al.[6] We observed a recombination between markers D22S273 and D22S280 in person III:21 that was transmitted to her affected daughter (IV:11), indicating the telomeric boundary of the locus in our family. This end is at the same marker as observed by Xiong et al. The boundary of the disease locus on the centromeric side in our family is unknown but our region of linkage on this side is larger than that in the published French-Canadian family[6]. The linked region, therefore, overlaps their region of linkage; we were unable to reduce the area.

The candidate region of at least 7.93 cM defined by haplotype reconstruction and linkage analysis has a high density of known and putative genes (more than 100) and is physically large (approximately 6.5 Mb) (Genemap; www.ncbi.nlm.nih.gov/genemap99). Candidate genes in this area are the seizure related gene 6 homolog-like (SEZ6L) and the genes encoding synapsin III and tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein (YWHAH). In humans, the SEZ6L gene on chromosome 22q11-q12 is quite similar to the seizure related 6 homolog gene (SEZ6) on chromosome 17q11.2, which encodes a brain-specific membrane protein. In mice, *Sez6* is located on chromosome 11B5 [38]. It was identified with linkage analysis in mice, showing tonic-clonic seizures after pentylenetetrazol injection[38]. Pentylenetetrazol acts as a convulsant via the $GABA_A$ benzodiazepine-receptor

complex. By determining the minimal dose to induce convulsions in mice, a good estimate of the general excitability of the central nervous system can be obtained. Synapsin III is a neuron-specific synaptic vesicle-associated phosphoprotein, involved in the regulation of neurotransmitter release and synaptogenesis[39]. Knockout mice for synapsin I and/or II experience seizures with a frequency proportional to the number of mutant alleles[40]. D22S280 is located within an intron between exons 6 and 7 of the synapsin III gene. YWHAH encodes a protein controlling intracellular signaling and neurotransmitter release[41]. The protein is located exclusively in the cytoplasm of neurons in the cerebral cortex. This protein may be associated with neuropsychiatric disorders[42]. Whether one of these genes is involved in the epilepsy of this family will have to be determined by sequence analysis.

Acknowledgements

We would like to thank the family for their co-operation. Furthermore, we would like to thank M. Boel, J.M. Diemel, G.J. de Haan, D.J. Kamphuis, M.J. Jongsma, G.J. van der Linden, F.G.A. van der Meché, M. Stein, E. Thiery, E.P.L.A. Timmermans, W. Verslegers, M. Vervaeck, H.J. Vroon, and P.J.M. van Wensen for providing us clinical information about affected relatives. We thank M. Hoekstra and E.E. Kors for assistance with sample processing and genotyping. P.M.C. Callenbach is a research fellow supported by a grant from the Netherlands Organization for Health, Research and Development (ZonMw, 940-33-030) and the Dutch National Epilepsy Fund-'The power of the small' (98-14).

References

1. Scheffer IE, Bhatia KP, Lopes-Cendes I et al. Autosomal dominant frontal epilepsy misdiagnosed as sleep disorder. Lancet 1994;343:515-7.
2. Scheffer IE, Bhatia KP, Lopes-Cendes I et al. Autosomal dominant nocturnal frontal lobe epilepsy. A distinctive clinical disorder. Brain 1995;118:61-73.
3. Ottman R, Risch N, Hauser WA et al. Localization of a gene for partial epilepsy to chromosome 10q. Nat Genet 1995;10:56-60.
4. Poza JJ, Saenz A, Martinez-Gil A et al. Autosomal dominant lateral temporal epilepsy:

clinical and genetic study of a large Basque pedigree linked to chromosome 10q. Ann Neurol 1999;45:182-8.

5. Scheffer IE, Phillips HA, O'Brien CE et al. Familial partial epilepsy with variable foci: a new partial epilepsy syndrome with suggestion of linkage to chromosome 2. Ann Neurol 1998;44:890-9.
6. Xiong L, Labuda M, Li DS et al. Mapping of a gene determining familial partial epilepsy with variable foci to chromosome 22q11-q12. Am J Hum Genet 1999;65:1698-710.
7. Phillips HA, Scheffer IE, Berkovic SF, Hollway GE, Sutherland GR, Mulley JC. Localization of a gene for autosomal dominant nocturnal frontal lobe epilepsy to chromosome 20q13.2. Nat Genet 1995;10:117-8.
8. Steinlein OK, Mulley JC, Propping P et al. A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. Nat Genet 1995;11:201-3.
9. Steinlein OK, Magnusson A, Stoodt J et al. An insertion mutation of the CHRNA4 gene in a family with autosomal dominant nocturnal frontal lobe epilepsy. Hum Mol Genet 1997;6:943-47.
10. Saenz A, Galan J, Caloustian C et al. Autosomal dominant nocturnal frontal lobe epilepsy in a Spanish family with a Ser252Phe mutation in the CHRNA4 gene. Arch Neurol 1999;56:1004-9.
11. Hirose S, Iwata H, Akiyoshi H et al. A novel mutation of CHRNA4 responsible for autosomal dominant nocturnal frontal lobe epilepsy. Neurology 1999;53:1749-53.
12. Steinlein OK, Stoodt J, Mulley J, Berkovic S, Scheffer IE, Brodtkorb E. Independent occurrence of the CHRNA4 Ser248Phe mutation in a Norwegian family with nocturnal frontal lobe epilepsy. Epilepsia 2000;41:529-35.
13. Phillips HA, Marini C, Scheffer IE, Sutherland GR, Mulley JC, Berkovic SF. A de novo mutation in sporadic nocturnal frontal lobe epilepsy. Ann Neurol 2000;48:264-7.
14. Gambardella A, Annesi G, De Fusco M et al. A new locus for autosomal dominant nocturnal frontal lobe epilepsy maps to chromosome 1. Neurology 2000;55:1467-71.
15. De Fusco M, Becchetti A, Patrignani A et al. The nicotinic receptor beta2 subunit is mutant in nocturnal frontal lobe epilepsy. Nat Genet 2000;26:275-6.
16. Phillips HA, Favre I, Kirkpatrick M et al. CHRNB2 Is the second acetylcholine receptor subunit associated with autosomal dominant nocturnal frontal lobe epilepsy. Am J Hum Genet 2001;68:225-31.
17. Phillips HA, Scheffer IE, Crossland KM et al. Autosomal dominant nocturnal frontal-lobe epilepsy: genetic heterogeneity and evidence for a second locus at 15q24. Am J Hum Genet 1998;63:1108-16.
18. Winawer MR, Ottman R, Hauser WA, Pedley TA. Autosomal dominant partial epilepsy with auditory features: defining the phenotype. Neurology 2000;54:2173-6.
19. Michelucci R, Passarelli D, Pitzalis S, Dal Corso G, Tassinari CA, Nobile C. Autosomal dominant partial epilepsy with auditory features: description of a new family. Epilepsia 2000;41:967-70.
20. Winawer MR, Boneschi FM, Barker-Cummings C et al. Four new families with autosomal dominant partial epilepsy with auditory features: clinical description and linkage to chromosome 10q24. Epilepsia 2002;43:60-67.
21. Brodtkorb E, Gu W, Nakken KO, Fischer C, Steinlein OK. Familial temporal lobe epilepsy with aphasic seizures and linkage to chromosome 10q22-q24. Epilepsia 2002;43:228-35.
22. Kalachikov S, Evgrafov O, Ross B et al. Mutations in LGI1 cause autosomal-dominant partial epilepsy with auditory features. Nat Genet 2002;30:335-41.

23. Morante-Redolat JM, Gorostidi-Pagola A, Piquer-Sirerol S et al. Mutations in the LGI1/Epitempin gene on 10q24 cause autosomal dominant lateral temporal epilepsy. Hum Mol Genet 2002;11:1119-28.
24. Gu W, Brodtkorb E, Steinlein OK. LGI1 is mutated in familial temporal lobe epilepsy characterized by aphasic seizures. Ann Neurol 2002;52:364-7.
25. Wallace RH, Izzillo P, MacIntosh AM, Mulley JC, Berkovic SF. Mutations in LGI1 in an Australian family with familial temporal lobe epilepsy with auditory features. Am J Hum Genet 2002;71 (suppl. 4):472.
26. Pizzuti A, Flex E, Di Bonaventura C et al. Epilepsy with auditory features: a LGI1 gene mutation suggests a loss- of-function mechanism. Ann Neurol 2003;53:396-9.
27. Fertig E, Lincoln A, Martinuzzi A, Mattson RH, Hisama FM. Novel LGI1 mutation in a family with autosomal dominant partial epilepsy with auditory features. Neurology 2003;60:1687-90.
28. Chernova OB, Somerville RP, Cowell JK. A novel gene, LGI1, from 10q24 is rearranged and downregulated in malignant brain tumors. Oncogene 1998;17:2873-81.
29. Commission on Classification and Terminology of the International League Against Epilepsy. Proposal for revised clinical and electroencephalographic classification of epileptic seizures. Epilepsia 1981;22:489-501.
30. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16:1215.
31. Lathrop GM, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci U S A 1984;81:3443-6.
32. Hayman M, Scheffer IE, Chinvarun Y, Berlangieri SU, Berkovic SF. Autosomal dominant nocturnal frontal lobe epilepsy: demonstration of focal frontal onset and intrafamilial variation. Neurology 1997;49:969-75.
33. Nakken KO, Magnusson A, Steinlein OK. Autosomal dominant nocturnal frontal lobe epilepsy: an electroclinical study of a Norwegian family with ten affected members. Epilepsia 1999;40:88-92.
34. Ito M, Kobayashi K, Fujii T et al. Electroclinical picture of autosomal dominant nocturnal frontal lobe epilepsy in a Japanese family. Epilepsia 2000;41:52-8.
35. Nakken KO and Brodtkorb E. Two Norwegian families illustrating the problem of genotype and phenotype in frontal lobe epilepsy. Acta Neurol Scand Suppl 2000;102 (suppl. 175):25-7.
36. Oldani A, Zucconi M, Ferini-Strambi L, Bizzozero D, Smirne S. Autosomal dominant nocturnal frontal lobe epilepsy: electroclinical picture. Epilepsia 1996;37:964-76.
37. Provini F, Plazzi G, Tinuper P, Vandi S, Lugaresi E, Montagna P. Nocturnal frontal lobe epilepsy. A clinical and polygraphic overview of 100 consecutive cases. Brain 1999;122:1017-31.
38. Wakana S, Sugaya E, Naramoto F et al. Gene mapping of SEZ group genes and determination of pentylenetetrazol susceptible quantitative trait loci in the mouse chromosome. Brain Res 2000;857:286-90.
39. Kao HT, Porton B, Czernik AJ et al. A third member of the synapsin gene family. Proc Natl Acad Sci U S A 1998;95:4667-72.
40. Rosahl TW, Spillane D, Missler M et al. Essential functions of synapsins I and II in synaptic vesicle regulation. Nature 1995;375:488-93.
41. Muratake T, Hayashi S, Ichikawa T et al. Structural organization and chromosomal assignment of the human 14-3-3 eta chain gene (YWHAH). Genomics 1996;36:63-9.

42. Toyooka K, Muratake T, Tanaka T et al. 14-3-3 protein eta chain gene (YWHAH) polymorphism and its genetic association with schizophrenia. Am J Med Genet 1999;88:164-7.

CHAPTER 9

General discussion

JJ Hottenga

In this section the various methods that have been applied in this thesis for analyzing the correlation between genotype and phenotype in complex neurological disorders will be discussed together with some possible future perspectives.

Association & stratification

For complex neurological disorders association studies are a straightforward method to quickly assess the involvement of candidate genes. Furthermore, the design is powerful for testing polymorphisms with small effects, which is essential for studying the genetics of complex disorders[1]. However, association studies are often criticized because of the failure of replicating positive results[2]. In literature, the advantages, as well as the inherent problems of the study design have been thoroughly discussed[2-4]. One of the major concerns for association studies has been population stratification[1,5]. Spurious association due to population stratification occurs when both the disorder and gene frequency differs between two populations. If cases and controls are selected in different proportions from these two populations, association will occur without a causal relation between the tested polymorphism and the disorder.

In *chapter two* it was shown that population stratification surprisingly has a small effect, except for extreme situations; the unlikely situation that the gene frequency and/or the selection of two populations is extremely different in cases and controls. Although this may lead to the conclusion that one does not need to take population stratification into account, there are reasons why such stratification should be avoided. Population stratification may be relevant in case of studying very large samples of cases and controls, or searching for gene variants with limited effect size[6-8]. Only a small number of empirical studies is currently available, concerning the presence and use of testing population stratification. In addition, studies report contradictory results based on what is considered to be a substantial increase in false-positive findings[9-11]. To use an analogy, if one would perform a case-control study involving lung-

cancer and a given risk factor, confounding factors like smoking, gender and age should be taken into account when designing such a study. The confounding of population stratification can taken into account as well, as several tests and correction methods have been proposed[12-17].

If population stratification is not a major issue, the question remains why association studies are often false-positives and cannot easily be replicated. There is good reason to argue that major factors are the sample size of the study population, and statistical problems due to multiple testing. As shown indirectly in *chapters three* and *four*, sample size can largely influence the outcome of studies. Unfortunately, many studies suffer from too small sample sizes to evaluate genuine associations (and gene-gene interactions)[18]. Therefore, increasing the sample size (to several hundreds or even thousands), and increasing the statistical significance level (well below $\alpha = 0.05$) seem logical remedies to reduce false-positive association results. Replication studies, meta-analyses and more in-depth functional research should be applied to confirm initial association findings. Recently, it has been proposed that data for association studies should be made available online[19]. In this way other researchers can add data and check their results in combined data sets. Such an approach would make more effective use of resources and would help in avoiding publication bias towards false-positive findings.

The use of parametric linkage analysis in complex neurological disorders

Although parametric linkage analysis is often considered to be less efficient than non-parametric methods for localizing genes in complex traits, the results of this thesis show that it can be a useful approach given that efforts to study homogenous material are taken. In *chapters seven* and *eight* parametric linkage analysis was applied to study single Mendelian families affected with epilepsy. In *chapter seven* a Dutch family with familial cortical tremor and

epilepsy (FCTE) showed no evidence for linkage to the Japanese 8q23.3-24.1 locus, which indicated heterogeneity for this phenotype[20,21]. In *chapter eight* linkage analysis was performed in a single family for several loci involved in autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE) (1p21, 15q24, 20q.13.3) and familial partial epilepsy with variable foci (FPEVF) (2q36, 22q11-q12)[22-26]. Linkage to chromosome 22q11-q12 favored the diagnosis of FPEVF, showing that linkage analysis can be used to support a diagnosis of rare familial epilepsy syndromes.

In addition to the epilepsy families, parametric linkage analysis was applied in seven Dutch migraine without aura (MO) families (*chapter six*). In order to increase the homogeneity, the MO families were selected based on the criterion that (nearly) all affected individuals should have MO. Branches in which a spouse was affected with migraine were not used. Despite these efforts, none of the individual families showed significant - or suggestive evidence for linkage. LOD scores were substantially lower than the expected simulated LOD scores, assuming a single segregating disorder gene per family. Allelic heterogeneity and/or presence of phenocopies were found, as a number of affected individuals in the large families did not carry the specific haplotype segregating with the disorder. Locus heterogeneity seemed to be present as well but could not be confirmed. Given the complexity and prevalence of migraine, the heterogeneity is not surprising. Interestingly, suggestive evidence for linkage was found at a locus on chromosome 4q21-q24, replicating the results of the Iceland and Finnish genome scans[27,28].

Family selection and effects of heterogeneity in the MO linkage analysis

In *chapter six* several strategies were employed to account for the heterogeneity that is obviously a large problem in the genetics of migraine. The selection of specific phenotypes to reduce heterogeneity has often been a

successful approach for mapping genes, for example in early-onset cases of Alzheimer's disease and familial hemiplegic migraine[29,30]. However, the method has an effect only if the selection criterion contributes to a more homogenous sample. The possibility that multiple risk factors segregate in the studied families remains, which probably occurred in *chapter six* when selecting the families with MO. The selection of these highly loaded families may even have contributed to the heterogeneity because it becomes more probable to select families, in which multiple disorder genes segregate. Some evidence for this was found in the segregation analyses of *chapter five*, where the polygenic model fitted as good as the general dominant single locus model.

In contrast to the epilepsy families, the migraine families did not show strong evidence for linkage when analyzed individually. Two families were large enough to detect significant linkage. Ironically, the family size also increases the probability of heterogeneity, as the married-in spouses may contribute new risk factors. This was avoided as much as possible by excluding branches with two affected parents but probably combinations of risk factors still contributed to the migraine development in the remaining branches. An alternative approach is to analyze a number of (smaller) families in a single analysis but it will likely increase the probability of locus heterogeneity. In *chapter six* this approach was employed as well without a substantial change in conclusions. Analyzing a single large family or multiple smaller families remains a dilemma. In the field of migraine the use of both methods has led to positive results recently[27,31]. Given a substantial publication bias, it is difficult to determine, which method is optimal. Theoretically, there is only a limited number of genes involved in a disorder, as a limited number of biochemical pathways are affected. Therefore, studying a very large sample of smaller families should eventually have more power to detect the responsible gene variant(s), as compared to a single family approach.

Another frequently used approach to account for heterogeneity is a test with an iterated mixture parameter added to the statistic, which determines the probability of a given family to be linked to the locus[32-34]. Also this approach was applied in *chapter six,* using the program HOMOG[33,34]. Although this method is increasing the power of locus detection under heterogeneity, it is far from perfect. For example, the linkage model parameters are assumed to be equal for the loci. Also, the estimations of the mixture parameter can be wrong, depending on the parametric linkage model used and the number of families that is tested. Furthermore, the likelihood statistic has a difficult distribution with one or two degrees of freedom[35,36]. The mixture method has been extended for the non-parametric analysis, dealing with some of the problems that are inherent to the use of a linkage model[37]. Whether parametric or non-parametric linkage analysis is the best approach to detect linkage is an issue of discussion[38,39].

Alternative approaches accounting for heterogeneity

To test linkage in migraine without aura, alternative strategies could have been employed to reduce heterogeneity as well. These include study design changes like testing association or sib-pair analysis instead of (parametric) linkage[1]. Another way to cope with heterogeneity using a linkage approach is to divide larger families into smaller nuclear families and analyze them as being independent, using sib-pair analysis or non-parametric methods[39]. When the mode of inheritance is specified as dominant for parametric linkage, while the true mode of inheritance is recessive, this method will increase the detection probability of recessive loci[27,40,41]. In case the mode of inheritance was specified correctly, some power is probably lost because pedigrees have been split into nuclear families[42]. Some genome scans analyzing the data with both methods show that the LOD scores and detected locations are often very similar[27,43,44].

In addition to increasing the homogeneity of the phenotype, the homogeneity of the whole genome in a studied sample can be increased as well. Families can be selected from an isolated population, in which it is assumed that genetic drift, disease bottlenecks and founder effects have reduced the heterogeneity of the genetic risk factors[45]. Preferably, the genealogy of the population is known as well, so that selected families or persons can be related to each other[45-47]. Currently, a number of isolate studies applying different design- and statistical approaches have been published with optimistic results[47-49]. It should be noted, however, that the heterogeneity may not be reduced for some disorders of interest. In addition, the found loci may be unimportant risk factors in other populations.

Nowadays, research should be aimed at developing more specific methods to detect linkage under heterogeneity. Correcting for linkage evidence at other loci may be such an option and various methods to employ this strategy have been developed[50-52]. The use of ordered subset analysis, in which families are rank-ordered based on a covariate (phenotype) and then permuted until the maximum LOD score of a given subset is found, may be extended for heterogeneity as well[53].

The selection of a proper family-based association test

With the current availability of dense single nucleotide polymorphisms maps and the increased number of mapped susceptibility loci, the emphasis of future genetic research for complex diseases will more often be focused on fine mapping genes with family-based association studies. A simple question, though sometimes difficult to answer, is how to select the proper statistical method used to analyze the data. Of course, a limitation is the study sample characteristics but the possibilities are large when a sample of sib-pairs for a dichotomous trait with unaffected, affected siblings and parents has been genotyped. If parent-affected child trios are selected from these data, then the haplotype relative risk (HRR) method, transmission disequilibrium test (TDT),

reconstruction combined-transmission disequilibrium test (RC-TDT) or TRANSMIT test can be applied[54-57]. In addition, tests like the sib transmission disequilibrium test (S-TDT), the discordant alleles test (DAT) and discordant sibship test (SDT) that make use of the known allele sharing / transmission in siblings can be used as well[58-60]. Furthermore, complete families can be analyzed, using a weighting function for the familial relations in tests like the pedigree disequilibrium test (PDT) and family based association test (FBAT)[61,62]. These examples were taken from a much larger list of family-based association tests that was thoroughly reviewed in Schulze and McMahon in 2002[63]. Since then, the number of tests has still increased[64,65]. Finally, other possibilities can be taken under consideration, such as the use of covariates in the analysis, the use of multiple markers or (tag) haplotypes and the use of quantitative traits[66-68].

In *chapter four* three statistics (TDT, Mantel-Haenszel extension, Z'score TDT/S-TDT) were applied to study if there was an association between the HVR haplotype, causing retinopathy, and co-morbid migraine and Raynaud phenomenon in a Dutch family[55,58.,69]. The data were generated using only one large pedigree, which caused problems in several of the proposed tests (PDT and FBAT) because the statistics are based on the weighting of multiple families[61,62]. However, splitting the family into multiple nuclear families resolved this issue, and some family-based statistics could be employed. Differences that were observed between the use of trios and sib-pairs may be explained by the increase of sample size using sibling-based approaches. With the different approaches, changing sub-samples from a single family are studied, therefore, the results may differ in outcome based on the selection. If the relation between the genotype and phenotype is strong and sample size is relatively large, these effects will probably not alter the outcome. In smaller samples, however, this might not be true. It is therefore important to know the properties of a given test. Here, literature becomes less extensive: many tests are developed but are tested only for a limited number of situations. In

addition, a limited number of tests is compared and mainly for the power to detect association. Experience in implementation and support for many of these tests is difficult to obtain, which may lead to a reduced use of the most optimal method. Future research should be aimed at more careful comparison of tests with empirical and simulated data. In addition, a more user-friendly program combining multiple tests, like SPSS for example, would likely be helpful for many (epidemiological) geneticists.

Conclusion

Different complex neurological disorders require different mapping strategies and study design to successfully locate genes involved in the disorder. Linkage analysis and association analysis are applied to contribute to these findings. In this thesis the results were often dependent on a selection either in samples or statistics used for the analysis. In case-control studies the selection of cases and controls may sometimes lead to confounding. In linkage analysis the family selection and method of analysis can be the difference between failure and success of a study. Essential is to know the effects and limitations within a study, even more when other possibilities of research are limited, which is often the case in complex neurological disorders.

References

1. Lander ES, Schork NJ. Genetic dissection of complex traits. Science. 1994;265:2037-48.
2. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. Lancet. 2003;8;361:865-72.
3. Little J, Bradley L, Bray MS et al. Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. Am J Epidemiol. 2002;15;156:300-10.
4. Ioannidis JP. Genetic associations: false or true? Trends Mol Med. 2003;9:135-8.
5. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet. 2003;15;361:598-604.
6. Marchini J, Cardon LR, Phillips MS, Donnelly P. The effects of human population structure on large genetic association studies. Nat Genet. 2004;36:512-7.
7. Freedman ML, Reich D, Penney KL et al. Assessing the impact of population stratification on genetic association studies. Nat Genet. 2004;36:388-93.
8. Wang Y, Localio R, Rebbeck TR. Evaluating bias due to population stratification in case-control association studies of admixed populations. Genet Epidemiol. 2004;27:14-20.

9. Koller DL, Peacock M, Lai D et al. False positive rates in association studies as a function of degree of stratification. J Bone Miner Res. 2004;19:1291-5.
10. Thomas DC, Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene associations? Cancer Epidemiol Biomarkers Prev. 2002;11:505-12.
11. Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. Cancer Epidemiol Biomarkers Prev. 2002;11:513-20.
12. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999;55:997-1004.
13. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet. 1999;65:220-8.
14. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945-59.
15. Satten GA, Flanders WD, Yang Q. Accounting for unmeasured population substructure in case-control studies of genetic association using a novel latent-class model. Am J Hum Genet. 2001;68:466-77.
16. Ripatti S, Pitkaniemi J, Sillanpaa MJ. Joint modeling of genetic association and population stratification using latent-class models. Genet Epidemiol. 2001; 21 Suppl 1:S409-14.
17. Pfaff CL. Adjusting for population structure in admixed populations. Genetic Epidemiology. 2002;1. 22:196–201.
18. Ioannidis JP, Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG. Genetic associations in large versus small studies: an empirical assessment. Lancet. 2003;15;361:567-71.
19. Ioannidis JP, Rosenberg PS, Goedert JJ, O'Brien TR; International Meta-analysis of HIV Host Genetics. Commentary: meta-analysis of individual participants' data in genetic epidemiology. Am J Epidemiol. 2002;156:204-10.
20. Plaster NM, Uyama E, Uchino M et al. Genetic localization of the familial adult myoclonic epilepsy (FAME) gene to chromosome 8q24. Neurology. 1999;53:1180-3.
21. Mikami M, Yasuda T, Terao A et al. Localization of a gene for benign adult familial myoclonic epilepsy to chromosome 8q23.3-q24.1. Am J Hum Genet. 1999;65:745-51.
22. De Fusco M, Becchetti A, Patrignani A et al. The nicotinic receptor beta 2 subunit is mutant in nocturnal frontal lobe epilepsy. Nat Genet. 2000 Nov;26:275-6.
23. Scheffer IE, Phillips HA, O'Brien CE et al. Familial partial epilepsy with variable foci: a new partial epilepsy syndrome with suggestion of linkage to chromosome 2. Ann Neurol. 1998;44:890-9.
24. Phillips HA, Scheffer IE, Crossland KM et al. Autosomal dominant nocturnal frontal-lobe epilepsy: genetic heterogeneity and evidence for a second locus at 15q24. Am J Hum Genet. 1998;63:1108-16.
25. Phillips HA, Scheffer IE, Berkovic SF et al. Localization of a gene for autosomal dominant nocturnal frontal lobe epilepsy to chromosome 20q13.2. Nat Genet. 1995;10:117-8.
26. Xiong L, Labuda M, Li DS et al. Mapping of a gene determining familial partial epilepsy with variable foci to chromosome 22q11-q12. Am J Hum Genet. 1999;65:1698-710.
27. Wessman M, Kallela M, Kaunisto MA et al. A susceptibility locus for migraine with aura, on chromosome 4q24. Am J Hum Genet. 2002;70:652-62.
28. Bjornsson A, Gudmundsson G, Gudfinnsson E et al. Localization of a gene for migraine without aura to chromosome 4q21. Am J Hum Genet. 2003;73:986-93.

29. Van Broeckhoven C, Backhovens H, Cruts M et al. Mapping of a gene predisposing to early-onset Alzheimer's disease to chromosome 14q24.3. Nat Genet. 1992;2:335-9.
30. Joutel A, Bousser MG, Biousse V et al. A gene for familial hemiplegic migraine maps to chromosome 19. Nat Genet. 1993;5:40-5.
31. Soragna D, Vettori A, Carraro G et al. A locus for migraine without aura maps on chromosome 14q21.2-q22.3. Am J Hum Genet. 2003;72:161-7.
32. Smith CAB. Testing for heterogeneity of recombination fraction values in human genetics. Ann Hum Genet. 1963;27:175–182.
33. Ott J. Linkage analysis and family classification under heterogeneity. Ann Hum Genet. 1983;47:311–20.
34. Bhat A, Heath SC, Ott J. Heterogeneity for multiple disease loci in linkage analysis. Hum Hered. 1999;49:229-31.
35. Whittemore AS, Halpern J. Problems in the definition, interpretation, and evaluation of genetic heterogeneity. Am J Hum Genet. 2001;68:457-65.
36. Hodge SE, Vieland VJ, Greenberg DA. HLODs remain powerful tools for detection of linkage in the presence of genetic heterogeneity. Am J Hum Genet. 2002;70:556-9.
37. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet. 1996;58:1347–63.
38. Abreu PC, Greenberg DA, Hodge SE. Direct power comparisons between simple LOD scores and NPL scores for linkage analysis in complex diseases. Am J Hum Genet. 1999;65:847-57.
39. Goring HH, Terwilliger JD. Linkage analysis in the presence of errors IV: joint pseudomarker analysis of linkage and/or linkage disequilibrium on a mixture of pedigrees and singletons when the mode of inheritance cannot be accurately specified. Am J Hum Genet. 2000;66:1310-27.
40. Knapp M, Seuchter SA, Baur MP. Linkage analysis in nuclear families. 1: Optimality criteria for affected sib-pair tests. Hum Hered. 1994;44:37-43.
41. Knapp M, Seuchter SA, Baur MP. Linkage analysis in nuclear families. 2: Relationship between affected sib-pair tests and lod score analysis. Hum Hered. 1994;44:44-51.
42. Huang J, Vieland VJ. Comparison of 'model-free' and 'model-based' linkage statistics in the presence of locus heterogeneity: single data set and multiple data set applications. Hum Hered. 2001;51:217-25.
43. Das SK, Hasstedt SJ, Zhang Z, Elbein SC. Linkage and association mapping of a chromosome 1q21-q24 type 2 diabetes susceptibility locus in northern European Caucasians. Diabetes. 2004;53:492-9.
44. Angius A, Petretto E, Maestrale GB et al. A new essential hypertension susceptibility locus on chromosome 2p24-p25, detected by genomewide search. Am J Hum Genet. 2002;71:893-905.
45. Peltonen L, Palotie A, Lange K. Use of population isolates for mapping complex traits. Nat Rev Genet. 2000;1:182-90.
46. Enserink M. Physicians wary of scheme to pool Icelanders' genetic data. Science. 1998;281:890-891.
47. Vaessen N, Heutink P, Houwing-Duistermaat JJ, et al. A genome-wide search for linkage-disequilibrium with type 1 diabetes in a recent genetically isolated population from the Netherlands. Diabetes. 2002;51:856-9.
48. Freimer NB, Reus VI, Escamilla MA et al. Genetic mapping using haplotype, association and linkage methods suggests a locus for severe bipolar disorder (BPI) at 18q22-q23. Nat Genet. 1996;12:436-41.

49. Gretarsdottir S, Sveinbjornsdottir S, Jonsson HH et al. Localization of a susceptibility gene for common forms of stroke to 5q12. Am J Hum Genet. 2002;70:593-603.
50. Schaid DJ, McDonnell SK, Thibodeau SN. Regression models for linkage heterogeneity applied to familial prostrate cancer. Am J Hum Genet. 2001;68:1189–96.
51. Cordell HJ, Jacobs KB, Wedig GC, Elston RC. Improving the power for disease locus detection in affected-sib-pair studies by using two-locus analysis and multiple regression methods. Genet Epidemiol. 1999;17 Suppl 1:S521-6.
52. Vieland VJ, Wang K, Huang J. Power to detect linkage based on multiple sets of data in the presence of locus heterogeneity: comparative evaluation of model-based linkage methods for affected sib pair data. Hum Hered. 2001;51:199–208.
53. Schmidt S, Scott WK, Postel EA et al. Ordered subset linkage analysis supports a susceptibility locus for age-related macular degeneration on chromosome 16p12. BMC Genet. 2004;5:18.
54. Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. Ann Hum Genet. 1987;51:227-33.
55. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet. 1993;52:506-16.
56. Knapp M. Reconstructing parental genotypes when testing for linkage in the presence of association. Theor Popul Biol. 2001;60:141-8.
57. Clayton D. A generalization of the transmission/disequilibrium test for uncertain-haplotype transmission. Am J Hum Genet. 1999;65:1170-7.
58. Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. Am J Hum Genet. 1998;62:450-8.
59. Boehnke M, Langefeld CD. Genetic association mapping based on discordant sib pairs: the discordant-alleles test. Am J Hum Genet. 1998;62:950-61.
60. Horvath S, Laird NM. A discordant-sibship test for disequilibrium and linkage: no need for parental data. Am J Hum Genet. 1998;63:1886-97.
61. Martin ER, Monks SA, Warren LL, Kaplan NL. A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am J Hum Genet. 2000;67:146-54.
62. Rabinowitz D, Laird N. A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. Hum Hered. 2000;50:211-23.
63. Schulze TG, McMahon FJ. Genetic association mapping at the crossroads: which test and why? Overview and practical guidelines. Am J Med Genet. 2002;114:1-11.
64. Lange C, Silverman EK, Xu X et al. A multivariate family-based association test using generalized estimating equations: FBAT-GEE. Biostatistics. 2003;4:195-206.
65. Qian D. Haplotype sharing correlation analysis using family data: a comparison with family-based association test in the presence of allelic heterogeneity. Genet Epidemiol. 2004;27:43-52.
66. Horvath S, Wei E, Xu X et al. Family-based association test method: age of onset traits and covariates. Genet Epidemiol. 2001;21 Suppl 1:S403-8.
67. Abecasis GR, Cardon LR, Cookson WO. A general test of association for quantitative traits in nuclear families. Am J Hum Genet. 2000;66:279-92.
68. Clayton D, Jones H. Transmission/disequilibrium tests for extended marker haplotypes. Am J Hum Genet. 1999;65:1161-9.
69. Laird NM, Blacker D, Wilcox M. The sib transmission/disequilibrium test is a Mantel-Haenszel test. Am J Hum Genet 1998;62:450-8.

CHAPTER 10

Summary in English, Dutch and Polish

JJ Hottenga

Summary

Chapter one: Introduction

In the introduction chapter, issues of mapping genes in complex neurological disorders were discussed. The complexity of these disorders often depends on the high prevalence, multi-factorial aetiology and heterogeneity. Genetic risk factors explain to a certain degree the aetiology of disorders. The impact of some genetic risk factors may only slightly increase the risk, whereas the impact of other factors can be much more prominent; the Mendelian forms of the disorder. The mapping of (susceptibility) genes has been conducted by using different methods, and dependent on the risk, it led to various outcomes. A summary of these methods, including linkage, sib-pair, association and TDT analysis, was given together with the advantages and disadvantages of the approaches. Finally, a brief layout was presented on the genetics of neurological disorders that are discussed in the thesis namely migraine, Alzheimer's disease (AD) and epilepsy.

Chapter two: Population stratification

Association studies have frequently been criticized because of the failure to replicate results. Population stratification in cases and controls is often cited as being one of the major causes of this problem. The aim of the study was to examine how much population stratification and diversity, which is caused by genetic drift, is required to lead to spurious associations. Genetically isolated populations were simulated with various degrees of founder effects and genetic drift. Our study shows that in case one marker is tested, the probability of finding a spurious association with an increased risk of 1.5 is often less than 5%. Unless multiple markers are tested, population stratification is likely not a major issue in both study replication, as well as causing false-positive studies. Only when the genetic drift is very strong, or in case when stratification of the two populations is extremely discordant, the risk for spurious association

exceeds 5%. The application of methods that test and correct population stratification should then be applied to correct the confounding.

Chapter three: A straightforward approach to overcome possible false-positive associations in studies of gene-gene interaction

Research of gene-gene interactions is important in unraveling risk factors involved in complex traits. However, association-based gene interaction studies are susceptible to false-positive – and false-negative findings. One of the reasons may be stratification of a limited number of cases and controls, leading to a small number of subjects in each stratum and a large gene frequency variation over strata. A straightforward approach to study this problem is the testing of association between two gene variants in controls. Here, an association between two unlinked genes should not be present. A large odds ratio or finding association is therefore a good indication for aberrant changes in control allele frequencies of the two genes. From this approach it also follows that one may improve the statistical power of the study, and reduce the probability of false-positive findings, by genotyping extra controls for the second gene in the limiting stratum; the control carriers of the risk allele of the first gene studied. This may be useful in large-scale epidemiological studies, in which multiple genes often have been characterized. In this chapter the approach was applied in empirical data of an AD study with the Apolipoprotein E (APOE) and presenilin-1 (PSEN1) genes. The results showed that most of the evidence for gene interaction between APOE and PSEN genes was indeed present in the allele frequency variation of controls.

Chapter four: The 3p21.1-p21.3 hereditary vascular retinopathy locus increases the risk for Raynaud phenomenon and migraine

In some families with rare cerebrovascular disorders that have Mendelian segregation, migraine can be a co-occurring phenomenon. Based on this co-occurrence, it can be assumed that the gene(s) involved in these Mendelian

disorders are also risk factors for migraine. In a large Dutch family with hereditary vascular retinopathy (HVR) , migraine and Raynaud phenomenon, a locus for HVR was identified on chromosome 3p21.1-21.3. As Raynaud phenomenon, migraine and HVR all share a vascular aetiology, we tested if this locus increased the susceptibility for Raynaud phenomenon and migraine. A problem with testing association in families is that the individual observations are related. TDT analyses on family members affected with migraine and/or Raynaud phenomenon showed no significant risk increase (probably due to low power) but the discordant sibling transmission disequilibrium analyses revealed that the HVR haplotype harbors a susceptibility factor for Raynaud phenomenon and migraine. The identification of the HVR gene will improve the understanding of the pathophysiology of HVR, Raynaud phenomenon and migraine.

Chapter five: Segregation analysis in Dutch migraine families

A homogenous phenotype can largely improve the performance of linkage studies. However, in migraine literature, the use of separate or combined analysis of migraine with (MA) – and without aura (MO) types, as being affected, has been a controversial issue. Another issue is the inclusion of patients being affected with both types of migraine attacks. Studies considering both a single migraine type, as well as combined migraine types have been successful in mapping migraine genes. In this study, the segregation of migraine types was studied in 55 Dutch families, in order to evaluate whether people with MA and MO should be considered affected for MO linkage analysis. A trio (parents-patient) approach, as well as a complex segregation analysis with POINTER was employed. The results in trios show that focusing on a specific migraine type in linkage analysis may be favorable, based on the number of transmissions of related migraine types, compared to mismatching types. Furthermore, adding MA and MO affected persons appeared to have only little effect on conclusions about the segregation of MO in migraine families.

Chapter six: Involvement of the 4q21-24 migraine locus in Dutch migraine without aura families

Gene mapping of the common forms of migraine, MA and MO, has been challenging because of the complex genetics and the high frequency of these disorders. Recently, however, several loci for MA and MO have been identified. In this study, seven Dutch families with apparent dominantly inherited MO were selected for a genome-wide scan (at 9 cM marker interval). In total, 392 markers were tested, and suggestive evidence for linkage was found for chromosomal region 4q21-q24. For marker D4S2361, the maximum multipoint LOD score of 1.98 was observed when analyzing all families combined. This study presents some evidence for replication of two previous studies in Finnish and Icelandic families, showing linkage to a region involved in MA and MO, respectively. It is tempting to speculate that the chromosome 4 locus might be important for both migraine types, although there still may be two genes within this locus. Future studies should shed more light on the susceptibility gene(s) in this region.

Chapter seven: A Dutch family with 'familial cortical tremor with epilepsy': clinical characteristics and exclusion of linkage to chromosome 8q23.3-q24.1

Familial Cortical Tremor with Epilepsy (FCTE) is an idiopathic generalized epilepsy of adult onset with autosomal dominant inheritance. FCTE is characterized by kinesigenic tremor and myoclonus of the limbs, generalized seizures, and electrophysiological findings consistent with cortical reflex myoclonus. Genetic analysis has been performed in five Japanese families. In all these families, linkage was shown to chromosome 8q23.3-q24.1. Here, we describe a Dutch family with clinical characteristics of cortical tremor with epilepsy. We tested genetic linkage to chromosome 8q23.3-q24.1. The clinical and electrophysiological findings were consistent with a diagnosis of FCTE, however linkage with chromosome 8q23.3-q24.1 was excluded. This finding led to the assumption that genetic heterogeneity for FCTE exists.

Chapter eight: Familial partial epilepsy with variable foci in a Dutch family: clinical characteristics and confirmation of linkage to chromosome 22q

Linkage analysis in a four-generation Dutch family with epilepsy fulfilling criteria of both ADNFLE and FPEVF was performed for known loci in these disorders. ADNFLE loci (located on chromosomes 1p21, 15q24, and 20q13.3) and FPEVF loci (located on chromosomes 2q36, and 22q11-q12) were tested. Epilepsy in this family was diagnosed in ten relatives. Seizures were mostly tonic, tonic-clonic, or hyperkinetic with a wide variety in symptoms and severity. Most interictal EEGs showed no abnormalities but some showed frontal, central, and/or temporal spikes and spike-wave complexes. Of two patients, an ictal EEG was available, showing fronto-temporal abnormalities in one and frontal and central abnormalities in the other. Genetic analysis revealed linkage to chromosome 22q11-q12 in this family, using a parametric approach with a reduced penetrance model. The frequent occurrence of seizures during daytime and the observation of interictal EEG abnormalities, originating from different cortical areas, were more in agreement with FPEVF. The observed linkage on chromosome 22q11-q12 supported this diagnosis and confirmed that the locus is responsible for this syndrome.

Chapter nine: General discussion

The general discussion of the thesis presents an overview of the results discussed in the previous chapters. Issues include the effect of population stratification related to recent findings, the feasibility of linkage analysis in various linkage studies and the effects of locus heterogeneity and family selection. In all issues some future perspectives were also given. Finally, some thoughts are given on the use of family-based association tests.

Samenvatting

Hoofdstuk één: Inleiding

In het inleidende hoofdstuk worden problemen bij het vinden van genen in complexe neurologische aandoeningen besproken. De complexiteit van deze aandoeningen hangt vaak af van de hoge prevalentie, de vele risico factoren die het ontstaan beïnvloeden en de grote heterogeniteit. Genetische risico factoren verklaren het ontstaan van deze aandoeningen voor een deel. Voor sommige genetische factoren zal het risico op de aandoening slechts gering toenemen, voor andere zal dit aandeel aanzienlijk hoger zijn; de Mendeliaanse vormen van de aandoening. Het vinden van risico genen met verschillende strategieën heeft, afhankelijk van de bijdrage aan het risico, geleid tot verschillend succes. Een samenvatting van deze strategieën, linkage, sib-pair, associatie en TDT analyse is gegeven samen met hun voor- en nadelen. Verder is ook een beperkt overzicht gegeven van de, tot nu toe, bekende genen voor de aandoeningen besproken in dit proefschrift namelijk, migraine, Alzheimer en epilepsie.

Hoofdstuk twee: Populatie stratificatie

Associatie studies zijn frequent onderhevig aan kritiek vanwege het feit dat veel resultaten niet kunnen worden gerepliceerd. Vaak wordt als oorzaak populatie stratificatie in patiënten en controles gegeven als de belangrijkste oorzaak van dit probleem. Het doel van de studie in dit hoofdstuk was om te onderzoeken hoeveel stratificatie en diversiteit, veroorzaakt door genetische drift van populaties, nodig is voor het ontwikkelen van fout positieve associaties. Genetisch geïsoleerde populaties werden hierom gesimuleerd met verschillende mate van founder effecten en genetische drift. De studie toont aan dat, wanneer één marker wordt getest, de kans op het vinden van een fout positieve associatie met een relatief risico van 1.5 vaak minder dan 5% is. Alleen wanneer de genetische drift erg sterk is, of wanneer de stratificatie van de populaties extreem discordant is in patiënten en controles, dan is het risico voor fout positieve associaties hoger dan 5%. De conclusie is dat populatie

stratificatie waarschijnlijk niet zo een groot effect heeft op de replicatie en kans op fout positieve studies, tenzij meerdere markers worden getest. Methoden die voor stratificatie testen en corrigeren, kunnen in dit geval worden gebruikt om deze confounding te voorkomen.

Hoofdstuk drie: Een eenvoudige methode voor het verhelpen van fout positieve associaties in gen-gen interactie studies

Het onderzoek naar gen-gen interacties is erg belangrijk voor het vinden van risico factoren voor complexe aandoeningen. Echter, op associatie gebaseerde gen interactie studies zijn gevoelig voor fout positieve - en negatieve bevindingen. Eén van de redenen kan het opdelen van patiënten en controles zijn, dat leidt tot te kleine aantallen en grote frequentie veranderingen in de verschillende strata. Een simpele methode om dit probleem te analyseren is het testen van associatie tussen twee genen in controles. Er zou geen associatie moeten bestaan tussen twee niet gekoppelde genen. Het vinden van een groot relatief risico is daarom een goede indicatie dat de genfrequenties van de twee genen in controles afwijkend zijn. Uit deze methode blijkt tevens dat de power en de kans op fout positieve bevindingen kan worden gereduceerd door het testen van contoles voor het specifieke stratum waar weinig waarnemingen zijn voor het tweede gen. Dit zijn de controledragers van het risico allel voor het eerste gen dat is onderzocht. Dit kan nuttig zijn in grote epidemiologische studies, waar verschillende genen zijn gekarakteriseerd. In dit hoofdstuk is deze methode toegepast op empirische data, waarbij de interactie tussen de genen Apolipoproteine E en Preseniline-1 in relatie tot Alzheimer is onderzocht. De resultaten laten zien dat het bewijs voor de interactie inderdaad afhankelijk was van de verschillen in allel frequentie in de controles.

Hoofdstuk vier: Het 3p21.1-p21.3 erfelijke vasculaire retinopathie locus verhoogt het risico op Raynaud fenomeen en migraine

In sommige families met een Mendeliaanse segregatie van zeldzame cerebrovasculaire aandoeningen kan migraine een bijkomende aandoening

zijn. Genen betrokken bij de cerebrovasculaire aandoeningen kunnen hierom tevens worden beschouwd als risico factoren voor het ontwikkelen van migraine. In een grote Nederlandse familie met erfelijke vasculaire retinopathie (HVR), migraine en Raynaud fenomeen werd een locus geïdentificeerd op chromosoom 3p21.1-21.3 voor HVR. Omdat migraine, Raynaud fenomeen en HVR alle drie een vasculaire etiologie hebben is getest of het HVR locus ook de gevoeligheid voor migraine en Raynaud fenomeen verhoogt. Een probleem bij het testen van associaties in families is dat de individuele observaties niet onafhankelijk zijn van elkaar. TDT analyses van de familie leden met migraine en Raynaud toonden niet aan dat het risico voor deze aandoeningen verhoogd is (waarschijnlijk door de lage detectie kans van de test). Echter, de discordante broer / zus transmissie disequilibrium analyses toonden wel aan dat het HVR haplotype waarschijnlijk een gevoeligheids factor voor Raynaud fenomeen en migraine bevat. Identificatie van het gen dat een rol speelt bij HVR zal dus ook leiden tot het geven van inzicht in de migraine en Raynaud fenomeen pathofysiologie.

Hoofdstuk vijf: Segregatie analyse in Nederlandse migraine families

Een homogeen fenotype kan de prestatie van linkage studies sterk verbeteren. Echter in de migraine literatuur blijft het gebruik van migraine met (MA) en zonder aura (MO) apart, of met beide aanvalstypen gecombineerd, een controversieel punt voor linkage analyse. Ook het toevoegen van personen met beide aanvalstypen is een probleem. Studies die alleen een enkel type als aangedaan hebben beschouwd, en ook studies die beide typen als aangedaan hebben beschouwd zijn succesvol geweest in het vinden van gen locaties. In deze studie is de segregatie van de verschillende migraine types bestudeerd in 55 Nederlandse families. Dit om te zien of mensen met zowel MA als MO als aangedaan beschouwd moeten worden in een MO linkage analyse. Een trio (ouders-kind) benadering en complexe segregatie analyse met POINTER is uitgevoerd om dit te testen. De resultaten geven aan dat het selecteren van een enkel migraine type voordeel kan geven, gebaseerd op het aantal keer dat

hetzelfde type wordt overgegeven ten opzichte van het discordante migraine type. Verder blijkt dat het toevoegen van personen met een gemengd migraine aanvalstype weinig effect heeft op de gegeven segregatie van MO in migraine families.

Hoofdstuk zes: Betrokkenheid van het 4q21-q24 migraine locus in Nederlandse migraine zonder aura families

Het identificeren van genlocaties voor de frequente vormen van migraine is een uitdaging geweest vanwege de complexe genetica en de hoge frequentie van deze aandoeningen. Momenteel zijn er echter een aantal locaties geïdentificeerd voor beide aanvalstypen MA en MO. In deze studie zijn 7 Nederlandse MO families geselecteerd voor een genoom scan (met een 9 cM marker interval). In totaal zijn er 392 markers getest en suggestief bewijs voor linkage is gevonden voor de regio 4q21-q24 op chromosoom 4. Voor marker D4S2361 is een maximum LOD score gevonden van 1.98 wanneer alle families gecombineerd zijn geanalyseerd. Deze studie geeft beperkt bewijs voor replicatie van migraine locaties die ook zijn gevonden in twee andere studies uit Finland en IJsland voor respectievelijk MA en MO. Het is aantrekkelijk om te speculeren dat deze locatie dus betrokken is bij beide typen migraine aanvallen. Het kan echter ook zo zijn dat er twee genen zijn die een effect hebben op ieder aanvalstype apart. Toekomstige studies moeten meer informatie geven over het gen of genen op deze locatie die bijdragen aan het risico op migraine.

Hoofdstuk zeven: Een Nederlandse familie met familiaire corticale tremor met epilepsie: klinische karakteristieken en exclusie van linkage op chromosoom 8q23.3-q24.1

Familiaire corticale tremor met epilepsie (FCTE) is een autosomaal dominante ideopatische gegeneraliseerde epilepsie die ontstaat in volwassenen. FCTE wordt gekarakteriseerd door een kinesiogene tremor en myoclonus in de ledematen, infrequente myclonische en generaliseerde tonisch clonische

aanvallen en elektrofysiologische resultaten die passen bij een corticale reflex myclonus. Genetische linkage analyse is gedaan in vijf Japanse families. In al deze families is linkage aangetoond op chromosoom 8q23.3-q24.1. In deze studie hebben we een Nederlandse familie onderzocht met klinische verschijnselen van corticale tremor met epilepsie voor linkage op chromosoom 8q23.3-q24.1. De klinische en elektrofysiologische verschijnselen kwamen overeen met de diagnose van FCTE, echter linkage op chromosoom 8q23.3-q24.1 werd uitgesloten. Deze bevinding leidde tot de assumptie dat er waarschijnlijk heterogeniteit is voor FCTE.

Hoofdstuk acht: Familiaire partiele epilepsie met variabele foci in een Nederlandse familie: klinische karakteristieken en bevestiging van linkage op chromosoom 22q

Linkage analyse is gedaan voor locaties in een Nederlandse familie van vier generaties met epilepsie die voldeed aan zowel nachtelijke frontaal kwab epilepsie (ADNFLE), als wel familiaire partiele epilepsie met variabele foci (FPEVF). Epilepsie in deze familie werd gediagnosticeerd in tien familieleden. De aanvallen waren voornamelijk tonisch, tonisch clonisch, of hyperkinetisch met variabele symptomen en ernst. Het merendeel van de interictale elektroencefalogrammen (EEGs) liet geen afwijkingen zien, maar soms werden frontale, centrale en/of temporale pieken en piekgolfcomplexen waargenomen. Van twee patiënten was een ictaal EEG beschikbaar. Deze vertoonde bij de één frontotemporale afwijkingen en bij de ander frontale en centrale afwijkingen. ADNFLE locaties werden getest op chromosomen 1p21, 15q24 en 20q13.3 en FPEVF locaties op chromosomen 2q36 en 22q11-q12. Analyse van de chromosoom gebieden, gebruik makend van een parametrisch model met gereduceerde penetrantie, leidde tot linkage op chromosoom 22q11-q12 in deze familie. De frequente aanvallen tijdens de dag, de observatie van interictale EEG afwijkingen met een oorsprong in diverse corticale gebieden samen met de gevonden linkage op chromosoom 22q11-q12 geven aan dat de diagnose FPEVF meer waarschijnlijk is.

Hoofdstuk negen: Algemene discussie

De algemene discussie van dit proefschrift geeft een korte algemene samenvatting van de resultaten gevonden in de verschillende onderzoeken. De discussie omvat verschillende onderwerpen, namelijk het effect van populatie stratificatie gerelateerd aan recente bevindingen, het nut van parametrische linkage analyse in complexe aandoeningen en het effect van locus heterogeniteit en selectie. Verder worden wat gedachten gegeven over het gebruik van op familie gebaseerde associatie testen. Voor alle punten worden wat verwachtingen gegeven voor de toekomst.

Podsumowanie

Rozdział pierwszy: Wstęp

W rozdziale wprowadzającym zostały przedyskutowane kwestie lokalizacji genów w kompleksowych zaburzeniach neurologicznych. Kompleksowość takich zaburzeń często zależy od wysokiej częstotliwości, wieloczynnikowego pochodzenia oraz heterogeniczności. Genetyczne faktory ryzyka wyjaśniają do pewnego stopnia pochodzenie zaburzeń. Wpływ niektórych genetycznych faktorów ryzyka może tylko nieznacznie zwiększyć ryzyko, podczas gdy wpływ innych czynników może być dużo bardziej znaczący; Mendeliańskie formy zaburzenia. Lokalizowanie genów (podatności) zostało przeprowadzone przy użyciu rożnych metod i w zależności od ryzyka, zakończyło się różnorodnym powodzeniem. Podsumowanie tych metod, wliczając sprzężenie, pary rodzeństwa, kojarzenie i analiza TDT, zostało podane razem z zaletami i wadami tych metod. Na końcu został przedstawiony krótki zarys zaburzeń genetycznych i neurologicznych, które są przedyskutowane w tej pracy, a mianowicie migrena, choroba Alzheimera (AD) oraz epilepsja.

Rozdział drugi: Stratyfikacja ludności

Badania skojarzeniowe były częstokrotnie krytykowane, ze względu na niepowodzenie w replikowaniu wyników. Stratyfikacja ludności w grupie dotkniętych i porównywalnej grupie nie- dotkniętych jest często cytowana jako będąca jedną z ważniejszych przyczyn tego problemu. Celem tego badania było sprawdzenie na ile stratyfikacja ludności i różnorodność, spowodowana przez genetyczny dryft, jest potrzebna do wyciągnięcia pozornych skojarzeń. Populacje odizolowane genetycznie były symulowane z różnorodnym stopniem efektów założyciela i genetycznego dryftu. Nasze badanie pokazuje, że w przypadku, gdy jeden marker jest testowany, prawdopodobieństwo znalezienia pozornego skojarzenia z podniesionym ryzykiem wynoszącym 1.5, jest często mniejsze niż 5%. Tylko w przypadku, gdy dryft genetyczny jest bardzo mocny, lub, gdy stratyfikacja dwóch populacji jest skrajnie niezgodna, ryzyko pozornego skojarzenia przekracza

5%. Stratyfikacja populacji jest prawdopodobnie nie najważniejszą kwestią w przypadku replikacji badania lub przyczynianiu się do pozornie- pozytywnych badań, chyba, że wielorakie markery są testowane. Zastosowanie metod, które testują i poprawiają stratyfikacje ludności powinny wtedy być użyte do korekcji nieprawidłowości.

Rozdział trzeci: Bezpośrednie podejście do pokonania możliwych pozornie-pozytywnych skojarzeń w badaniach interakcji gen-gen.

Badanie interakcji gen- gen jest ważne w rozkładaniu faktorów ryzyka w kompleksowych cechach genetycznych. Jednakże, badania interakcji genetycznej na podstawie skojarzenia są podatne na pozornie- pozytywne i pozornie- negatywne znalezienia. Jednym z powodów może być stratyfikacja ograniczonej liczby członków z grupy dotkniętych i nie- dotkniętych, prowadząc do małej ilości podmiotów w każdej linii i dużego wahania w częstotliwości genu w liniach. Bezpośrednie podejście do przestudiowania tego problemu jest przez testowanie skojarzenia pomiędzy odmianami dwóch genów w grupie nie- dotkniętych. W tym przypadku skojarzenie pomiędzy dwoma niepołączonymi genami nie powinno być obecne. Dość duże ryzyko względne lub skojarzenie znalezienia jest dlatego dobrym wskazaniem zaskakujących zmian w częstotliwościach allele w grupie nie- dotkniętych tych dwóch genów. Z tego podejścia wynika również, że ktoś może poprawić moc statystyczną tego badania i zmniejszyć możliwość pozornie-pozytywnych znalezisk poprzez ustalenie genotypów dodatkowych członków z grupy nie- dotkniętych dla drugiego genu w linii ograniczającej; nosiciele dotknięci odmianą genu ryzyka pierwszego genu przebadanego. To może być przydatne w badaniach epidemiologicznych na dużą skale, w których wielokrotność genów często była charakteryzowana. W tym rozdziale badanie było zastosowane w danych empirycznych badania AD z genami Apolipoprotein E (APOE) i presenilin-1 (PSEN1). Wyniki pokazały, ze większość dowodów na interakcje genów pomiędzy genami APOE i PSEN

były rzeczywiście obecne w różnorodności częstotliwości odmiany genu w grupie nie- dotkniętych.

Rozdział czwarty: Lokalizacja 3p21.1-p21.3 dziedzicznego naczyniowego uszkodzenia siatkówki podnosi ryzyko zjawiska Raynauda oraz migreny

W przypadku niektórych rodzin z rzadkimi zaburzeniami mózgowo-naczyniowymi, które maja segregacje Mendeliana, migrena może być zjawiskiem współwystępującym. Na podstawie takiego współwystępowania można założyć, ze gen(y) związane z tymi zaburzeniami Mendeliana sa również współczynnikami ryzyka dla migreny. W obszernej holenderskiej rodzinie z dziedzicznym naczyniowym uszkodzeniem siatkówki (HVR), migreną i zjawiskiem Raynauda, lokalizacja dla HVR został zidentyfikowany na chromosomie 3p21.1-21.3. Jako że zjawisko Raynauda, migrena i HVR maja wspólne pochodzenie naczyniowe, przetestowaliśmy czy ta lokalizacja podniosła podatność na zjawisko Raynauda i migrenę. Problem z testowaniem skojarzenia w rodzinach jest taki, że indywidualne obserwacje są połączone. Analizy TDT na członkach rodziny z migreną i/lub zjawiskiem Raynauda nie pokazały znaczącego wzrostu ryzyka (prawdopodobnie ze względu na niska moc), ale analizy braku równowagi w transmisji pomiędzy przeciwstawnym rodzeństwem pokazały, że haplotyp HVR ukrywa czynnik podatności na zjawisko Raynauda i migrenę. Identyfikacja genu HVR poprawi zrozumienie patofizjologii HVR, zjawiska Raynauda, migreny oraz HVR.

Rozdział piąty: Analiza segregacji w holenderskich rodzinach z migreną.

Homologiczny fenotyp może w dużej mierze poprawić wydajność badań sprzężenia. Jednakże w literaturze o migrenie, użycie osobnej lub mieszanej analizy migreny typu z (MA) i bez aury (MO) jako będącej dotkniętą było i nadal jest zagadnieniem kontrowersyjnym. Inną kwestią jest włączenie pacjentów będących dotkniętymi z obydwoma rodzajami ataków migreny. Badania dotyczące obu rodzajów migreny, pojedynczej jak i połączonej, były pomyślne w zlokalizowaniu genów migreny. W tym badaniu segregacja typów

migreny była przestudiowana w 55-ciu holenderskich rodzinach, żeby ocenić czy ludzie z MA i MO powinni być uważani za dotkniętych do analizy sprzężenia MO. Zostało użyte podejście trójkowe (rodzice-dziecko) zarówno jak analiza kompleksowej segregacji na programie POINTER. Rezultaty w trójkach pokazują, ze skupianie się na specyficznym rodzaju migreny w analizie sprzężenia może być korzystne opierając się na liczbie transmisji spokrewnionych rodzajów migreny porównanych do rozbieżnych rodzajów. Co więcej, dodanie osób dotkniętych MA i MO okazało się mające tylko mały wpływ na wnioski o segregacji MO w rodzinach z migreną.

Rozdział szósty: Udział locusu migreny 4q21-24 w holenderskich rodzinach bez aury.

Lokalizacja genów powszechnej odmiany migreny, MA i MO, było ambitne ze względu na kompleksową genetykę i wysoką częstotliwość tych zaburzeń. Jednakże niedawno, kilka locusów dla MA i MO zostało zidentyfikowanych. W tym badaniu, siedem holenderskich rodzin z wyraźnie dominującym odziedziczonym MO zostało wyselekcjonowanych do szeroko- genomowego skanowania (co interwa 9 cM marker). W sumie zostały przetestowane 392 markery i sugestywne dowody dla sprzężenia zostały znalezione w przypadku chromosomowego obszaru 4q21-q24. Dla markera D4S2361, maksymalny wielopunkt LOD o wyniku 1.98 został zaobserwowany przy analizowaniu wszystkich rodzin łącznie. To badanie ukazuje pewne dowody na powtórzenie dwóch wcześniejszych badań na finlandzkich i islandzkich rodzinach, pokazując powiązanie odpowiednio z obszarem związanym z MA i MO. Kusząca jest spekulacja, ze locus chromosomu 4 może być ważny dla obu rodzajów migreny, pomimo tego, że mogą tam być nadal dwa geny w obrębie tego locusu. Dalsze badania powinny rzucić więcej światła na wrażliwość genu/genów w tym obszarze.

Rozdział siódmy: Holenderska rodzina z „rodzinnym korowym wstrząsem z epilepsją": kliniczne cechy i wykluczenie sprzężenia do chromosomu 8q23.3-q24.1

Rodzinny korowy wstrząs z epilepsja (FCTE) jest idiopatyczną uogólnioną epilepsją początku dojrzałości z autosomalnie dominująca dziedzicznością. FCTE charakteryzuje się kinesygenicznym wstrząsem i myoklonusem kończyn, ogólnym atakiem, i elektrofizjologicznymi znaleziskami zgodnymi z korowym odruchem myoklonusu. Analiza genetyczna została przeprowadzona na pięciu rodzinach japońskich. We wszystkich rodzinach, zostało pokazane sprzężenie z chromosomem 8q23.3-q24.1. Tutaj opisujemy rodzinę holenderską z klinicznymi cechami korowego wstrząsu z epilepsją. Przetestowaliśmy genetyczne sprzężenie dla chromosomu 8q23.3-q24.1. Kliniczne elektrofizjologiczne odkrycia były zgodne z diagnozą FCTE, jednakże sprzężenie z chromosomem 8q23.3-q24.1 zostało wykluczone. To odkrycie poprowadziło do założenia, że heterogeniczność genetyczna dla FCTE istnieje.

Rozdział ósmy: Rodzinna częściowa epilepsja ze zmiennym ogniskiem u holenderskiej rodziny: kliniczne cechy i potwierdzenie sprzężenia do chromosomu 22q

Analizy sprzężenia w czterogeneracyjnej rodzinie holenderskiej z epilepsją spełniającą wymagania zarówno ADNFLE, jak i FPEVF zostały wykonane dla znanych locusów towarzyszących takim zaburzeniom. Przetestowane zostały ADNFLE locusy (usytuowane na chromosomie 1p21, 15q24, i 20q13.3) oraz FPEVF locusy (usytuowane na chromosomie 2q36, i 22q11-q12). W tej rodzinie epilepsja została zdiagnozowana w przypadku dziesięciu członków rodziny. Ataki były przeważnie toniczne, toniczno-kloniczne, oraz hyperkinetyczne z dużą różnorodnością w symptomach i surowości. Większość miedzyudarowych badań EEG nie pokazało żadnych nieprawidłowości, lecz parę pokazało przednich, środkowych, i/lub skroniowych nagłych skoków oraz zespołów skokowo- falowych.

Udostępnione zostało udarowe EEG dwóch pacjentów pokazujące przednio-skroniowe nieprawidłowości w jednym i środkowe w drugim przypadku. Analiza genetyczna ujawniła sprzężenie dla chromosomu 22q11-q12 w tej rodzinie używając metody parametrycznej z modelem niemającym całkowitego prawdopodobieństwa uaktywnienia się choroby, będąc nosicielem tego genu. Częste pojawianie się ataków podczas dnia i obserwacja nieprawidłowości w miedzyudarowych badaniach EEG rodzących się w rożnych obszarach korowych były w zgodzie z FPEVF. Zaobserwowane sprzężenie na chromosomie 22q11-q12 podtrzymało ta diagnozę i potwierdziło, że locus jest odpowiedzialne za ten syndrom.

Rozdział dziewiąty: Ogólne omówienie

Ogólne omówienie pracy doktorskiej prezentuje ogólny przegląd rezultatów przedyskutowanych w rożnych rozdziałach. Kwestie te obejmują rezultat stratyfikacji ludności mający związek z niedawnymi odkryciami, przydatności analizy sprzężenia w rożnych badaniach sprzężenia i skutki heterogenetyczności locusu i selekcja rodziny. We wszystkich kwestiach były również dane perspektywy na przyszłość. Na koniec wspomniane jest parę myśli o użyciu testów kojarzenia na podstawie rodzin.

Curriculum vitae

List of abbreviations

Acknowledgements

Curriculum vitae

Jouke- Jan was born on the 20th of October 1974 in Leiden. In 1993 he graduated from his Voortgezet Wetenschappelijk Onderwijs at the Pieter Groen College in Katwijk. In September 1993, he started his studies of Biomedical Sciences at the University of Leiden. He obtained his 'doctoraal' degree in 1998. During that time he did three projects related to genetic epidemiology. With Prof. Dr. P.E. Slagboom (TNO, Leiden) he did an association study, examining the relation between the PAI 4/5 polymorphism and risk of septic shock in children having severe meningococcal infection. Subsequently, he studied the relevance of the Calcium channel subunits in common forms of migraine at the departments of Neurology and Human Genetics in Leiden under guidance of Prof. Dr. R.R. Frants and Prof. Dr. M.D. Ferrari. His third project was at the Erasmus University Rotterdam, under guidance of Prof. Dr. C.M. van Duijn, and involved the effects of population stratification for some commonly tested polymorphisms. His interest led to a proposal from his additional mentor for the last two projects, Dr. L.A. Sandkuijl, which involved continuation of the work, at both Rotterdam and Leiden Universities to apply genetic epidemiologic approaches in population and family based designs. He received his Master's Degree in Genetic Epidemiology at the Erasmus University of Rotterdam in 2001, which subsequently led to the defense of his thesis "Genetic epidemiological approaches in complex neurological disorders" in 2005. Currently, Jouke- Jan is working at the Netherlands Twin Registry under the guidance of Prof. Dr. D.I. Boomsma and Prof. Dr. E.J.C. de Geus.

List of abbreviations

AD	Alzheimer's disease
ADLTE	Autosomal dominant lateral temporal epilepsy
ADNFLE	Autosomal dominant nocturnal frontal lobe epilepsy
ADPEAF	Autosomal dominant partial epilepsy with auditory features
AIC	Akaike information criterion
APOE	Apolipoprotein E
APP	Amyloid precursor protein
BAFME	Benign adult familial myoclonic epilepsy
BFNC	Benign familial neonatal convulsions
CADASIL	Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy
CI	Confidence interval
CRV	Autosomal dominant cerebroretinal vasculopathy
CT	Computed tomography
DAT	Discordant alleles test
DNA	Deoxyribonucleic acid
EEG	Electroencephalogram
EMG	Electromyogram
FAME	Familial adult myoclonic epilepsy
FBAT	Family based association test
FCTE	Cortical tremor with epilepsy
FHM	Familial hemiplegic migraine
FPEVF	Familial partial epilepsy with variable foci
GABA	G-aminobutyric acid
GDB	Genome database
GEFS+	Generalized epilepsy with febrile seizures
g-SEPs	Giant somatosensory evoked potentials
GSL	General single locus model

HERNS	Hereditary endotheliopathy with retinopathy, nephropathy and stroke
HRR	Haplotype relative risk method
HVR	Hereditary vascular retinopathy
HWE	Hardy-Weinberg equilibrium
IBD	Identity by descent
IBS	Identity by state
IHS	International Headache Society
LUMC	Leiden University Medical Center
MA	Migraine with aura
MA/MO	Mixed MA and MO migraine type
MERRF	Mitochondrial encephalomyopathy with ragged-red-fibres
MO	Migraine without aura
MRI	Magnetic resonance imaging
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Diseases Association
OR	Odds ratio
PCR	Polymerase chain reaction
PDT	Pedigree disequilibrium test
QTLs	Quantitative trait loci
RC-TDT	Reconstruction combined-transmission disequilibrium test
SCA	Spinocerebellar ataxia
SDT	Discordant sibship test
SEP	Somatosensory evoked potentials
SNP	Single nucleotide polymorphism
S-TDT	Sib transmission disequilibrium test
TDT	Transmission disequilibrium test

Acknowledgements

Many people have worked with me on the development of this thesis. I want to thank all for their contributions, guidance, discussions, suggested improvements, patience and friendly support. I wish all of you the best in future life and science. In addition, I want to thank all study participants and families for their co-operation.