Association mapping of genomic microdeletions and common susceptibility variants predisposing to genetic generalized epilepsies

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Holger Sebastian Trucks

aus Köln

2013

Berichterstatter:	Prof. Dr. Peter Nürnberg
(Gutachter)	Prof. Dr. Bernd Wollnik

Tag der mündlichen Prüfung: 04.07.2013

Table of contents

Publications

Parts of this study have been published previously:

Rare exonic deletions of the RBFOX1 gene increase risk of idiopathic generalized epilepsy. Lal D, **Trucks H**, Møller RS, Hjalgrim H, Koeleman BP, de Kovel CG, Visscher F, Weber YG, Lerche H, Becker F, Schankin CJ, Neubauer BA, Surges R, Kunz WS, Zimprich F, Franke A, Illig T, Ried JS, Leu C, Nürnberg P, Sander T; EMINet Consortium; EPICURE Consortium. Epilepsia. 2013 Feb; 54(2):265-71. doi: 10.1111/epi.12084. Epub 2013 Jan 25.PMID: 23350840

Exon-disrupting deletions of NRXN1 in idiopathic generalized epilepsy. Møller RS, Weber YG, Klitten LL, **Trucks H**, Muhle H, Kunz WS, Mefford HC, Franke A, Kautza M, Wolf P, Dennig D, Schreiber S, Rückert IM, Wichmann HE, Ernst JP, Schurmann C, Grabe HJ, Tommerup N, Stephani U, Lerche H, Hjalgrim H, Helbig I, Sander T; EPICURE Consortium. Epilepsia.2013 Jan 7.doi: 10.1111/epi.12078. [Epub ahead of print] PMID:23294455

Genome-wide association analysis of genetic generalized epilepsies implicates susceptibility loci at 1q43, 2p16.1, 2q22.3 and 17q21.32. EPICURE Consortium, EMINet Consortium, Steffens M, Leu C, Ruppert AK, Zara F, Striano P, Robbiano A, Capovilla G, Tinuper P, Gambardella A, Bianchi A, La Neve A, Crichiutti G, de Kovel CG, Kasteleijn-Nolst Trenité D, de Haan GJ, Lindhout D, Gaus V, Schmitz B, Janz D, Weber YG, Becker F, Lerche H, Steinhoff BJ, Kleefuß-Lie AA, Kunz WS, Surges R, Elger CE, Muhle H, von Spiczak S, Ostertag P, Helbig I, Stephani U, Møller RS, Hjalgrim H, Dibbens LM, Bellows S, Oliver K, Mullen S, Scheffer IE, Berkovic SF, Everett KV, Gardiner MR, Marini C, Guerrini R, Lehesjoki AE, Siren A, Guipponi M, Malafosse A, Thomas P, Nabbout R, Baulac S, Leguern E, Guerrero R, Serratosa JM, Reif PS, Rosenow F, Mörzinger M, Feucht M, Zimprich F, Kapser C, Schankin CJ, Suls A, Smets K, De Jonghe P, Jordanova A, Caglayan H, Yapici Z, Yalcin DA, Baykan B, Bebek N, Ozbek U, Gieger C, Wichmann HE, Balschun T, Ellinghaus D, Franke A, Meesters C, Becker T, Wienker TF, Hempelmann A, Schulz H, Rüschendorf F, Leber M, Pauck SM,

Trucks H, Toliat MR, Nürnberg P, Avanzini G, Koeleman BP, Sander T. Hum Mol Genet. 2012 Dec 15;21(24):5359-72. doi: 10.1093/hmg/dds373. Epub 2012 Sep 4. PMID: 22949513

Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. deKovel CG*, **Trucks H***, Helbig I*, Mefford HC, Baker C, Leu C, Kluck C, Muhle H, von Spiczak S, Ostertag P, Obermeier T, Kleefuss-Lie AA, Hallmann K, Steffens M, Gaus V, Klein KM, Hamer HM, Rosenow F, Brilstra EH, Trenité DK, Swinkels ME, Weber YG, Unterberger I, Zimprich F, Urak L, Feucht M, Fuchs K, Møller RS, Hjalgrim H, De Jonghe P, Suls A, Rückert IM, Wichmann HE, Franke A, Schreiber S, Nürnberg P, Elger CE, Lerche H, Stephani U, Koeleman BP, Lindhout D, Eichler EE, Sander T. Brain. 2010 Jan; 133(Pt 1):23-32. Epub 2009 Oct 20. PMID: 19843651 (* shared first authorship)

Abstract

Approximately 3% of the general population is affected by epilepsy during lifetime, making epilepsy one of the most common neurological diseases. Genetic generalized epilepsies (GGE) are the most common of genetic epilepsies and account for 20-30% of all epilepsies. GGE is subdivided into genetically determined subgroups with gradual transition, including genetic absence epilepsies (GAE), juvenile myoclonic epilepsy (JME), and epilepsy with generalized tonic-clonic seizures (EGTCS). In spite of a high heritability rate of 80% and a predominant genetic etiology, the genetic factors predisposing to GGE are still mostly unknown. In the present study, we carried out association studies to investigate whether genomic microdeletions and common susceptibility variants increase risk for GGE.

To test the common disease/common variant hypothesis, genome-wide association studies (GWAS) were performed in several GGE cohorts using case-control and family-based study designs. For analysis, all patients were either pooled or stratified according to the subgroup they belong to in order to detect common or subgroup-specific risk factors, respectively. The GWAS comprised a case-control cohort of 1,523 European GGE patients and 2,454 German controls and a sample cohort of 566 European parent-offspring trios. Meta-GWAS analyses revealed significant association ($P < 5.0 \times 10^{-8}$) with GGE at 2p16.1 (rs35577149, meta-analysis $P = 1.65 \times 10^{-8}$, OR[C] = 0.78, 95% CI 0.71 – 0.86). Significant association with JME was detected at 1q43 (rs12059546, meta-analysis $P = 2.27 \times 10^{-8}$, OR[G] = 1.53, 95% CI 1.33 – 1.78). Suggestive evidence for association ($P < 1.0 \times 10^{-5}$) was found for GGE at 8q12.2 (rs6999304, meta-analysis $P= 1.77 \times 10^{-6}$, OR[G] = 1.33, 95% CI 1.17 - 1.51) and for GAE at 2q22.3 (rs75917352, meta-analysis $P = 1.41 \times 10^{-7}$, OR[T] = 0.67, 95% CI 0.58 – 0.79). The associated regions harbor high-ranking candidate genes: *CHRM3* at 1q43, *VRK2* at 2p16.1, and *ZEB2* at 2q22.3. Further replication efforts are necessary to elucidate whether these positional candidate genes contribute to the heritability of the common GGE syndromes.

Exploring the rare variant/common disease hypothesis, we investigated the impact of six recurrent microdeletions on the genetic risk of GGE at the genomic hotspot regions 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.11, and 22q11.2, which had been

implicated as rare genetic risk factors in a wide range of neurodevelopmental disorders. Recurrent microdeletions were assessed in 1,497 European GGE patients, 5,374 controls, and 566 GGE trios using high-resolution SNP microarrays. Considering all six microdeletion hot spots together, we found a significant excess of these microdeletions in 2,563 GGE patients versus 5,940 controls ($P < 2.20 \times 10^{-16}$, OR = 7.65, 95% CI 4.59 – 13.18). Individually, significant associations with GGE were observed for the microdeletions at 15q11.2 ($P = 1.12 \times 10^{-4}$, OR = 3.59, 95% CI 1.80 – 7.25), 15q13.3 ($P = 5.48\times 10^{-9}$) and 16p13.11 ($P = 4.42\times 10^{-6}$, OR = 17.39, 95% CI 3.86 – 159.88).

In a candidate-gene approach, we tested whether exon-disrupting/removing microdeletions in the genes encoding *NRXN1* and *RBFOX1* confer susceptibility for GGE. We found a significant association with GGE at both loci (*NRXN1*: P = 0.0049; *RBFOX1*: P = 0.0083). However, high phenotypic variability and incomplete penetrance, resulting in apparently imperfect segregation, indicate that partial *NRXN1* and *RBFOX1* deletions represent susceptibility factors rather than highly penetrant mutations.

The present study substantiates a role of both genomic microdeletions and common susceptibility variants in the genetic predisposition of common GGE syndromes. We strengthened the statistical evidence for associations of genetic variants at 1q43, 2p16.1, and 2q23.2 with GGE syndromes and identified a novel susceptibility locus at 8q12.2. Although individually rare, the associations of all microdeletions at 15q11.2, 15q13.3, 16p13.3, *NRXN1*, and *RBFOX1* taken together contribute significantly to the genetic variance of GGE.

Zusammenfassung

Etwa 3% der allgemeinen Bevölkerung ist zu Lebzeiten von Epilepsie betroffen. Somit ist Epilepsie eine der häufigsten neurologischen Erkrankungen. Genetische generalisierte Epilepsien (GGE) sind die häufigsten genetisch bedingten Epilepsien und verantwortlich für 20-30% aller Epilepsien. Die GGE lassen sich in Subgruppen mit fließenden Übergängen im Phänotyp unterteilen. Zu diesen zählen die genetische Absence-Epilepsie (GAE), die juvenile myoklonische Epilepsie (JME) und die Epilepsie mit generalisierten tonisch-klonischen Anfällen (EGTCS). Trotz einer hohen Erblichkeit von 80% und einer überwiegend genetischen Ätiologie, sind die genetischen Faktoren zur Entstehung von GGE immer noch weitgehend unbekannt. In der vorliegenden Studie führten wir Assoziationsstudien durch, um zu untersuchen, ob genomische Mikrodeletionen und häufige Suszeptibilitäts-Varianten das Risiko für GGE erhöhen.

Um die „common disease/common variant"-Hypothese zu testen, wurden genomweite Assoziationsstudien (GWAS) in mehreren GGE-Kohorten unter Verwendung von fall-kontroll- und familien-basierten Studiendesigns durchgeführt. Für die Analyse wurden die Patienten sowohl alle zusammen als auch nach der Zugehörigkeit zur jeweiligen Untergruppe separat betrachtet, um gemeinsame oder untergruppen-spezifische Risikofaktoren zu erkennen. Die GWAS umfasste eine Fall-Kontroll-Kohorte mit 1.523 GGE-Patienten europäischen Ursprungs und 2.454 deutschen Kontrollen, sowie einer Probenkohorte aus 566 europäischen Eltern-Kind-Trios. Meta-GWAS-Analysen ergaben signifikante Assoziation ($p < 5{,}0 \times 10^{-8}$) mit GGE auf 2p16.1 (rs35577149, Meta-Analyse $P = 1{,}65 \times 10^{-8}$, oder [C] = 0,78, 95% CI von 0,71 bis 0,86). Signifikante Assoziation mit JME wurde auf 1q43 nachgewiesen (rs12059546, Meta-Analyse $P = 2{,}27 \times 10^{-8}$, OR [G] = 1,53, 95% CI 1,33 bis 1,78). Suggestive Assoziation ($P < 1{,}0 \times 10^{-5}$) mit GGE wurde auf 8q12.2 gefunden (rs6999304, Meta-Analyse $P = 1{,}77 \times 10^{-6}$, oder [G] = 1,33, 95% CI 1,17 bis 1,51), sowie für GAE auf 2q22.3 (rs75917352, Meta-Analyse $P = 1{,}41 \times 10^{-7}$, oder [T] = 0,67, 95% CI 0,58 bis 0,79). Die damit verbundenen Regionen beherbergen hochrangige Kandidatengene: *CHRM3* auf 1q43, *VRK2* auf 2p16.1 und *ZEB2* auf 2q22.3. Weitere Replikationsstudien sind notwendig, um herauszufinden, ob diese

positionellen Kandidatengene zur Erblichkeit von häufigen GGE-Syndromen beitragen.

Zur Erforschung der „common disease/rare variant"-Hypothese untersuchten wir den Einfluss von sechs rekurrenten Mikrodeletionen auf das genetische Risiko von GGE. Diese treten in genomischen „Hotspot"-Regionen bei 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.11 und 22q11.2 auf und waren zuvor als seltene genetische Risikofaktoren in einer Vielzahl von neurologischen Erkrankungen erkannt worden. Die Untersuchung der rekurrenten Mikrodeletionen erfolgte primär mit hochauflösenden SNP-Mikroarrays in 1.497 europäischen GGE-Patienten, 5.374 Kontrollen und 566 GGE-Trios. Wenn alle sechs Mikrodeletions-Hotspots zusammen betrachtet wurden, zeigte sich eine signifikante Anhäufung dieser Mikrodeletionen in einer erweiterten Stichprobe von 2.563 GGE-Patienten im Vergleich zu 5.940 Kontrollen (P <2.20 × 10-16, OR = 7.65, 95% CI 4,59 bis 13,18). Einzeln betrachtet wurden signifikante Assoziationen mit GGE für die Mikrodeletionen auf 15q11.2 (P = 1.12×10^{-4}, OR = 3.59, 95% CI 1.80 – 7.25), 15q13.3 (P = 5.48×10^{-9}) und 16p13.11 (P = 4.42×10^{-6}, OR = 17.39, 95% CI 3.86 – 159.88) gefunden.

In einem Kandidaten-Gen-Ansatz haben wir getestet, ob Multi-Exon-Deletionen in den Genen *NRXN1* und *RBFOX1* zur Suszeptibilität von GGE beitragen. An beiden Loci (*NRXN1*: P = 0,0049; *RBFOX1*: P = 0,0083) fanden wir eine signifikante Assoziation entsprechender Mikrodeletionen mit GGE. Allerdings deuten die beobachtete hohe phänotypische Variabilität und unvollständige Penetranz und die daraus resultierende unvollkommene Segregation in den betroffenen Familien darauf hin, dass partielle Deletionen von *NRXN1* und *RBFOX1* eher als Suszeptibilitäts-Faktoren denn als hoch-penetrante Mutationen zu betrachten sind.

Die vorliegende Studie belegt, dass sowohl genomische Mikrodeletionen als auch häufige Suszeptibilitätsvarianten eine wichtige Rolle für die genetische Prädisposition von häufigen GGE-Syndromen spielen. So gelang es, die statistische Evidenz für Assoziationen von genetischen Varianten auf 1q43, 2p16.1 und 2q23.2 mit GGE-Syndromen zu erhöhen und einen neuen Suszeptibilitätslokus auf 8q12.2 zu identifizieren. Ferner zeigte sich, dass seltene Mikrodeletionen auf 15q11.2, 15q13.3 und 16p13.3 sowie an den Genloci *NRXN1* und *RBFOX1* in ihrer Gesamtheit erheblich zur genetischen Varianz von GGE beitragen.

List of figures

List of tables

List of abbreviations

Abbreviation	Term
aCGH	Array Comparative Genome Hybridization
ADHD	Attention Deficit/Hyperactivity Disorder
ASD	Autism spectrum disorder
BAC	Bacterial artificial chromosome
bp	Base pairs
CAE	Childhood absence epilepsy
CDCV	"Common disease, common variant" hypothesis
CDRV	"Common disease, rare variant" hypothesis
CI	Confidence Interval
CN	Copy Number
CNP	Copy Number Polymorphism
CNV	Copy Number Variation
cQC	Contrast Quality Control value
CR	Call Rate
df	Degree of freedom
DNA	Deoxyribonucleic acid
DZ	Dizygotic twins
EEG	Electroencephalography
EGTCS	Epilepsy with genetic tonic-clonic seizures
GTCS	Genetic tonic-clonic seizures
EM	Expectation-Maximization algorithm
EPI	Epilepsy
FoSTeS	Fork Stalling and Template Switching
FRET	Fluorescence resonance energy transfer
FS	Febrile seizure
GAE	Genetic absence epilepsy
GC	Genomic control
GEFS+	Genetic epilepsy with febrile seizures plus

Abbreviation	Term
GGE	Genetic generalized epilepsy
GWAS	Genome-wide association study
HMM	Hidden Markov Model
HWE	Hardy Weinberg Equilibrium
IBD	Identity-By-Decent
ID	Intellectual disability
ILAE	International League Against Epilepsy
IMBIE	Institute of Medical Biometry, Informatics and Epidemiology
JME	Juvenile myoclonic epilepsy
kb	Kilo-base pairs
KORA	Kooperative Gesundheitsforschung in der Region Augsburg
LCR	Low-copy repeats
LD	Linkage Disequilibrium
LMM	Linear Mixed Model
LOH	Loss-Of-Heterozygosity
MAF	Minor Allele Frequency
Mb	Mega-base pairs
MCMC	Markov chain Monte Carlo method
MDS	Multidimensional scaling
MZ	Monozygotic twins
NAHR	Non-allelic homologous recombination
NCBI	National Center for Biotechnology Information
ng	Nano gram
NHEJ	Non-homologous end-joining
OR	Odds Ratio
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PopGen	Populations Genetik Biobank
QC	Quality control
qPCR	Quantitative Polymerase Chain Reaction

Abbreviation	Term
QQ-plot	Quantile-Quantile-plot
RR	Relative risk or risk ratio
SD	Standard deviation
SHIP	Study of Health in Pomerania
SNP	Single Nucleotide Polymorphism
SZ	Schizophrenia
TDT	Transmission Disequilibrium Test
UCSC	University of California, Santa Cruz

1 Introduction

1.1 Introduction to Epilepsy

Epilepsy is a remediless disorder of the brain, defined by the occurrence of at least one unprovoked epileptic seizure. These seizures appear as symptoms due to abnormal excessive or synchronous neuronal activity in the brain (Fisher et al., 2005). Approximately 3% of the general population are affected by an epilepsy during lifetime, making epilepsy one of the most common neurological diseases (Hauser et al., 1993; Freitag et al., 2001). The clinical condition may be seriously troubling for the patient and for his social environment, and result in increased morbidity and premature mortality (Duncan et al., 2006). Up to 50% of epilepsy patients show a comorbidity of other cognitive, behavioral, psychiatric and neurologic disorders, and even sudden death (Tellez-Zenteno et al., 2007; Jensen, 2011; Kanner, 2013).

A genetic background of epilepsy is unquestioned and has been represented in twin and family studies (Berkovic et al., 1998; Callenbach et al., 1998; Bianchi et al., 2003; Johnson et al., 2003; Kjeldsen et al., 2003). Although 50% of epilepsies are genetically determined, monogenic epilepsies account only for 1-2% of all human epilepsies (Pandolfo, 2011). A number of causative genes have been identified for those forms with a clear family history and simple inheritance, many of them encoding for ion channels or other components of neuronal signaling (Meisler et al., 2001; Reid et al., 2009; Yalçın, 2012a). Despite these distinct findings, variable penetrance and expressivity could be shown for known epileptogenic mutations in animal models, depending on the genetic background (Bergren et al., 2005; Yu et al., 2006; Tan et al., 2008).

For sporadic epilepsies, a complex genetic component is obvious and a population-specific contribution most likely, although difficult to verify (Greenberg et al., 2000; Cavalleri et al., 2007; Guo et al., 2012). Nonetheless, recent genome-wide association studies (GWAS) were able to successfully identify susceptibility loci for epilepsy (EPICURE Consortium et al., 2012a; Guo et al., 2012). Computational simulations indicate that even small simultaneous changes in several ion channel genes may

have a severe effect, modeling a complex genetic background for common epilepsies (Thomas et al., 2009).

A still increasing number of pathogenic genes, copy number variants and gene regulatory elements have been discovered that are not directly or not at all connected with ion channels or neuronal signaling (Gurnett and Hedera, 2007). Therefore, the list of potential candidates for epilepsies is extraordinary large. Despite technical innovations and their remarkable impact on epilepsy research in the past 3 to 15 years (Rees, 2010; Poduri and Lowenstein, 2011a; Hildebrand et al., 2013; Vorstman and Ophoff, 2013), the genetic factors predisposing to epilepsy remain mostly elusive (Helbig and Lowenstein, 2013).

The identification of additional genes influencing the risk for epilepsies has large potential for clinical applications. Diagnostic and predictive testing based on genetic information may help patients to make informed decisions about managing their health care, despite a primary benefit to pharmacogenomics (Ottman et al., 2010).

1.2 Classification of genetic generalized epilepsy (GGE) syndromes

The classification of epileptic syndromes is an outstanding challenge that has been faced for almost 50 years (Gastaut, 1964, 1969; ILAE, 1981, 1985, 1989). Progressions in the understanding of epilepsies induced several revisions and updates to the traditional classification (Engel, 2001, 2006; Nordli, 2005; Capovilla et al., 2009), and have now outdated former concepts. In 2010, the International League Against Epilepsy (ILAE) Classification Commission introduced a new proposal for new terms and concepts to describe seizures and epilepsies (Berg et al., 2010). The terms idiopathic, symptomatic and cryptogenic have been replaced by the terms genetic, structural/metabolic and unknown (Berg et al., 2010). In contrast to the structural/metabolic (former symptomatic) epilepsies, idiopathic or genetic epilepsies occur in the absence of neurological deficits, intellectual disability, or brain lesions and have no known or suspected external cause, and genetic factors are implicated (Figure 1-1) (Berkovic et al., 2006; Helbig et al., 2008).

Generalized epilepsies are divided into subgroups by seizure type. Those groups are represented by seizures with tonic and/or clonic manifestations (tonic-clonic seizures, clonic seizures, and tonic seizures), absences (typical absences, atypical

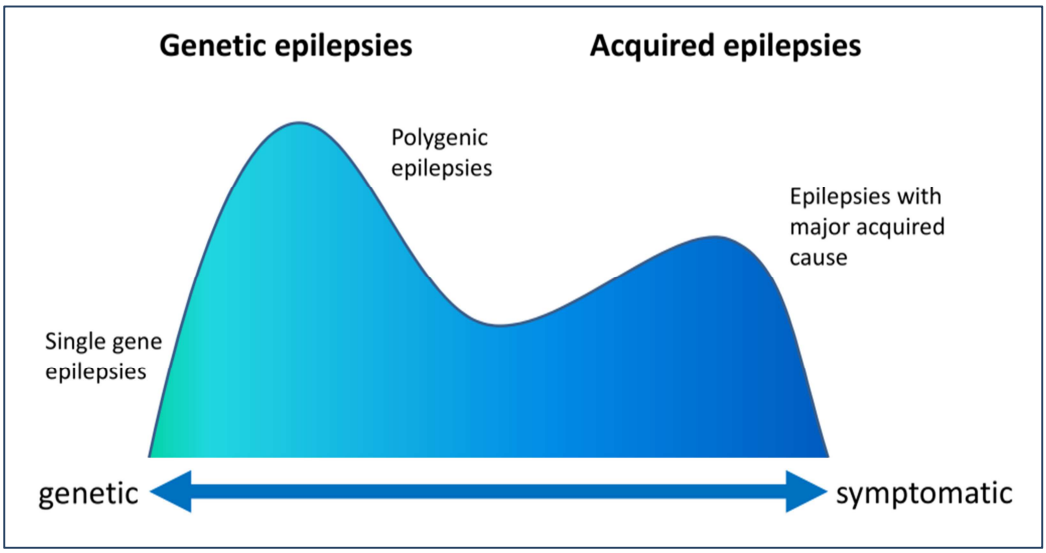


Figure 1-1 | Biological spectrum of epilepsy (Adapted from Berkovic et al., 2006)

absences, and myoclonic absences), myoclonic seizure types (myoclonic seizures, myoclonic astatic seizures, and eyelid myoclonia), epileptic spasms and atonic seizures (Classification from Engel, 2006. See Suppl. 6-1 for a more detailed overview, including non-generalized forms). These groups may be further divided by age of onset and additional cofactors. Four of these subtypes account for more than 90% of GGEs: Childhood absence epilepsy (CAE), juvenile absence epilepsy (JAE), juvenile myoclonic epilepsy (JME) and epilepsy with generalized tonic-clonic seizures (EGTCS) (Mullins et al., 2007). Those subgroups are often easy to distinguish by age of onset (CAE: 4-8 years, avg. 7.5 years; JAE: 9-13 years, avg. 13.3 years; JME: 12-18 years, avg. 14.6 years; EGTCS: 6-28 years, peak at 17.1 years), electroencephalography (EEG) patterns and by type and occurrence of seizures (Janz, 1997; Nordli, 2005). There has been evidence that those four generalized epilepsy syndromes are closely related to each other, showing similar EEG patterns and excitability in neuronal networks (Moeller et al., 2011; Zambrelli and Canevini, 2011), analogous to family studies suggesting a close genetic relation between CAE and JAE (Marini et al., 2004). Despite these findings and a high rate for remission of 65% until adolescence in CAE, 15-18% of all CAE patients develop JME, demonstrating some connection between the syndromes (Wirrell et al., 1996;

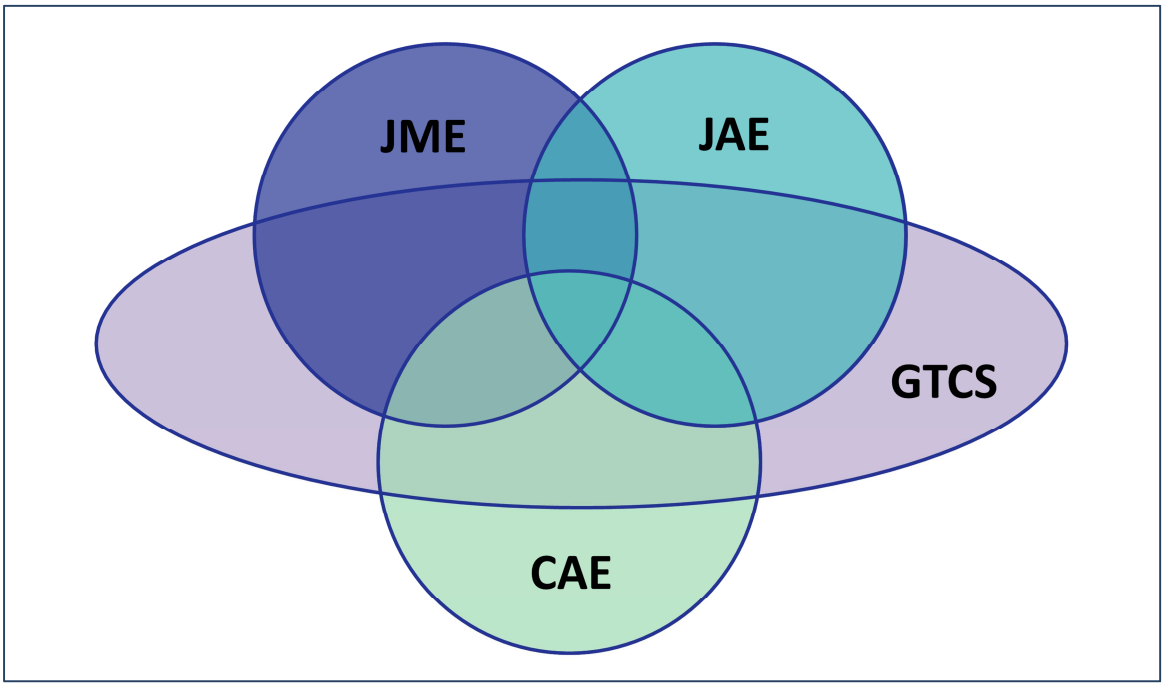


Figure 1-2 | Overlap of the complex phenotypes of most common genetic generalized epilepsies (modified from Janz et al., 1997; Crunelli and Leresche, 2002).

Martínez-Juárez et al., 2006). However, phenotypic heterogeneity can be observed within families, that are affected by differing seizure types (Kinirons et al., 2008). GTCS alone is present in about 50% of the cases, while about 70% of patients with other syndromes experience GTCS at least once in their lifetime (Janz, 1997; Mullins et al., 2007). Several linkage studies support the hypothesis, that interactions and different combinations of common disease loci for GGE are responsible for the spectrum of heterogeneous phenotypes (Figure 1-2) (Sander et al., 2000; Durner et al., 2001; Hempelmann et al., 2006; EPICURE Consortium et al., 2012b).

1.3 Genetic factors of GGE syndromes

The incidence of GGEs is about 20-30%, depending on the tested population (Berg et al., 1999; Freitag et al., 2001; Jallon et al., 2001). About 4-8% of GGE patients have a family history of epilepsy, with most affected relatives having GGE as well (Callenbach et al., 1998; Bianchi et al., 2003).

Furthermore, offspring of GGE patients is reported to be three times more likely affected than expected by population incidence rates, with a recurrence risk of 4-10% in siblings (Ottman et al., 1989; Bianchi et al., 2003). Several twin studies showed high concordance rates of generalized epilepsy phenotypes in monozygotic

twins (concordance rate of 0.65-0.82 in MZ and 0.12-0.26 in DZ), suggesting the strong influence of syndrome-specific susceptibility genes in the etiology of epilepsy (Berkovic et al., 1998; Kjeldsen et al., 2003). On the contrary, extensive family studies found high phenotype variability within pedigrees (Scheffer and Berkovic, 1997; Ottman et al., 1998). Concordance rates in families with myoclonic and/or absence seizure of 58-65% with seizure types led to the conclusion, that some genes confer susceptibility to only one specific GGE syndrome, while others contribute to a broader range of phenotypes (Winawer et al., 2002, 2005).

To date, the best characterized genes in GGE are *EFHC1* (myoclonin-1), *GABRA1* (gamma-aminobutyric acid (GABA) A receptor, alpha 1 subunit), *GABRG2* (gamma-aminobutyric acid (GABA) A receptor, gamma 2 subunit), *CACNA1A* (calcium channel, voltage-dependent, P/Q type, alpha 1A subunit), *CACNB4* (calcium channel, voltage-dependent, beta 4 subunit) and *SLC2A1* (solute carrier family 2 (facilitated glucose transporter), member 1) (Pandolfo, 2011).

EFHC1 encodes for a calcium binding protein and is supposed to be involved in regulating cell division and neuronal migration during cortical development (de Nijs et al., 2009; Léon et al., 2010) and *EFHC1* mutations were found co-segregating in unrelated families with JME (Suzuki et al., 2004). Disruption of *EFHC1* causes subtle malformations occurring in cortical and subcortical development associated with JME (de Nijs et al., 2009, 2012; Wong, 2010).

$GABA_A$ receptor function was expected to play a role in epileptogenesis for a long time, because of its key function as in inhibitory synaptic neurotransmission (Olsen et al., 1999). The first genetic evidence was a *GBARG2* mutation in a family with generalized epilepsy with febrile seizures (FS) (Baulac et al., 2001). Additional *GBARG2* mutations were later reported in patients with CAE (Wallace et al., 2001; Crunelli and Leresche, 2002), followed by a report of a *GABRA1* mutation in family members with JME (Cossette et al., 2002). Both mutations of *GABRA1* and *GBARG2* has been found in families with different phenotypes including CAE, CAE with FS and generalized epilepsy with febrile seizures plus (GEFS+) (Crunelli and Leresche, 2002; Lachance-Touchette et al., 2011). Other $GABA_A$ receptor subunit mutations in *GABRB3* and *GABRD* have been identified in diverging epilepsy phenotypes as well

(Macdonald et al., 2010), highlighting the strong impact of the $GABA_A$ receptor genes in epileptogenesis.

Genes of the voltage-dependent calcium channel complex proteins were identified in GGE syndromes. The effect of *CACNB4* mutations were described in mice and in families with heterogeneous GGE phenotypes, respectively CAE with JME and GTCS, JME with GTCS, and ataxia (Escayg et al., 2000). Similarly, mutations in *CACNA1A* were found in patients with CAE and ataxia (Jouvenceau et al., 2001).

Several other mutations in ion channel genes were described for GGE subtypes that are not included in this study, appreciable mutations in *KCNQ2* and *KCNQ3* in benign familial neonatal convulsions and *SCN1A*, *SCN2A*, *SCN1B* mutations in GEFS+ and Dravet syndrome (Claes et al., 2001; Gardiner, 2005).

An important non-ion channel gene to the etiology of GGE is *SLC2A1*, encoding the GLUT1 glucose transporter. Mutations altering gene function have been detected in patients with early-onset absence epilepsy (Suls et al., 2009). Again, *SLC2A1* shows a phenotypic spectrum besides GGE, including epileptic encephalopathy, GTCS as well as absence, myoclonic and atonic seizures associated with generalized spike-wave on EEG (Harkin et al., 2007; Mullen et al., 2011).

While the dogma of epilepsy in general as a channelopathy is put aside, newly discovered genes illustrate the importance of unstudied pathways in the etiology of epilepsy (Turnbull et al., 2005; Greenberg and Subaran, 2011; Poduri and Lowenstein, 2011b). As part of the GABA synthesis pathway, the genome-coded mitochondrial enzyme *ME2* is strongly suggested to be predisposing to GGE (Pal et al., 2003). Accordingly, association with the *BRD2* gene locus could be identified in patients with JME, although gene function is yet unknown (Pal et al., 2003). Other reports introduce other novel risk factors, for instance *ARX* to absence seizures and GTCS (Marsh et al., 2009), *STXBP1* to genetic subtypes of Ohtahara syndrome (Saitsu et al., 2008), *PNKP* to a previously unknown syndrome (microcephaly, early-onset, intractable seizures and developmental delay, MCSZ) (Shen et al., 2010) and *PCDH19* to epilepsy and mental retardation limited to females (Dibbens et al., 2008).

In most cases the relation between mutation and epilepsy phenotype is unclear, hindered by genetic heterogeneity with variable expressivity, emphasizing the

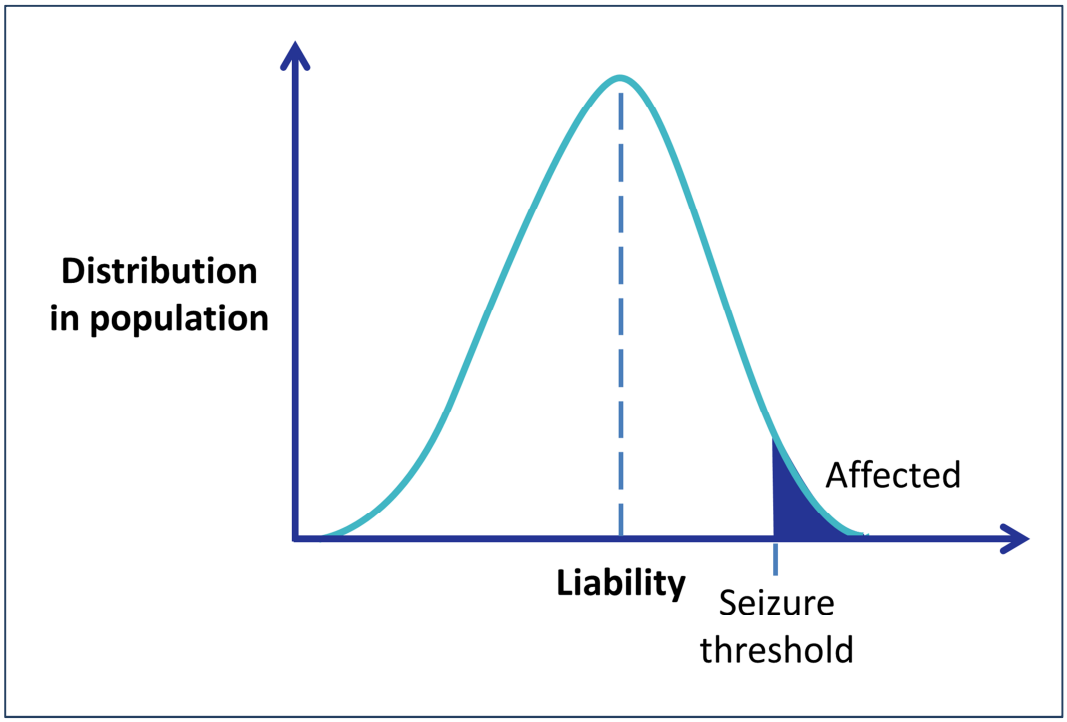


Figure 1-3 | Multifactorial-polygenic model of disease. Liability, an immeasurable quantitative trait, is normally distributed in the population. Individuals with liability above a threshold value are affected. With epilepsy, the threshold could be conceived of as a seizure threshold. Modified from Ottman et al., 2005.

importance of genetic background to specific epilepsy phenotypes (Reid et al., 2009). Hence most epilepsy patients have no affected relatives, a genetically complex etiology for the majority of all epilepsies is most likely, with multiple genes of moderate effect contributing to disease risk and clinical outcome (Ottman, 2005; Johnson, 2011). This assumption favors the model of a polygenic and multifactorial etiology of common epilepsy syndromes, with multiple genes and environmental factors involved (Figure 1-3).

Genome wide linkage and association studies have reported a number of different loci for GGEs. Early linkage mapping studies found evidence for susceptibility loci on chromosomes 2q36, 6p21, 14q23 and 20q13 (Sander, 1996; Sander et al., 2000). Association was found both in *CACNA1A* and *CACNA1H*, supporting the findings in Mendelian epilepsies (Chioza et al., 2001; Vitko et al., 2005). A recently published linkage study identified a novel risk loci at 5q34, close to several genes of the $GABA_A$ receptor (*GABRB2*, *GABRA6*, *GABRA1*, *GABRG2*) (EPICURE Consortium et al., 2012b), and consenting evidence of prior reported susceptibility loci at 2q34 (Ratnapriya et

al., 2010) and 13q31.3 (Tauer et al., 2005; Hempelmann et al., 2006). Replication in independent families/cohorts usually failed and no 'common epilepsy genes' could be identified (Mulley et al., 2005).

Association studies of candidate genes are based on the 'common disease, common variant' model (CDCV), assuming that common variations with moderate effect contribute to disease risk. The validity of this model was disputed after more than 50 studies without consisting results (Tan et al., 2004). In contrast, the 'common disease, rare variant' (CDRV) model, with rare variants with strong effects on disease outcome, was considered to be more fitting to complex epilepsy (Mulley et al., 2005). Besides, inadequate study designs might have contributed to the lack of success in epilepsy research (Kasperaviciūte et al., 2010; Tan and Berkovic, 2010; Heinzen et al., 2012). Nonetheless, the identification of indisputable genetic risk factors warrant further association studies (Pandolfo, 2011; Hildebrand et al., 2013). Association studies in Chinese population identified variants in *CACNA1H* for CAE (Liang et al., 2007) as well as *CAMSAP1L1* and *GRIK2* for general epilepsy (Guo et al., 2012). The to date largest GWAS in European population has found significant association in Linkage Disequilibrium (LD) blocks, including *VRK2* and *PNPO* as potent candidate genes, as well as genetic susceptibility in *SCN1A* for GGE in general, as well as susceptibility alleles in *ZEB2* for GAE, *CHRM3* for JME (EPICURE Consortium et al., 2012a).

In the same way, gene-gene interactions and modifiers are important factors for epileptogenesis. As an example, combinations of mild phenotype alleles can result in a severe clinical phenotype (Reid et al., 2009). Mouse models have demonstrated the effect of gene-gene interactions (Kearney et al., 2006), as well as the impact of protective and modifier variants (Kim et al., 2001; Glasscock et al., 2007; Martin et al., 2007c; Jorge et al., 2011).

1.4 Structural genomic variations in human disease

Although pathogenic structural genomic variations are known for quite some time (given the 15q11-13 deletion and duplication syndromes as a prominent example) (Butler et al., 1986; Bundey et al., 1994), investigation was often difficult and limited to large structural variations with severe effects. Due to technological advances the

understanding of this variations, especially duplications and deletions that are collectively called Copy Number Variations (CNVs), has made huge progress. First databases were established in 2004 (Iafrate et al., 2004), soon followed by a better understanding of the underlying genetic framework of CNVs (Sebat et al., 2004; Tuzun et al., 2005). Despite as assumed before, further mapping approaches showed that CNVs are not rare events, but are far more common than expected (Redon et al., 2006). CNVs contribute more to genomic diversity than any other genomic variation previously discovered (Redon et al., 2006; Korbel et al., 2007; Stranger et al., 2007).

Great efforts have been made to reveal the mechanisms underlying genomic rearrangements (Sebat et al., 2004; Tuzun et al., 2005), identifying three major principles that may cause CNVs (Gu et al., 2008). Regions with low-copy repeats (LCRs) can mediate non-allelic homologous recombination (NAHR) events. NAHRs mostly account for recurrent CNVs, which share the same breakpoint positions and sizes, resulting in almost identical events in multiple individuals. Non-recurrent CNVs, that may be stimulated by LCRs, do not share the same breakpoints or size, but still may show some overlap or redundant core regions in different samples. In rare cases, no-recurrent NAHRs may be caused as well by highly homologous repetitive sequences like Alu or LINE elements (Gu et al., 2008; Slavotinek, 2008).

Non-recurrent CNVs may as well be caused by non-homologous end-joining (NHEJ), which occurs during DNA double-strand break repair. NHEJ may lead to modification of broken DNA strands to allow for ligation of the breakpoints, that may add or delete several base pairs (Gu et al., 2008).

Other non-recurrent events may be mediated by replication Fork Stalling and Template Switching (FoSTeS) (Lee et al., 2007; Gu et al., 2008). During DNA replication, the replication fork may pause, the lagging strand disconnecting from the template and annealing to another replication fork due to microhomology. The replication would then continue at the new position, creating a duplication by switching to an upstream replication fork, or a deletion by switching to a downstream position. The orientation of the replication fork may cause inversions in addition any may occur several times in a row.

Large CNVs may encompass a high number of genes, without any apparent phenotype (Iafrate et al., 2004; Sebat et al., 2004; Redon et al., 2006). On the contrary, CNVs may affect gene functions in different ways (Cook and Scherer, 2008), commonly considered as changes in gene dosage or expression (Gu and Lupski, 2009; Itsara et al., 2009). Pathogenic CNVs that are too small to be detected with conventional cytogenic methods, but are above level of InDel-variants, are called microdeletions or microduplications vice versa. The simplest scenario is the deletion or disruption of one or more genes, which result in a functional loss. A microdeletion may lead to haploinsufficiency, unmasking pathogenic recessive mutations on the remaining homologues segment that have been silent in previous generations. Furthermore, regulatory elements may be disrupted or deleted, fusion genes can be generated at CNV breakpoints and lead to a gain-of-function mutation, or positional effects may influence gene expression of more distant genes unaffected by any CNV (Lupski and Stankiewicz, 2005; Feuk et al., 2006). Even a "two-hit" model of additional pathogenic CNVs contributing to a phenotype is conceivable (Girirajan et al., 2010, 2012). Pathogenic genomic duplications appear to be less common and seem to develop less severe phenotypes than microdeletions, although they might be under-diagnosed so far (Cook and Scherer, 2008; Slavotinek, 2008). Since only a limited number of rare CNVs have been successfully associated with disease, it is most likely that considerable amount of rare CNVs account for the 'common disease-rare variant' hypothesis (Itsara et al., 2009).

1.5 Genomic microdeletions in genetic generalized epilepsy

The first pathogenic CNV locus that was associated with GGE was a recurrent microdeletion at 15q13.3, which was found in 12 of 1,223 and none of 3,699 controls (Helbig et al., 2009). Despite being the first CNV associated with GGE, this microdeletion is still the most prevalent risk factor for common epilepsies. The reported 15q13.3 microdeletion has been reported before in association with mental retardation and seizures (Sharp et al., 2008), and schizophrenia (Schizophrenia Consortium, 2008; Stefansson et al., 2008). Following studies emphasized the pathogenicity of the reported microdeletion (Dibbens et al., 2009) and highlighted its phenotypic variability (Miller et al., 2009; Pagnamenta et al., 2009).

Association with GGE was subsequently detected with microdeletions at 15q11.2 and 16p13.1 (de Kovel et al., 2010), both microdeletions previously reported in neuropsychiatric disorders; microdeletions at 15q11.2 associated with schizophrenia (Stefansson et al., 2008; Kirov et al., 2009; Need et al., 2009), microdeletions at 16p13.11 associated with autism spectrum disorder (Sebat et al., 2007; Ullmann et al., 2007), mental retardation (Ullmann et al., 2007; Hannes et al., 2009), and with schizophrenia (Need et al., 2009). In addition, microdeletions at 1q21.1 previously reported in schizophrenia and mental retardation (Brunetti-Pierri et al., 2008; Mefford et al., 2008; Schizophrenia Consortium, 2008; Stefansson et al., 2008; Need et al., 2009), microdeletions at 16p11.2 identified in autism and mental retardation (Sebat et al., 2007; Kumar et al., 2008; Marshall et al., 2008; Weiss et al., 2008), and microdeletions at 22q11.1 associated with schizophrenia, mental retardation and autism (Bassett et al., 2008; Schizophrenia Consortium, 2008; Kirov et al., 2009; Need et al., 2009), showed some suggestive evidence, but no association with GGE. Altogether, CNVs associated with GGE collectively explain a larger portion of the genetic variance epilepsy syndromes than any single gene (de Kovel et al., 2010; Poduri and Lowenstein, 2011b).

Recurrent microdeletions associated with genomic disorders share similar mechanisms and structures and are mostly mediated by NAHR (Mefford and Eichler, 2009), although they show a high variability in phenotypes (Mefford et al., 2010). Therefore other risk factors, environmental and genetic modifiers are likely to contribute to the outcome of disease (Mefford and Eichler, 2009; Mefford and Mulley, 2010).

In addition, smaller structural aberrations have been reported that only affect single genes or only single exons. Deletions in *NRXN1* have been associated with a broad spectrum of neurodevelopmental and psychiatric disorders including schizophrenia, autism spectrum disorder, and intellectual disability (Kirov et al., 2009; Rujescu et al., 2009; Ching et al., 2010). In some cases, comorbidity of epilepsy was reported (Ching et al., 2010; Gregor et al., 2011) and compound *NRXN1* mutations combining heterozygous exonic microdeletions and nonsense or splice-site mutations have been described in individuals with severe early onset epilepsy and mental retardation (Harrison et al., 2011; Duong et al., 2012). Neurexins are neuronal

adhesion molecules, required for synaptic contacts and efficient neurotransmission in the brain. They are located in the presynaptic terminal where they interact with postsynaptic neuroligins to form a transsynaptic complex (Südhof, 2008).

Structural variations disrupting the gene encoding the neuronal splicing regulator *RBFOX1* have been reported in three patients exhibiting epilepsy in comorbidity with autism, intellectual disability, or pontocerebellar hypoplasia (Bhalla et al., 2004; Martin et al., 2007a; Gallant et al., 2011). Additionally, a linkage locus for photoparoxysmal response in GGE families has been mapped to the genomic region of *RBFOX1* at 16p13.3 (Pinto et al., 2005). *RBFOX1* is a splicing regulator gene, involved in the splicing of many neuronal transcripts, binding the sequence (U)GCAUG in introns flanking alternative exons (Jin et al., 2003; Auweter et al., 2006; Voineagu et al., 2011; Fogel et al., 2012). It is crucial for regulation of neuronal excitation and has a notably impact on susceptibility of epilepsy (Gehman et al., 2011; Voineagu et al., 2011), as several target transcripts of *RBFOX1* (for example, *DCX*, *GABRG2*, *GAD2*, *GRIN1*, *KCNQ2*, *SCN8A*, *SLC12A5*, *SNAP25*, *SV2B*, and *SYN1*) are correlated with epileptogenesis (Barnby et al., 2005; Corradini et al., 2009; Papale et al., 2009; Pandolfo, 2011; Fogel et al., 2012; Veeramah et al., 2012). Brain-specific homozygous and heterozygous *Rbfox1* knockouts in mice do not alter brain morphology, but show differentially spliced RNA transcripts and display spontaneous seizures and a dramatic epileptogenic response to kainic acid resulting in status epilepticus (Gehman et al., 2011). Consequently, RNA interference–mediated 50% knockdown of *RBFOX1* transcripts in human neurons changes the alternative splicing pattern and expression of primarily neuronal genes involved in synapse formation and function (Voineagu et al., 2011; Fogel et al., 2012).

1.6 Objectives

This study was aimed at the identification of novel genomic loci associated with common GGE syndromes. To achieve this goal, two strategies for identifying genetic factors underlying complex disease were pursued:

i) GWAS meta-analyses of two primary GWAS GGE data sets comprising a GGE case-control cohort and a family-based study group. In addition to the global analysis, phenotypically distinct GGE subgroups, such as genetic absence epilepsies (GAE,

comprising CAE and JAE) and juvenile myoclonic epilepsies (JME), were to be analyzed separately to distinguish genetic factors, which may differentially predispose to particular GGE subtypes.

ii) candidate variant/gene testing to investigate the impact of structural microdeletions on the etiology of common genetic epilepsies. Large recurrent microdeletions previously reported in neuropsychiatric disease were to be analyzed to detect association with GGE. Additionally, two potential candidate genes for GGE (*NRXN1*, *RBFO1*) were to be screened for exonic microdeletions to identify novel risk factors for epileptogenesis.

2 Methods and Materials

2.1 Study Participants

Epilepsy patients of European ancestry with common GGE syndromes (GAE with age-at-onset 3-20 years, JME with age-at-onset 6-20 years and EGTCS alone with age-at-onset 8-30 years) were recruited in a multi-center effort from the European EPICURE Project (http://www.epicureproject.eu). Phenotyping and diagnostic classification of GGE syndromes were carried out according to EPICURE guidelines and standardized phenotyping protocols (http://portal.ccg.uni-koeln.de/ccg/research/epilepsy-genetics/sampling-procedure/) (ILAE, 1989; Nordli, 2005; Berg et al., 2010). All study participants gave informed consent according to the regulations at their local institutional review boards. Individuals with a history of major psychiatric disorders (autism spectrum disorder, schizophrenia and affective disorder) or severe intellectual disability were excluded.

Two independent, unrelated sample cohorts were available for analysis. First, a case-control study cohort consisting of 1,569 unrelated GGE cases of European ancestry (600 males, 969 females; Austria, n = 197; Belgium, n = 53; Denmark, n = 95; Germany, n = 933; and The Netherlands, n = 291) and 6,201 German controls (3,142 males, 3,059 females). The epilepsy patients comprised the following GGE syndromes: 693 GAE, 625 JME, and 251 patients with EGTCS alone. The control cohort of 6,201 unscreened German controls were collected from Southern (PopGen, n = 1,625) (Wichmann et al., 2005), Northern (KORA, n = 1,163) (Krawczak et al., 2006), and North-Eastern (SHIP, n = 3,413) (Völzke et al., 2011) regions of Germany.

In addition, 566 parent-offspring trios of European origin with children affected by GGE were available for family-based association analysis (216 male trio children, 350 females). Trios were recruited from Australia (n = 98), Austria (n = 3), Bulgaria (n=4), Denmark (n = 20), Finland (n=1), France (n=16), Germany (n = 15), Italy (n = 272), Spain (n=4), and Turkey (n = 133). The trio sample cohort contained following syndromes: 317 GAE, 157 JME, and 92 trios with EGTCS alone.

2.2 Array-based SNP genotyping

Microarrays for SNP genotyping usually have a similar design. An acryl or glass surface is carrying a large number of unique oligomer probes that are specific to a defined position on the genome, for example a SNP or a copy number region. Large groups of unique probes are forming so called features on the array surface. Sample DNA is amplified, fragmented and hybridized to complementary probes on the array surface. DNA fragments are stained with fluorescent nucleotides and/or with fluorescent antibodies. A laser based imaging system is used to collect fluorescent signals. Genotypes can be distinguished by the position of the signal, corresponding to the position of a probe feature, by signal intensity and by color if a two colored assay is used. Imaging data can now be analyzed by the software provided by the array manufacturer and is transformed into genotype, copy number or other applicable data.

2.2.1 DNA preparation and quality control

All DNA samples provided by the participating centers were quantified using the NanoDrop 1000 photometer (NanoDrop products, Wilmington, DE, USA) and the presence of high-molecular weight genomic DNA was examined by gel electrophoresis on 1%-agarose gels. DNA samples exhibiting a high degree of degradation (no distinct fragment band > 10 kb) were excluded from the experimental investigations.

2.2.2 SNP genotyping using the Affymetrix Genome-Wide SNP Array 6.0

The Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA) offers oligonucleotide probe features for 1.8 million specific genetic markers. Of those, more than 906,600 account for SNPs and 946,000 for non-polymorphic probes to detect copy number variations (CNVs). A total of 500 ng high quality genomic DNA is split in two aliquots and digested in parallel with two restriction enzymes (Nsp and Sty). The fragments are ligated to adaptors that recognize the cohesive 4 bp restriction overhangs. A generic primer that recognizes the adaptor sequence is used to amplify DNA fragments by Polymerase Chain Reaction (PCR) (Kennedy et al., 2003). PCR conditions have been optimized to preferentially amplify fragments in the 200 to 1,100 bp size range. PCR amplification products for each

restriction enzyme digest are combined and purified using polystyrene beads. The amplified DNA is then fragmented, labeled, and hybridized to a SNP6.0 microarray (Matsuzaki et al., 2004). After ligation, the arrays are stained and imaged.

Raw data analysis and genotyping is performed by the Affymetrix Genotyping Console software (Affymetrix, Santa Clara, CA, USA). Raw data analysis is required to assure sufficient quality of the processed arrays, by evaluation of probe set signal intensity. Signal intensity is crucial for correct genotype calling. On the Affymetrix SNP6.0 array, the contrast QC value (cQC) represents the distribution of signal intensities for subset of representative SNPs. The same accounts for Nsp and Sty specific values, cQC-Nsp and cQC-Sty. Samples with any cQC value ≤ 0.4 have been excluded due to expected low raw signal quality. Genotyping is performed after filtering for quality of raw signal intensities.

SNP genotyping is performed using the Birdseed algorithm (Korn et al., 2008), integrated in the Affymetrix Genotyping Console. Birdseed uses a customized (EM) algorithm to fit two-dimensional Gaussians to SNP data, producing genotypes and confidence scores for every individual at every SNP.

2.2.3 SNP genotyping using the Affymetrix Axiom® Genome-Wide SNP array

The Affymetrix Axiom® Genome-Wide HU genotyping array (Affymetrix, Santa Clara, CA, USA) is a two color, ligation-based assay utilizing 30-mer oligonucleotide probes synthesized in situ on a microarray substrate, with automated, parallel processing of 96 samples per plate (Hoffmann et al., 2011). The Axiom® Genome-Wide Reagent v1 array used in this study is carrying specific probe features for 567,096 SNPs. An amount of 200 ng high quality DNA is amplified by Whole Genome Amplification and digested by restriction Enzyme into fragments of 25-125 bp in size. DNA fragments form a probe-target complex on the array surface and are hybridized with SNP-site specific probes and ligated for specificity. Arrays are washed, stained and imaged after ligation.

Raw data analysis is performed by the Affymetrix Genotyping Console software, similar to the Affymetrix SNP6.0 array. For evaluation of signal intensities of the Axiom® Genome-Wide array, the corresponding value is called DishQC. Manufacturer's threshold for sample exclusion is specified at ≤ 0.82, although the

given threshold is not definitive. For the Axiom® array, samples with a marginal DishQC values have been genotyped and excluded on account of the Call Rate (CR) per sample.

Genotyping is performed after filtering for quality of raw signal intensities by the Affymetrix Genotyping Console software, using the Birdseed algorithm (Korn et al., 2008).

2.3 Genome-wide association studies

2.3.1 Array and SNP and quality control

SNP microarray data required extensive quality control (QC) to ensure high data quality and to avoid formation of artifacts during imputing. Only samples with sufficient signal intensity ratio were used for further QC filtering (see 2.2.2, 2.2.3). Samples with CR < 95%, and excessive heterozygosity rate of autosomal SNPs > 29.5%, were excluded from analysis. Unsuccessful gender assignment led to exclusion of the sample as contamination or bad sample quality was expected. Remaining samples were genotyped again to improve cluster distribution of the high quality samples.

Within the family-based data set, trio pedigree structure and relationship of all trio members was checked with PedigreeExplorer (Steffens 2007; Institute for Medical Biometry, Informatics and Epidemiology (IMBIE) of the "Medizinische Einrichtungen der Universität Bonn", Bonn, Germany; http://pedigreeexplorer.meb.uni-bonn.de/) and by IBD-estimation in PLINK (Purcell et al., 2007; http://pngu.mgh.harvard.edu/purcell/plink/). Based on this information, pedigree structure and sample permutations were corrected if possible or pedigrees excluded if necessary.

SNPs with high genotyping accuracy were selected for SNP imputation according to the following QC criteria, with respect to the sample cohort and array type. Affymetrix SNP6.0 QC criteria: i) minor allele frequency (MAF) < 5% in cases or controls, ii) CR < 98% for SNPs with MAF > 10% and CR < 99% for SNPs with MAF < 10% in either cases or controls, ii) difference of missing data >1% between the cases and controls, iv) deviation from the Hardy–Weinberg equilibrium (HWE) with P <

1.0×10^{-4} in the controls. Affymetrix Axiom QC criteria: i) MAF < 5%, ii) CR < 95%, iii) HWE with $P < 1.0 \times 10^{-6}$, Mendelian error rate > 5% with remaining Mendelian errors set to missing.

2.3.2 Principal component analysis

A study cohort with samples from different populations may be prone to effects by reason of differing SNP allele frequencies in those populations, called population stratification. Population stratification may occur due to geographical and cultural barriers that lead to discrete mutation, selection and genetic drift in separated populations. Thereby the founding of subpopulations is likely. Subpopulations can often be distinguished as ethnic groups, but further stratification may be hidden beneath an obvious superstructure. Those groups may differ crucial in their allele frequencies for genetic variations. As association analysis statistics are based on the comparison of differing allele frequencies, distinct origins in samples and especially between cases and controls may lead to false-positive results (Thomas and Witte, 2002). In this case, association would display the difference between populations and no association with disease.

Existing genome-wide SNP genotypes can be used to detect and correct for population stratification. Established methods are principal component analysis (PCA) and multidimensional scaling (MDS). As both methods yield comparable results (Wang et al., 2009) and genetic and geographical distances are well correlated in European populations, we decided to perform a PCA analysis with the EIGENSTRAT v3.0 software (Patterson et al., 2006; Price et al., 2006). The principle used by both methods is the summarization of available data (i.e. genotype information of a sample cohort) in principal components (PCA) or dimensions (MDS). Those are sorted by their amount of variance and correlated. The first dimensions are carrying the mayor information about the sample, with later dimensions tending towards 0. To assure that true population effects are explored, a high quality data set (SNP quality filters: CR > 0.99, MAF > 0.15, HWE with $P > 10^{-3}$) is created with exclusion of markers in tight Linkage Disequilibrium (LD; $r^2 > 0.1$), markers that are located in known regions of long range LD patterns (Suppl. 6-2) and common inversion polymorphisms.

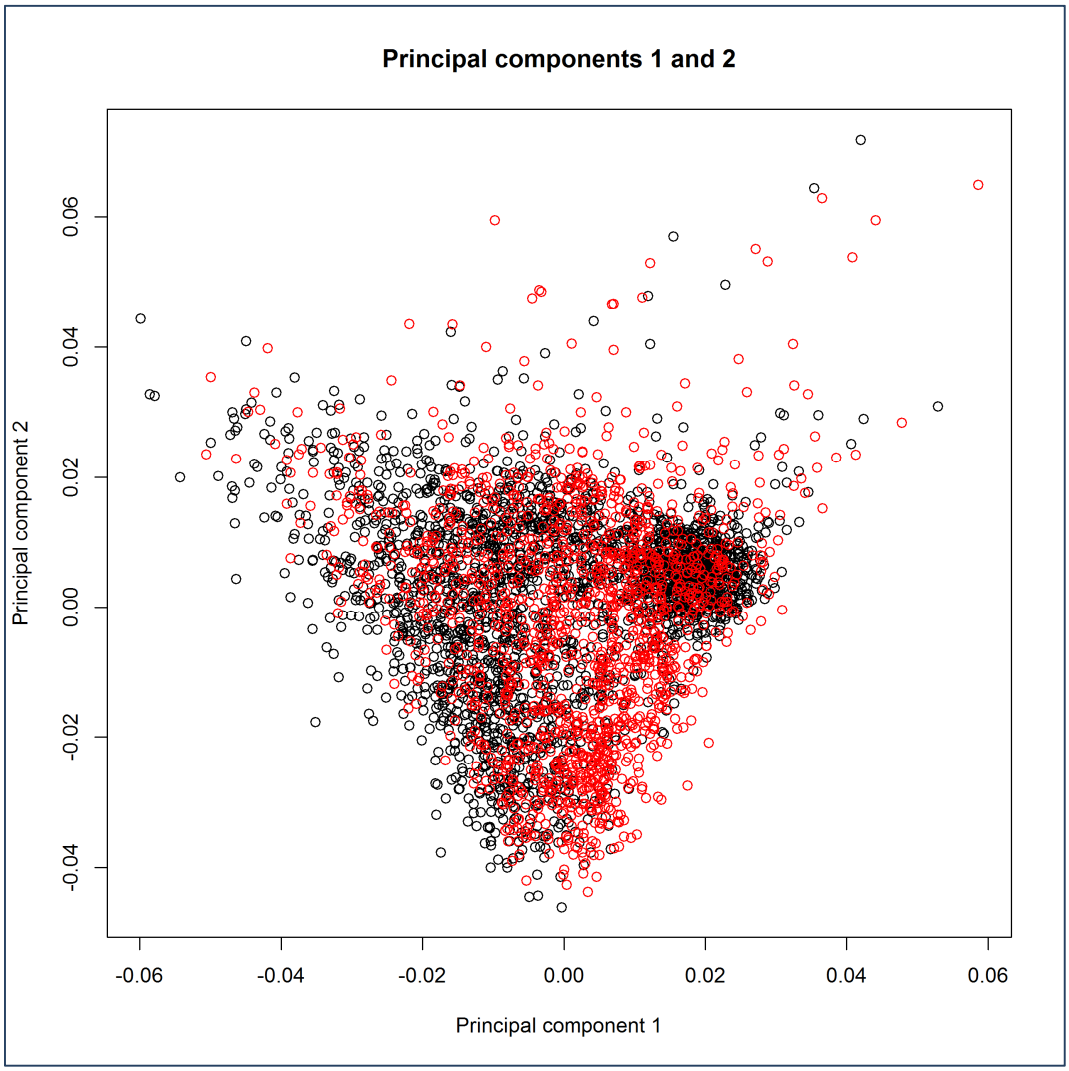


Figure 2-1 | Plot of principal components 1 and 2 from PCA. Exemplary data plotted of principal components with largest effect in PCA. Cases in red, controls in black.

2.3.3 SNP imputing

Imputing is a useful multipoint method that can detect causal variants that have not been directly genotyped by, for example, SNP-based DNA microarrays (Marchini et al., 2007). This approach improves statistical power of association analysis by predicting unobserved SNP genotypes using existing catalogs of variation and known haplotypes, such as the HapMap project (Altshuler et al., 2010) and the 1000genomes project (Abecasis et al., 2012), as a reference panel. Equally important, imputation is an indispensable step for later meta-analysis, if studies based on different genotyping methods or arrays are combined. Different genotyping arrays only have a fraction of their SNPs overlapping in general. The reduction of a joined data set is avoided by imputing to a much larger marker set with a considerably increased marker overlap.

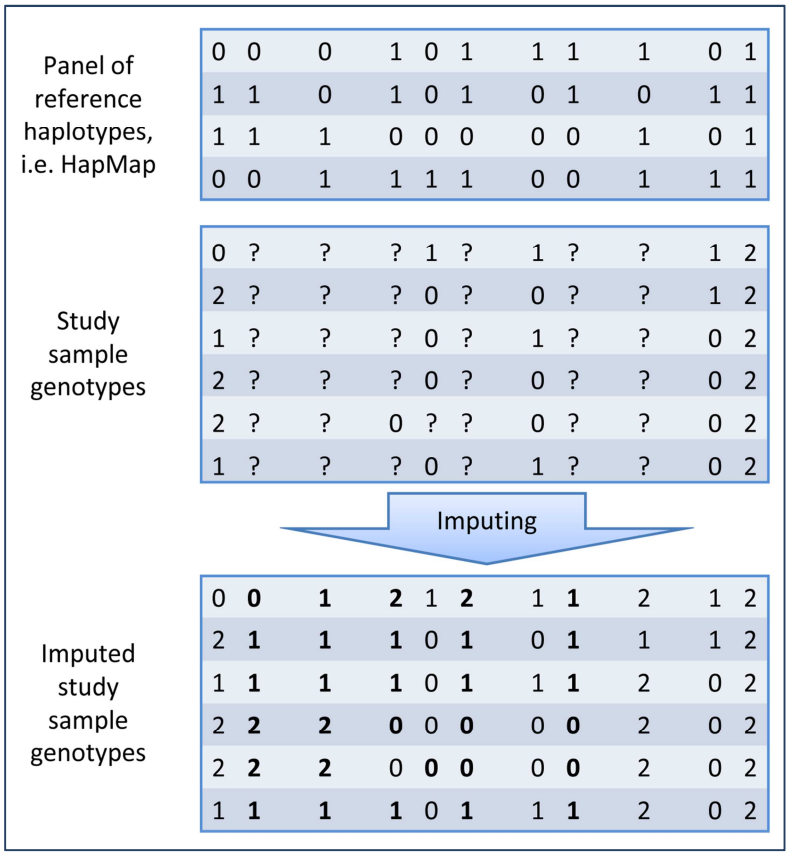


Figure 2-2 | Simplified procedure of imputing (modified from Marchini et al., 2007, https://mathgen.stats.ox.ac.uk/impute/impute_v1.html)

Imputation in this study was accomplished using IMPUTE v2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) (Howie et al., 2009), a method that uses a Markov chain Monte Carlo (MCMC) framework to predict untyped SNPs by combining a reference panel of individuals genotyped at a dense set of SNPs with a study sample collected from a genetically similar population and genotyped at a subset of these sites. The approach is based on an approximate population genetics model that gives more weight to genotypes that are consistent with the local patterns of LD. This approach uses information from all markers in LD with an untyped SNP, decreasing with genetic distance from the SNP being imputed (Marchini et al., 2007). The SNPTEST v2 software was used to calculate frequentist and bayesian tests used for quality filtering of SNP markers (Marchini et al., 2007). SNP markers were filtered for minor allele frequency, proportion of missing genotypes, and info score. The info score is a measure of the observed statistical information for the estimate of allele frequency of a SNP. All individuals in the sample are used to test each SNP with a maximum value of 1 for perfect information.

Quality control of imputed data

Imputed SNP genotypes were quality filtered again to maintain a high quality data set. Quality filtering of imputed data done with IMPUTE dosage data using SNPTEST v2 (Marchini et al., 2007). SNPs were excluded from further analysis according the following QC criteria: i) SNPtest info quality value < 0.9, ii) MAF < 3% (for cases and controls separately in the case-control cohort), iii) missing data proportion > 1%, iv) HWE deviation with $P < 10^{-6}$ (in the case-control cohort, with $P < 10^{-6}$ in the cases and $P < 10^{-4}$ in the controls).

Imputing of Affymetrix SNP 6.0 array case-control data

After SNP genotyping, the quality filtered data set consisted of 1523 cases and 2454 controls (579 males, 944 females; cases with 705 GAE, 579 with JME, and 239 with EGCTS alone) with 572,071 SNP markers. The data set was imputed using IMPUTE v2, based on the reference panel: 1000 Genomes Phase I (interim) release in NCBI build 37 (hg19) coordinate (update 05 Mar 2012) by Dr. Carolien deKovel at the Department of Medical Genetics, Section Complex Genetics, University Medical Center Utrecht, The Netherlands. Imputed data was transferred to Cologne for further QC and association analysis. The final data set consisted of 5, 564,553 SNPs after post-imputing QC.

Imputing of Affymetrix Axiom® SNP array trio data

The quality filtered data set consisted of 566 trios and 553,012 SNP markers after SNP genotyping. The data set was imputed using IMPUTE v2, based on the reference panel 1000 Genomes Phase I (interim) release in NCBI build 37 (hg19) coordinate (update 19 Apr 2012) by Dr. Markus Leber at the IMBIE of the "Medizinische Einrichtungen der Universität Bonn", Bonn, Germany. Imputed data was transferred to Cologne for further QC and analysis. The final data set consisted of 5,776,102 SNPs after post-imputing QC.

Data conversion for TDT analysis

Data output format of IMPUTE is so called dosage data, presenting frequencies of AA, AB and BB genotypes for every single SNP per sample. For some analysis data had to be converted into Linkage data format, that features genotype base calls for every SNP. Data conversion was done with a custom Perl script (by Dr. Markus Leber, IMBIE of the "Medizinische Einrichtungen der Universität Bonn", Bonn, Germany), using the best genotype method. Dosage data alleles are converted into Linkage format, if the probability of the most likely genotype is at least 40% higher than that of the second likely genotype. Otherwise, genotype is set to missing.

2.3.4 Case-control-based genome-wide association analysis

The case-control data set in our study is equivalent to the Stage-1 data cohort of previously published GWAS (EPICURE Consortium et al., 2012a). Stage-1 data of this publication was selected for a meta-analysis with the TDT trio data surveyed by our study. An update in the imputing reference set required rerun of imputation procedures, and recalculation of statistical analysis for actuality and full compatibility with the trio data.

While a high genomic inflation factor (GGE: $\lambda = 1.11$, GAE: $\lambda = 1.05$ and JME: $\lambda = 1.06$) was reported from prior GWAS for association results of logistic regression analysis of our data (EPICURE Consortium et al., 2012a), genome-wide association was calculated using FaST-LMM (http://fastlmm.codeplex.com/) (Lippert et al., 2011), a method using a factored spectrally transformed linear mixed-model that explicitly captures all sources of structure based on estimates of the genetic relatedness of individuals. Logistic regression for an additive model was performed by SNPTEST v2 to estimate Odds Ratios (OR) and confidence interval, using the first five principle components from PCA analysis and gender status as covariates. The signal intensity cluster plots of the most significantly associated SNPs ($P < 1.0 \times 10^{-5}$) were manually examined within the Affymetrix Genotyping Console. In total, 100 SNPs exhibiting artificial cluster plots were excluded.

2.3.5 Family-based genome-wide association analysis

SNP genotype dosage data were converted to best-genotype estimates due to the following conditional likelihoods of genotypes into linkage format with Perl. Data was quality filtered again, with CR ≥ 98% and a Mendelian error rate < 4% with remaining Mendelian errors set to missing. The final data set consisted of 566 complete parent-offspring trios and 5,089,023 SNPs (216 male trio children, 350 females; 317 GAE, 157 JME, and 92 trios with EGTCS alone); one trio was excluded due to increased rate of Mendelian errors.

Genome-wide association was calculated by using the Transmission-Disequilibrium Test (TDT) in PLINK (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell et al., 2007). This method is using data from families with at least one affected child to evaluate the transmission of a marker allele from a parent to an affected offspring (Spielman et al., 1993). The TDT considers parents who are heterozygous for an allele associated with disease and evaluates the frequency with which that allele or its alternate is transmitted to affected offspring. Test statistic is calculated by a variant of the χ^2-Test, the McNemar's test (McNemar, 1947).

In contrast to classic χ^2-test based approaches to test for association, the TDT remains valid and preserves the Type-I error rate, regardless of population history (Ewens and Spielman, 1995). However, population stratification will lead to a decrease in the number of heterozygous parents and will change the relative proportion of informative markers (Sebro and Rogus, 2010). This leads to a loss of statistical power by increasing the required sample sizes to achieve level of significance. TDT analysis was done in PLINK (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell et al., 2007). The signal intensity cluster plots of the most significantly associated SNPs ($P < 1 \times 10^{-5}$) were manually examined within the Affymetrix Genotyping Console. In total, 24 SNPs exhibiting artificial cluster plots were excluded.

2.3.6 Meta-analysis of case-control and family-based association studies

A well-established tool to perform fast and reliable meta-analysis is METAL (http://www.sph.umich.edu/csg/abecasis/metal/index.html) (Willer et al., 2010). A key factor for meta-analysis is assigning appropriate weights to statistic results for

association for each combined study to indicate the effect size of the individual P-values. The basic principle of meta-analysis is to combine the evidence for association from individual studies, using appropriate weights. METAL implements two different approaches for weighted analysis. For the first approach, the p-value and the direction of effect (in our study log(OR)) are converted into a Z-Score, with a negative Z-Score for large p-values and positive Z-Scores for small p-values. Z-scores for each allele are combined across samples in a weighted sum proportional to the square-root of the sample size for each study. The second approach weights the effect size estimates by their estimated standard errors. This second approach requires effect size estimates and their standard errors to be in consistent units across studies. Asymptotically, the two approaches are equivalent when the trait distribution is identical across samples (such that standard errors are a predictable function of sample size). Odds Ratios from logistic regression from case-control analysis and OR from trio TDT analysis were pooled using GWAMA (http://www.well.ox.ac.uk/gwama/index.shtml) (Mägi and Morris, 2010).

Sample sizes and number of overlapping SNPs for genome-wide association meta-analysis were as following: GGE 1523 cases, 2454 controls, 566 GGE trios, 4,224,196 overlapping SNPs; GGE subtype analyses: GAE: 705 cases and 2454 controls, 317 trios, 4,224,597 overlapping SNPs; JME: 579 cases, 2454 controls, 157 trios, 4,220,447 overlapping SNPs.

2.3.7 Relative risk and odds ratio

The relative risk is an effect size for comparison of two groups in a given cohort that are exposed to a certain trait. The number of individuals developing an effect is measured over time (Viera, 2008).

Table 2-1 | Calculating risk ratio and odds ratio in a cohort study

	Develop outcome	**Do not develop outcome**
Exposed	a	b
Not exposed	c	d

Modified from Viera et al., 2008

The relative risk or risk ratio (RR) can be calculated as follows:

$$RR = \frac{incidence\ in\ exposed}{incidence\ in\ not\ exposed} = \frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)}$$

The same table can be used to calculate an odds ratio (OR). The OR represents the ratio between the probability (or odds) of the two groups to develop a certain trait:

$$OR = \frac{\begin{array}{c}odds\ that\ an\ exposed\ \ person\\ develops\ the\ outcome\end{array}}{\begin{array}{c}odds\ that\ an\ unexposed\ \ person\\ develops\ the\ outcome\end{array}} = \frac{\left(\frac{a}{b}\right)}{\left(\frac{c}{d}\right)} = \frac{a\,d}{b\,c}$$

RR and OR estimate the rate of risk or probability to develop a specific outcome. For RR = 3, the risk to develop the outcome is 3 times higher. For an OR = 3, the same accounts for the odds to develop the outcome. To estimate the statistical significance and accuracy of RR and OR calculation, it is recommended to calculate a confidence interval (CI, usually 95%). The CI is represented in a spectrum of values. In 95 out of 100 observations, the OR would be between the upper and the lower CI values. ORs and CIs have been calculated using the R software environment (http://www.r-project.org/).

2.3.8 Genomic control

The genomic control (GC) is an analytic method used to control the false positive rate in population-based association studies (Devlin and Roeder, 1999; Devlin et al., 2004). To reduce the effect of stratification by experimental design and analysis, cases and controls should be matched for ethnicity and environmental covariates. GC then adjusts for the residual effects of stratification. A careful study design and implementation pay off in statistical power, though even small stratification can inflate association test statistics. The GC further eludes the need for Bonferroni correction for multiple testing, leading to better performance in many settings while still decreasing the risk for false positives. To apply GC to an association study, an inflation factor λ is calculated for the χ^2 distribution under the hypothesis of no population stratification:

$$\lambda = \frac{median\ (\chi_1^2, \chi_2^2, \chi_3^2, \ldots, \chi_n^2)}{0.456}$$

The association statistics of single-marker P-values are therefore corrected by:

$$\chi_{GC}^2 = \lambda * \chi^2$$

2.3.9 Power calculations

Power calculations for GWAS were performed using CaTS power calculator (http://www.sph.umich.edu/csg/abecasis/CaTS/) (Sebro and Rogus, 2010). Case-control GWAS for GGE has a power of 80% to detect a variant associated with disease on OR ≥ 1.40, with significance threshold of $P \leq 5 \times 10^{-8}$, assuming a prevalence of 0.3%, an additive genetic model, and frequency of disease-associated allele of ≥ 20% in unaffected controls. For the GAE subgroup, 80% power is achieved for OR ≥ 1.57 under identical parameters, for JME subgroup at OR ≥ 1.64. Within trio analysis, GGE trios have 80% power to detect a variant with OR ≥ 1.87, for GAE trios with OR ≥ 2.35, and for JME trios with OR ≥ 3.60 (estimation on a pseudo-case-control partition of transmitted to untransmitted alleles). Genome-wide meta-analysis obtained 80% power for GGE with OR ≥ 1.34, for GAE with OR ≥ 1.47, and for JME with OR ≥ 1.56. Sizes of sample cohorts are as follows: Case-control analysis of GGE: 1,523 cases vs. 2,454 controls, GAE: 705 cases vs. 2,454 controls, JME: 579 cases vs. 2454 controls; TDT analysis of GGE: 556 cases, GAE: 317, JME: 157; meta-analysis of GGE: 2,089 cases vs. 3,020 controls, GAE: 1,022 cases vs. 2,771 controls, JME: 736 cases vs. 2,611 controls.

Detailed power calculations of case-control, TDT and meta-analysis were performed with the PGA2 software (http://dceg.cancer.gov/tools/analysis/pga) (Menashe et al., 2008). Minimal detectable OR at 80% power for GGE, GAE and JME are shown in Figure 2-3. Power calculations were performed as above, assuming a disease prevalence of 0.3%, the additive risk model and r^2 = 0.9 between a causal variant and a genotyped.

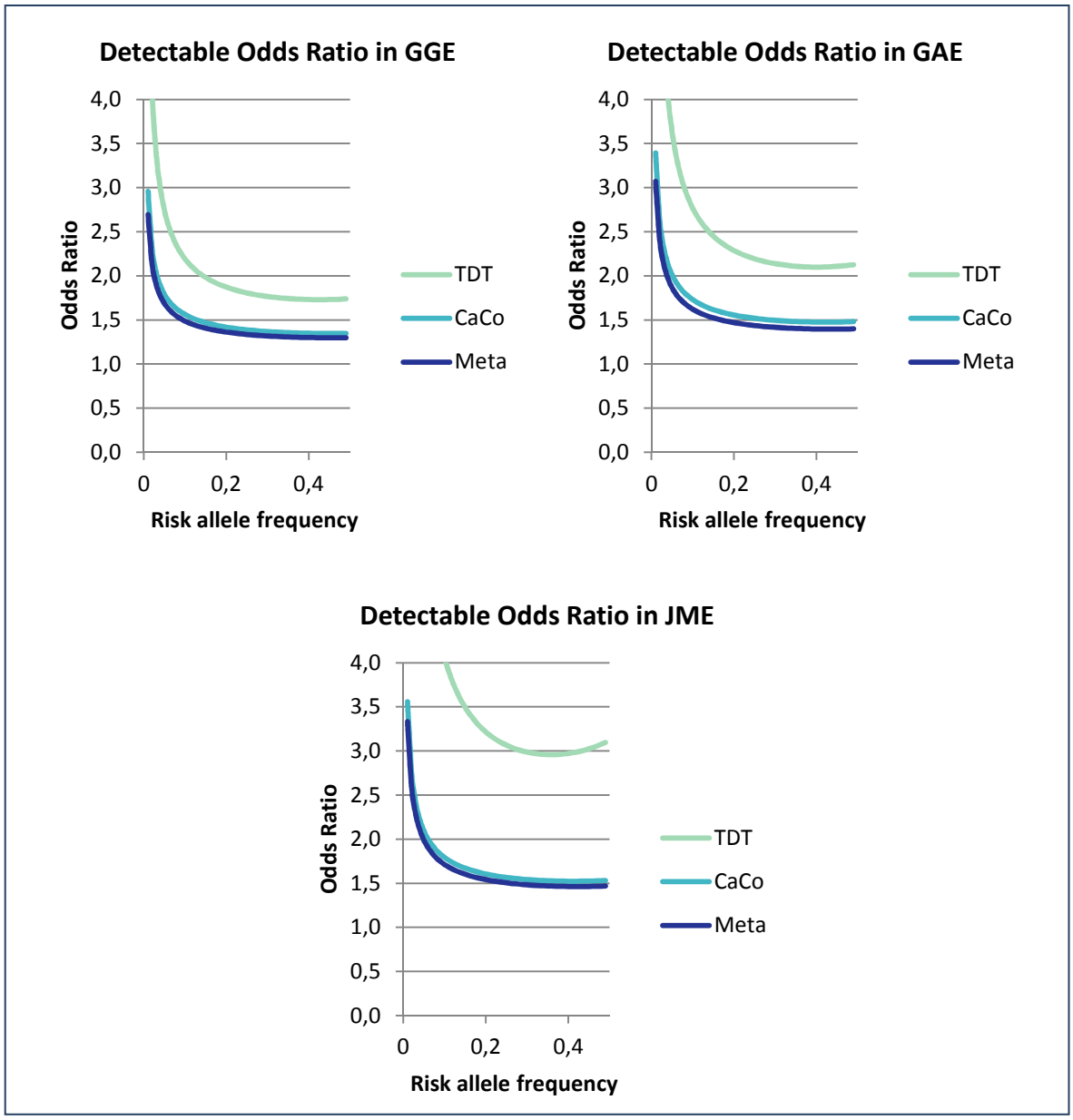


Figure 2-3 | Detectable Odds Ratios for genome-wide association of TDT, case-control and meta-analysis in GGE, GAE and JME

2.3.10 GWAS significance threshold and presentation of results

Threshold for genome-wide significance was set to $P \leq 5 \times 10^{-8}$, for suggestive association to $P \leq 1 \times 10^{-5}$ (McCarthy et al., 2008). Results were plotted as Manhattan and QQ-plots with a custom R script (*qqman.r* by Stephen Turner, 2011, University of Virginia, http://GettingGeneticsDone.blogspot.com/). Regional details of significant results to show local LD structure were created using LocusZoom (http://genome.sph.umich.edu/wiki/LocusZoom) (Pruim et al., 2010).

2.4 Copy number analysis

2.4.1 Copy number analysis using Affymetrix SNP6.0 arrays

Beside more than 900,000 SNP probes, the Affymetrix SNP6.0 array features 946,000 non-polymorphic Copy Number Polymorphism (CNP) markers to detect CNVs. CNV calling was performed by the Affymetrix Genotyping Software using the Birdseye algorithm (Korn et al., 2008). Birdseye is a Hidden Markov Model (HMM) to find regions of variable copy number in a sample. The hidden state is the true copy number of the individual's genome; the observed states are the normalized intensity measurements of each probe on the array. Intensity measurements therefore allow for differentiation between normal CNV state, CNV losses or CNV gains. Breakpoint estimation of CNV segments is not exact and depends on probe resolution of the observed segment and CNV calling performance is heavily dependent on the general number of markers called. Hence large CNVs with a high number of array probe sets and a high resolution yield higher detection confidence. Only samples arrays fulfilling the GWAS array QCs were included in CNV analyses.

After CNV calling and segmentation, CNV segments were quality filtered. Only segments with ≥ 20 SNP and/or CNP markers and ≥ 40 kb in size were kept for analysis. For CNVs with less markers and smaller in size, false-positive error rate is expected to exceed a reasonable level. CNVs with a minimum of 20 markers and 40 kb are expected to be relatively valid, even if the arrays were processed at different laboratories (Pinto et al., 2011). Centromere-spanning CNVs were eliminated from the data, as they represent typical technical artifacts.

2.4.2 Loss-of-Heterozygosity screening to predict recurrent microdeletions

In case of a genetic microdeletion, all heterozygous SNP genotypes are lost due to the loss of one homologous chromosomal strand in the affected region, leading as well to LOH. The co-occurrence of deletions and LOH is used as an additional marker together with signal intensity in various algorithms for CNV calling (Colella et al., 2007; Wang et al., 2007).

For the Affymetrix Axiom Genome-Wide array, CNV calling is not supported by the standard software. Technical limitations of the array, in particular the low

redundancy of probe features for the benefit of small array size, are complicating normalization procedures and CNV calling. Although a successful approach is reported for the Axiom Bovine array in cattle (Rincon et al., 2011), our own efforts were unsuccessful. To date, no method is available for analysis of human CNVs with Axiom Genome-wide arrays. Using the genotype information acquired with Axiom arrays, a screening for LOH is practical by implication. Individuals carrying a LOH in a defined region where genomic deletions have been reported previously are considered to be possible deletion carriers as well. Chances of success have been estimated by frequency of reported deletions in affected and unaffected individuals. In addition, Mendelian errors are additional evidence for occurrence of deletions in trios. LOH analysis was performed using PLINK (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell et al., 2007).

To increase consistency of the LOH analysis, we genotyped our samples with a decreased Confidence score. The Confidence score with a value between 0 and 1 represents the ratio of the distance from the closest cluster to the second closest cluster during genotyping. Lower values are more confident, larger values of the confidence threshold indicate less certain calls. Calls with Confidence scores above the threshold are assigned a no-call. To obtain higher genotype reliability, we changed the Confidence score for the LOH analysis data set from 0.15 to 0.005.

PLINK LOH analysis results in a summary table for each LOH region per sample, including position, size, marker density and rate of homozygous and heterozygous markers that were used for candidate selection. In addition, we calculated a score by taking into account the ratio of heterozygous to non-heterozygous (including missing genotypes). We assume, that for a LOH with a given set of markers the homozygosity rate is $P_{HOM} = 1$, together with a heterozygosity rate $P_{HET} = 0$ and vice versa. With an increasing number of missing (non-informative) markers (P_{MISS}), we expect the reliability of genotype calls to decrease as well as the impact of observed heterozygous makers. To take this into account, we calculated a score value called S_{HET} to represent the ratio of heterozygous markers P_{HET} multiplied with the ratio of heterozygous P_{HET} to non-heterozygous markers ($P_{HET} + P_{MISS}$):

$$S_{HET} = P_{HET} * \left(\frac{P_{HET}}{P_{HET} + P_{MISS}}\right)$$

where (P_{HET} + P_{MISS}) > 0.The more markers are non-informative or missing, the smaller the ratio P_{HET} versus (P_{HET} + P_{MISS}), the lesser the effect of P_{HET} in the observed region of LOH. By multiplying P_{HET} with the ratio P_{HET} versus (P_{HET} + P_{MISS}), the effect size of P_{HET} is shown under the impact of P_{MISS}. As follows:

$$S_{HET} = P_{HET} * \left(\frac{P_{HET}}{P_{HET} + P_{MISS}}\right) = \frac{P_{HET}^2}{1 - P_{HOM}}$$

where (P_{HET} + P_{MISS}) > 0, P_{HOM}< 1. As for P_{HET}, a S_{HET} = 1 represents a heterozygous and S_{HET} = 0 a homozygous site, but with regard to the occurrence of heterozygous markers in relation to the number of missing genotypes as an indicator for a low regional genotype quality.

Screening for LOHs to identify possible deletion carriers required verification of microdeletions by additional methods.

2.4.3 Quantitative real-time PCR

Reliable quantitative analysis of a specific gene or region is a critical aspect for Copy Number analysis. Although DNA microarrays are very efficient, they are often too expensive for an explicit screening of a small region of interest and their results require confirmation. Quantitative Polymerase Chain Reaction (qPCR) is a powerful tool for this task. We used TaqMan® Copy-Number assays (Applied Biosystems, Lifetech, USA) for a specific real-time qPCR setup with an internal reference for each single reaction.

The qPCR reaction is performed with the Applied Biosystems7900HT Sequence Detection System, according to standard protocol (*Determining Gene copy number using TaqMan® real-time PCR assay on the7900HT*, ABI 2006). TaqMan® CN assays consist of an oligonucleotide primer pair specific to the region of interest and an unique probe that will bind between primers during qPCR reaction. The probe works of the *Fluorescence resonance energy transfer* (FRET) principle and features fluorescent donor and molecules. During qPCR, the probe binds to the DNA target sequence and is destroyed during elongation phase. This separates the donor from the acceptor changing the emitted fluorescence signal, which is measured at the end of the elongation phases. For Copy Number estimation, a reference reaction with the highly conserved single copy *RNase P* gene is performed within the same reaction

well. Copy Number of the samples is calculated by comparison of signal intensity change over time for the target assay against the reference *RNase P* assay. Each sample is analyzed as a triplet (or quadruplet if required by assay performance) in order to obtain a sufficient statistical basis for data analysis The SDS software and CopyCaller software (both Applied Biosystems, Lifetech, USA) are used for analysis and presentation of results.

2.4.4 Array comparative genomic hybridization

Array comparative genome hybridization (aCGH) is a well-proven method for CNV detection and validation (Itsara et al., 2009). In contrast to SNP based microarray CNV analysis, aCGH supplies better results for breakpoint estimation and a higher confidence of results. All aCGH assays are based on the comparison of two distinct labeled samples, one target and one reference sample. Sample are hybridized to one array carrying long oligonucleotide (Sebat et al., 2004) or bacterial artificial chromosome (BAC) clone (Iafrate et al., 2004) target features. Signal intensity ratio of both samples is used as a proxy for CNV detection. A well described reference sample is essential, as unreported CNVs in the reference may lead to identification of false-positive variants in the target sample. Array CGH analysis applied in this study has been performed in the laboratory of Heather Mefford in the Department of Pediatrics, Division of Genetic Medicine at the University of Washington in Seattle.

2.4.5 Association tests based on contingency tables

A contingency table displays the frequency distribution of continuously variables to be tested. Those variables are tested to fit a null hypothesis, an expectation that is fulfilled if the observed values are equal to the expected ones. If applicable, Pearson's χ^2-test is used. As for frequency distributions of ≤ 5 Pearson's χ^2-test yields no exact results, Fisher's exact test is used. Calculation of p-values was performed with standard functions for χ^2-test and Fishers exact test within the R software environment (http://www.r-project.org/).

Where applicable, threshold for significance has to be adjusted for multiple testing. This is done using the method from Bonferroni (Bland and Altman, 1995), were the threshold for significance is 0.05 / number of tests.

2.4.6 Analysis of recurrent microdeletions

Six recurrent microdeletions were selected for evaluation of contributing to risk for IGE. An excess of these six large candidate microdeletions has been previously reported in a wide spectrum of neuropsychiatric disorders including autism, intellectual disability and schizophrenia (Ullmann et al., 2007; Cook and Scherer, 2008). The selection of microdeletions at 1q21.1 (Brunetti-Pierri et al., 2008; Mefford et al., 2008; Schizophrenia Consortium, 2008; Stefansson et al., 2008; Need et al., 2009), 15q11.2 (Stefansson et al., 2008; Kirov et al., 2009), 16p11.2 (Sebat et al., 2007; Kumar et al., 2008; Marshall et al., 2008; Weiss et al., 2008), 16p13.11 (Ullmann et al., 2007; Hannes et al., 2009; Need et al., 2009) and 22q11.2 (Bassett et al., 2008; Schizophrenia Consortium, 2008; Need et al., 2009) was based on previous copy number variation and a CNV meta-analyses (Itsara et al., 2009), and required following inclusion criteria: Recurrent microdeletion mediated by NAHR (equal in size and with defined breakpoints); previous association of the microdeletion with neuropsychiatric disorders ($P < 0.05$), and size of the deletion larger than >400 kb to ensure a reliable detection by the Affymetrix SNP 6.0 array. Microdeletions were considered to match the published deletions if they overlapped at least 85% of the genomic region of the candidate microdeletion (Table 2-1).

Initially, Affymetrix SNP6.0 array data of 1523 GGE samples and 2454 controls were selected for CNV analysis. Candidate microdeletion studies have been published previously for smaller subsamples of this ongoing research project (Helbig et al., 2009; de Kovel et al., 2010). The sample cohort used in our study is identical with the case-control cohort used for GWAS and meta-analysis, and thus did undergo same QC parameters that were applied prior to imputing. In addition, we added Affymetrix SNP6.0 array data of 3,413 additional unrelated German controls from the SHIP cohort, which underwent the same QC filtering procedures. SHIP samples were combined with the data set of the GWAS case-control cohorts and were examined for stratification by PCA.

Copy Number calling was performed with the Affymetrix Genotyping Console. To ensure quality of CNV calls, only CNV segments with more than 20 markers and at least 40 kb in size were taken into account. Samples with a high number of CNVs (number of CNVs per sample > average number of CNVs of all samples $\pm$ 3 SD) and

an excessive CNV load (total size of CNVs per sample > average total size of CNVs of all samples ± 3 SD) were excluded from the analysis. After filtering, 1497 IGE samples and 5374 controls (KORA n = 1348, PopGen n = 1079, SHIP n = 2947) remained.

Candidate regions were then explored for CNV segments. Changes of the heterozygosity state and log2 ratios along with candidate deletions were visually inspected to exclude technical artifacts. Identified CNV segments were validated with quantitative real-time PCR. Cross validation and exact breakpoint estimation were performed with array Comparative Genomic Hybridization. Association analyses between genotype and phenotype were carried out by two-sided $\chi 2$-tests or Fisher's exact tests where appropriate. If available, additional family members of microdeletion carriers were screened with either qPCR and/or aCGH for segregation analysis. Pedigree figures were generated with HaploPainter (http://haplopainter.sourceforge.net/) (Thiele and Nürnberg, 2005).

Table 2-2 | Candidate regions for recurrent microdeletions

Chr. segment	Chr. position (Mb)	MicroDel size (Mb)	Candidate gene	Neuropsychiatric disorder
1q21.1	146.5-147.9	1.35	GJA5, GJA8, HYDIN2	ID, SZ
15q11.2	22.75-23.2	0.45	CYFIP1	SZ
15q13.3	30.9-32.5	1.50	CHRNA7	ID/EPI, SZ, ASD
16p11.2	29.6-30.2	0.70	KCTD13, PRRT2 TAOK2	ASD, ID
16p13.11	14.8-16.4	1.60	NDE1	ID, SZ, ASD
22q11.2	19.2-22.2	3.00	COMT, SNAP29	SZ, ID, ASD

ID: Intellectual Disability; SZ: Schizophrenia; EPI: Epilepsy; ASD: Autism Spectrum Disorder; Chromosomal position in NCBI built 37.3, hg 19.

2.4.7 Family-based analysis of large recurrent microdeletions

A family-based study based on 566 parent-offspring trios with European origin was performed to further evaluate the frequency of the reported six recurrent microdeletions and to explore their pattern of inheritance. A standardized method for CNV calling was not available, due to the specifications of the Axiom array with

which the trios were called. Thus, screenings for LOH within candidate regions (at 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.11 and 22q11.2) was performed by PLINK (Purcell et al., 2007), to identify potential deletion carriers. For all 566 trios, genotype calling was repeated with an increased threshold for calling confidence (Affymetrix genotyping confidence value ≤ 0.005) to decrease the number of false genotype calls in the hemizygous regions. In total 326,123 SNPs fulfilling high QC filters were used for LOH screening.

LOH screening was performed within breakpoints of previously reported microdeletions, given by reported breakpoints and smallest possible overlap of known deletion carriers. In addition, an extended region of including each 1 Mb distal and proximal from the microdeletion was screened to identify deletions with deviating starting and/or ending positions. Well characterized deletion carriers from previous screenings were used to assess inclusion criteria for each region individually by rate of heterozygous genotypes and number of Mendelian errors per region (Table 2-3). Moreover, we applied a score called S_{HET} as auxiliary criteria to take the ratio of heterozygous to missing genotypes into account (see 2.4.2). Only samples with at least 85% overlap to a region with recurrent microdeletions were selected as microdeletion candidates. For validation of deletions, all trio members were genotyped with qPCR.

2.4.8 Penetrance estimations for copy number variants

For a detailed risk assessment, we assessed the penetrance of each candidate microdeletion. The penetrance of the microdeletions associated with GGE was estimated by a Bayesian approach previously reported for schizophrenia (Vassos et al., 2010). The probability P for developing a disease D for individuals carrying a CNV genotype G versus controls is calculated by the formula:

$$P(D|G) = \frac{P(G|D)P(D)}{P(G|D)P(D) + P(G|\dot{\mathrm{D}})\mathrm{P}(\dot{\mathrm{D}})}$$

Confidence intervals for penetrance estimations were calculated with the Clopper–Pearson exact tail area method (Rosenfeld et al., 2012) using the R software environment.

Table 2-3 | Position and inclusion criteria for LOH screening of deletion candidates

Region		Start Mb	End Mb	#SNPs	PHET	S_{HET}	Mendel errors
1q21.1	core	146.5	147.9	185	≤ 0.02	≤ 0.01	≥ 5
	extended	145.5	148.9	200	< 0.01	< 0.01	≥ 10
15q11.1	core	22.8	23.1	43	≤ 0.025	≤ 0.03	≥ 4
	extended	21.8	24.1	150	≤ 0.02	≤ 0.02	≥ 1
15q13.3	core	31.2	32.5	148	≤ 0.035	≤ 0.01	≥ 17
	extended	30.2	33.5	488	< 0.01	< 0.01	≥ 1
16p11.2	core	29.6	30.3	8	≤ 0.01	≤ 0.01	≥ 1
	extended	28.6	31.3	55	≤ 0.01	≤ 0.01	≥ 2
16p13.1	core	14.8	16.4	145	≤ 0.025	≤ 0.02	≥ 10
	extended	13.8	17.4	426	≤ 0.01	≤ 0.01	≥ 10
22q11.2	core	19.5	21.3	88	≤ 0.01	≤ 0.01	≥ 17

PHET: Proportion of heterozygous genotypes per LOH segment, S_{HET}: Heterozygosity score per LOH segment, #SNPs: Number of SNPs in the genomic region of interest. Physical position NCBI built 37.3, hg19.

2.4.9 Candidate gene approach to detect exon-ablating microdeletions

Exon-ablating microdeletions in *NRXN1*

To analyze the impact of exon-disrupting/removing microdeletions in *NRXN1* (2p16.3, 50,145,643 – 51,259,674, hg19), this candidate gene study includes 1,569 unrelated GGE cases of European ancestry and 6,201 German controls. All samples were genotyped with the Affymetrix SNP6.0 array, CNV analysis was performed by the algorithm implemented in the Affymetrix Genotyping Console. Segments with >20 markers and >40 kb in size were considered as highly confident CNV calls. Changes of the log2 signal intensity ratios of the *NRXN1* microdeletions were manually inspected to exclude technical artifacts. All exonic *NRXN1* microdeletions identified in the patient cohort were verified by TaqMan qPCR (*NRXN1* exon 4 Hs04683030_cn; Applied Biosystems, Foster City, CA, U.S.A.) and/or aCGH. For segregation analysis, all available family members were typed by qPCR and/or by aCGH. Pedigree figures were generated with HaploPainter (http://haplopainter.sourceforge.net/) (Thiele and Nürnberg, 2005).

Exon-ablating microdeletions in *RBFOX1*

Finally, the study consisted of 1,408 unrelated GGE patients of European ancestry (869 females, 539 males; CAE n = 413, JAE n = 207, GAE n = 7, JME n = 557, EGTCS n = 224) and 2,256 German population controls (1,077 females, 1,179 males). The control cohort of 2,256 German controls were collected from Southern (KORA, n = 1,185) (Wichmann et al., 2005) and Northern (PopGEN, n = 1,071) (Krawczak et al., 2006) regions of Germany.

All samples were genotyped with the Affymetrix SNP6.0 array, CNV analysis was performed by the algorithm implemented in the Affymetrix Genotyping. Segments with >20 markers and >40 kb in size were considered as highly confident CNV calls. We included only arrays which revealed genome-wide less than 50 microdeletions (> 20 markers, > 40 kb) (Elia et al., 2012).

All samples were screened for *RBFOX1* (16p13.3, 6,069,132 – 7,763,340, hg19) deletions affecting the genomic sequence of the gene. All potential *RBFOX1* microdeletions were manually inspected for the regional SNP heterozygosity state and log2 ratios of the signal intensities to exclude technical artifacts. Subsequently, the copy number state of all *RBFOX1* microdeletions identified was examined by qPCR, using seven TaqMan CNV assays covering the 5'-terminal *RBFOX1* exons 1-4 (Hs04461212_cn, Hs03952094_cn, Hs03929445_cn, Hs03930287_cn, Hs03948431_cn, Hs04457208_cn, Hs05452109_cn; Applied Biosystems, Foster City, CA, U.S.A.). For segregation analysis, all available family members were typed by qPCR and/or by aCGH. Pedigree figures were generated with HaploPainter (http://haplopainter.sourceforge.net/) (Thiele and Nürnberg, 2005).

3 Results

3.1 Genome-wide association and meta-analysis

3.1.1 Genome-wide association and meta-analysis of GGE patients

Genome-wide association studies were performed to identify common variants associated with GGE. Therefore, we analyzed two GWAS datasets: i) a case-control study with 1,523 European GGE cases and 2,454 German controls and ii) a family-based study with 566 parent offspring trios affected with GGE. Subsequently, GWAS meta-analysis was carried out by combining the SNP P-values of both GWAS datasets. The case-control data set in our study is equivalent to the Stage-1 data cohort of a previously published GWAS (EPICURE Consortium et al., 2012a).

Case-control GWAS of GGE

The GWAS of 1,523 GGE cases and 2,454 controls included 5,563,867 quality filtered SNPs. Using a LMM association statistic, no SNP reached the threshold for genome-wide significance (LMM $P < 5.0 \times 10^{-8}$). , but 44 SNPs reached the threshold for suggestive association (LMM $P < 1.0 \times 10^{-5}$) (Table 3-1, Figure 3-1). The most significant association occurred on chromosome 2p16.1 in an intergenic genomic region with no gene as close as 100 kb (rs35577149, chr2:57,931,347, LMM $P = 2.60 \times 10^{-6}$, OR[C] = 0.78, 95% CI 0.71 – 0.86). Genomic inflation factor was 0.998 and therefore no genomic correction applied (see QQ-plot in Suppl. 6.3.1).

Table 3-1 | Genome-wide association results of LMM P < 1.0×10^{-5} in GGE

SNP	CHR	POS	Allele A1/A2	Source	Info	LMM P	OR (95% CI)	Gene
rs35577149*	2	57,931,347	C/T	Imp	1.00	2.60E-07	0.78 (0.71-0.86)	
rs388556*	2	13,392,540	G/C	Geno	1.00	1.12E-06	0.76 (0.68-0.84)	
rs3106317	8	24,123,239	G/C	Geno	1.00	2.35E-06	0.76 (0.67-0.86)	ADAM28(28.3k)
rs12495872*	3	13,121,129	A/G	Imp	0.99	2.43E-06	0.66 (0.56-0.78)	IQSEC1(6.5k)
rs11890028	2	166,943,277	G/T	Imp	0.96	2.53E-06	0.77 (0.70-0.85)	SCN1A(13.1k)
rs2268330	14	23,795,976	G/C	Geno	1.00	2.89E-06	1.30 (1.17-1.44)	PABPN1(0.6k)
rs75414440	8	89,603,781	G/A	Imp	0.99	5.23E-06	1.43 (1.21-1.69)	
rs35803605*	2	38,760,917	A/G	Imp	0.97	5.36E-06	0.76 (0.67-0.85)	HNRPLL(29.4k)
rs3818509*	9	130,670,532	T/C	Imp	0.97	5.58E-06	0.70 (0.60-0.81)	ST6GALNAC4
rs2254349*	13	111,987,226	C/G	Imp	0.97	5.84E-06	0.81 (0.74-0.89)	C13orf16
rs1170543	1	34,744,384	C/G	Imp	1.00	9.50E-06	1.27 (1.14-1.42)	C1orf94(59.7k)
rs289034	5	114,122,165	C/T	Geno	1.00	9.60E-06	0.80 (0.73-0.88)	

CHR: Chromosome, POS: Physical position NCBI built 37.3 hg19, Allele A1: Minor allele, Allele A2: Major allele, Imp: SNP imputed, Geno: SNP genotyped, Info: SNPtest info value, Gene: Gene nearby the associated SNP, distance in brackets. The top-ranked SNP was represented in regions with several single marker association signals exceeding LMM P < 1.0×10^{-5} (*).

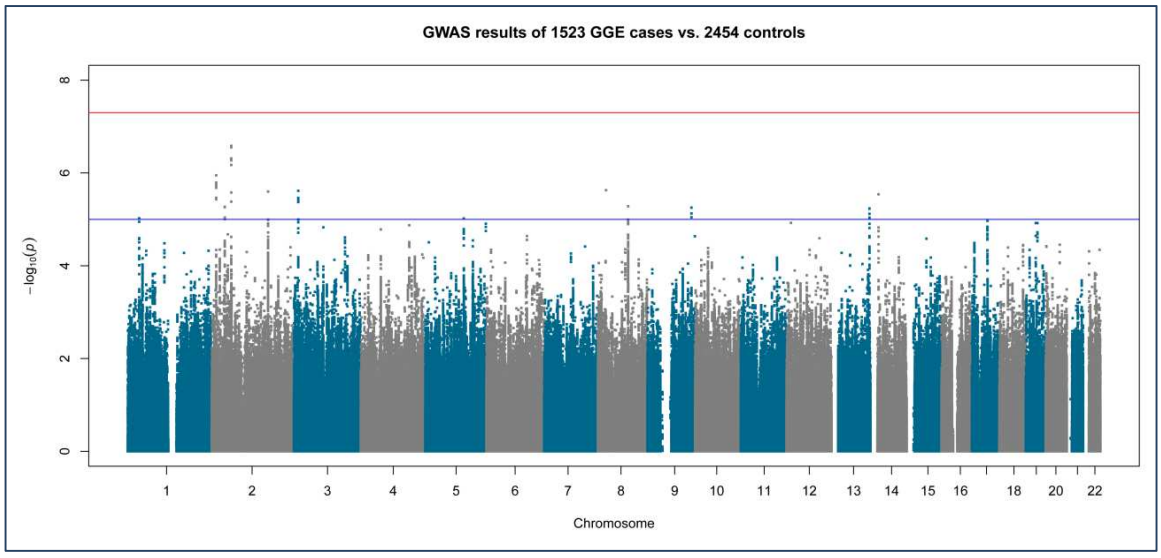


Figure 3-1 | Manhattan plot of genome-wide association results in 1,523 GGE cases versus 2,454 controls. The x–axis shows the genomic position of SNP markers on chromosomes 1 to 22 (NCBI built 37.3, hg19), the y-axis shows the negative log10 of the single marker P-value of the linear mixed model (LMM) association statistic. The blue line represents the threshold for a suggestive association with P < 1.0×10^{-5}, the red line for significant association with P < 5.0×10^{-8}.

Family-based GWAS of GGE

The transmission disequilibrium association test was performed for 566 GEE trios. The GWAS included 5,098,823 quality-filtered SNPs. No genome-wide significance was observed, but 64 SNPs showed suggestive association (Table 3-2, Figure 3-2). The lowest TDT P-value was detected on chromosome 8q12.2 (rs13266191, chr8:59,024,803, TDT P = 4.14×10^{-7}, OR[A] = 0.50, 95% CI 0.38 – 0.66) within the genomic sequence of *FAM110B*. All P-values were corrected for genomic inflation ($\lambda = 1.015$, see QQ-plot in Suppl. 6.3.1).

Table 3-2 | Genome-wide association results of TDT P < 10^{-5} in GGE

SNP	CHR	POS	Allele A1/A2	Source	Info	TDT P	OR (95% CI)	Gene
rs13266191*	8	59,024,803	A/G	Imp	0.97	4.14E-07	0.50 (0.38-0.66)	FAM110B
rs8071170*	17	53,420,489	G/A	Imp	0.99	2.89E-06	1.51 (1.27-1.80)	HLF(18.1k)
rs8002187*	13	61,602,325	C/T	Imp	0.99	3.71E-06	1.50 (1.26-1.78)	
rs2861858*	18	36,702,626	G/A	Imp	0.97	5.16E-06	0.67 (0.56-0.79)	LOC647946(84.3k)
rs71540757	7	16,653,094	A/G	Imp	0.96	6.16E-06	1.81 (1.40-2.34)	ANKMY2
rs2159129	19	29,049,808	C/T	Imp	0.98	7.09E-06	1.52 (1.27-1.83)	
rs10510254	3	3,279,498	C/T	Imp	0.99	7.15E-06	0.50 (0.37-0.68)	CRBN(58.1k)
rs9815296	3	178,152,351	T/G	Imp	0.99	7.91E-06	2.20 (1.55-3.13)	
rs7556080	1	154,638,826	A/G	Geno	0.98	8.45E-06	1.53 (1.27-1.85)	ADAR(38.4k)
rs4612821	11	134,444,050	T/C	Imp	0.97	9.21E-06	0.64 (0.53-0.78)	
rs3779323	7	77,684,897	C/T	Geno	0.99	9.93E-06	0.66 (0.55-0.79)	MAGI2

CHR: Chromosome, POS: Physical position NCBI built 37.3 hg19, Allele A1: Minor allele, Allele A2: Major allele, Imp: SNP imputed, Geno: SNP genotyped, Info: SNPtest info value, Gene: Gene nearby the associated SNP, distance in brackets. The top-ranked SNP was represented in regions with several single marker association signals exceeding TDT P < 1.0×10^{-5} (*).

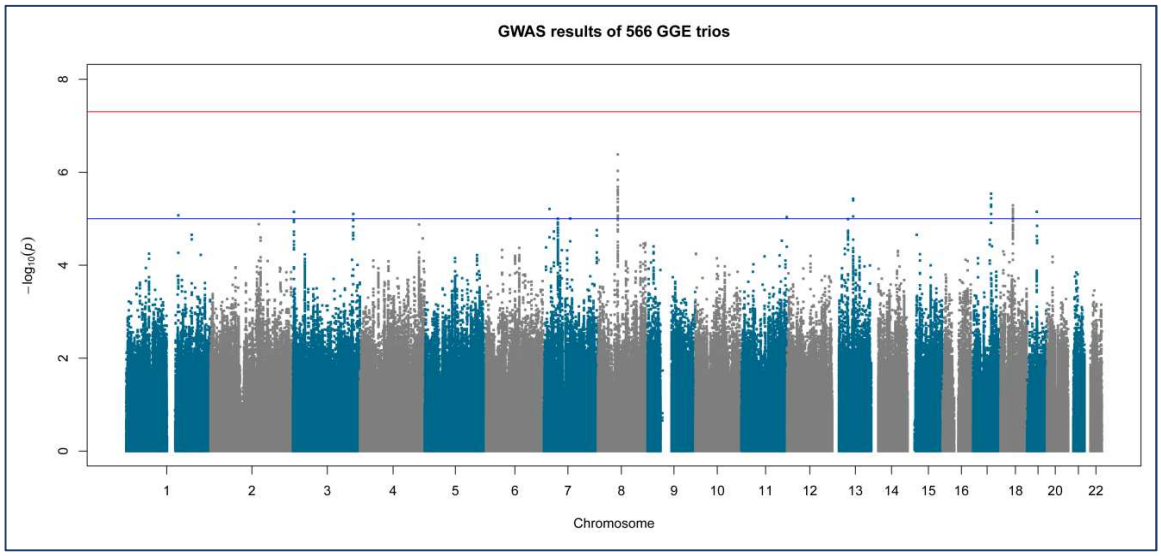


Figure 3-2 | Manhattan plot of genome-wide association in 566 GGE trios. The x–axis shows the genomic position of SNP markers on chromosomes 1 to 22 (NCBI built 37.3, hg19), the y-axis shows the negative log10 of the single marker P-value of the Transmission Disequilibrium Test (TDT) association statistic. The blue line represents the threshold for a suggestive association with $P < 1.0 \times 10^{-5}$, the red line for significant association with $P < 5.0 \times 10^{-8}$.

GWAS meta-analysis of GGE

The results of GWAS meta-analysis of 1,523 GGE cases, 2,454 controls and 566 GGE trios with 4,224,196 overlapping SNPs are shown in Table 3-3 and Figure 3-3. In total, seven SNPs showed genome-wide significance ($P < 5.0 \times 10^{-8}$). The top-ranked association was found in an intergenic region on chromosome 2p16.1 (rs35577149, chr2:57931347, meta-analysis $P = 1.65 \times 10^{-8}$, OR[C] = 0.78, 95% CI 0.71 – 0.86). In addition, 41 SNPs reached the threshold for suggestive association (LMM $P < 1.0 \times 10^{-5}$). All meta-analysis P-values were corrected for genomic inflation ($\lambda = 1.021$, see QQ-plot in Suppl. 6.3.1).

SNP markers with observed genome-wide significance were plotted with LocusZoom for regional view and to indicate the LD of neighboring SNPs with reference to the SNP with the most significant P-value. The top-ranked SNP rs35577149 is in strong LD with six significant SNPs at chromosome 2p16.1 (Figure 3-4).

Table 3-3 | Genome-wide association meta-analysis results of P < 10^{-5} in GGE

SNP	CHR	POS	Allele A1/A2	LMM P	TDT P	Meta P	OR_{meta} (95% CI)	Gene
rs35577149*	2	57,931,347	C/T	2.60E-07	0.01	1.65E-08	0.79 (0.73-0.86)	
rs6999304*	8	89,590,027	G/C	3.06E-05	0.01	1.77E-06	1.33 (1.17-1.51)	
rs12059546*	1	239,970,097	G/A	4.75E-05	0.01	2.31E-06	1.29 (1.17-1.43)	CHRM3
rs62059804*	17	7,453,505	A/C	3.70E-05	0.02	2.63E-06	0.79 (0.72-0.87)	TNFSF12
rs17734302*	7	21,092,494	A/G	5.60E-04	3.66E-04	2.99E-06	0.78 (0.70-0.87)	
rs17779783*	20	2,121,659	C/T	3.88E-05	0.04	4.95E-06	1.55 (1.28-1.88)	STK35
rs3101626	15	27,820,990	C/T	2.57E-04	5.61E-03	7.42E-06	1.27 (1.14-1.41)	GABRG3(42.9k)
rs113403571*	17	46,097,524	G/A	1.52E-05	0.14	7.81E-06	0.76 (0.68-0.84)	COPZ2(6.0k)

CHR: Chromosome; POS: Physical position NCBI built 37.3 hg19, Allele A1: Minor allele, Allele A2: Major allele, Gene: Gene nearby the associated SNP, distance in brackets. The top-ranked SNP was represented in regions with several single marker association signals exceeding Meta P < 1.0×10^{-5} (*).

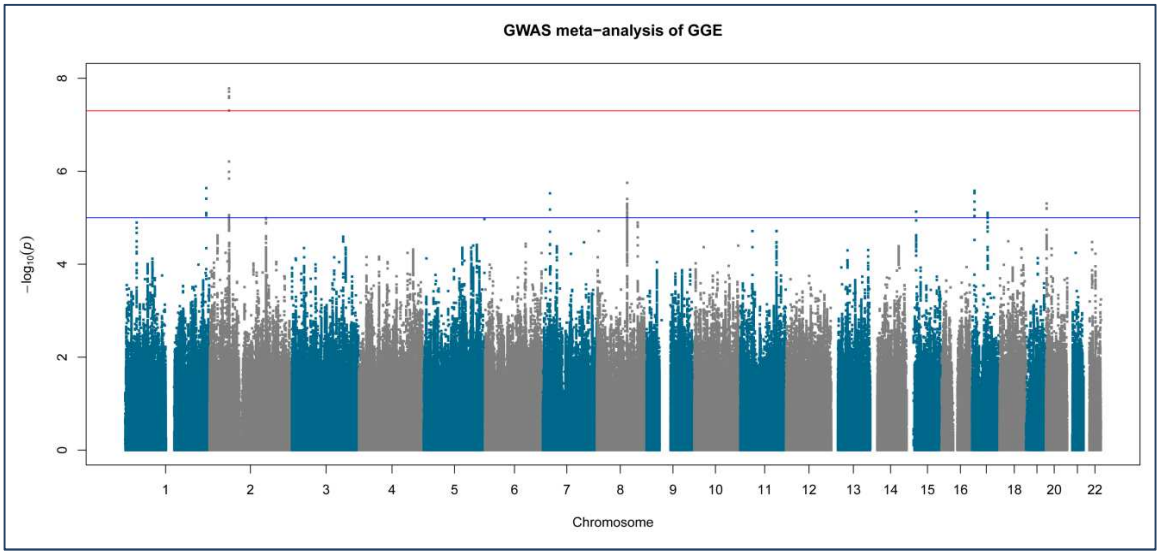


Figure 3-3 | Manhattan plot of genome-wide association meta-analysis in 1,523 GGE cases, 2,454 controls and 566 GGE trios. The x–axis shows the genomic position of SNP markers on chromosomes 1 to 22 (NCBI built 37.3, hg19), the y-axis shows the negative log10 of the meta P-value per SNP. The blue line represents the threshold for a suggestive association with P ≤ 1.0×10^{-5}, the red line for significant association with p ≤ 5.0×10^{-8}.

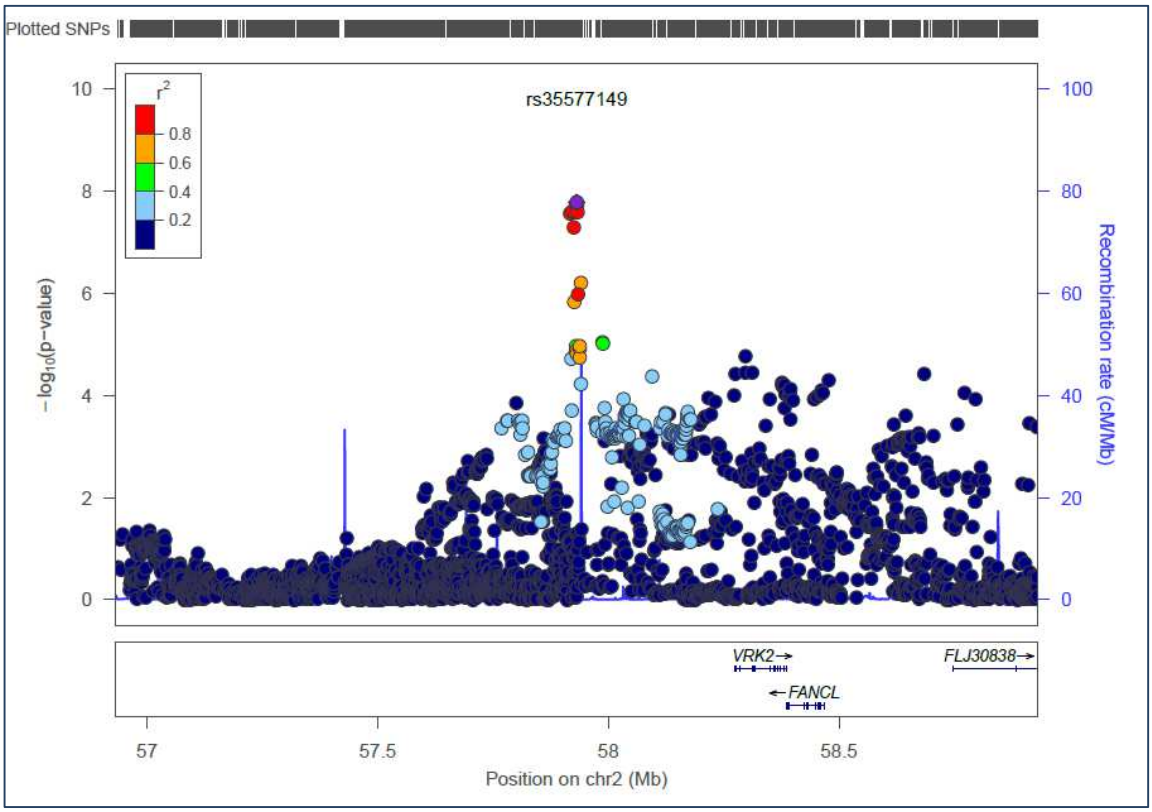


Figure 3-4 | Regional detail plot of the $-\log_{10}$ of LMM P-value of GGE for SNP rs35577149 on chromosome 2p16.1, presenting genes and LD of neighboring SNPs close to detected association. SNP color code is complied with r^2 value referring to rs35577149, physical position in NCBI built 37.3, hg19.

3.1.2 Genome-wide association and meta-analysis of GAE patients

GWAS analysis of the subgroup of patients with GAE included two datasets: i) 705 European GAE cases, 2,454 German controls, and ii) 317 European parent-offspring trios affected by GAE. The results of both association studies were combined into a genome-wide meta-analysis. The case-control data set in our study is equivalent to the Stage-1 data cohort of a previously published GWAS (EPICURE Consortium et al., 2012a).

Case-control GWAS of GAE

Genome-wide association analysis with 5,564,397 SNPs was performed using 705 GAE patients and 2,454 controls. No SNP was observed with a P-value lower than the threshold for significance. Suggestive association was detected in 86 SNPs (Table 3-4, Figure 3-5). Lowest P-value was found at 2q22.3 (rs75917352, chr2:145370978, LMM P = 3.54×10^{-7}, OR[T] = 0.63, 95% CI 0.53 – 0.76), with

LOC100131409 as the nearest gene locus (54.4k apart). Correction of P-values for genomic inflation was not required ($\lambda = 0.990$, see QQ-plot in Suppl. 6.3.2).

Table 3-4 | Genome-wide association results of LMM P < 10^{-5} in GAE

SNP	CHR	POS	Allele A1/A2	Source	Info	LMM P	OR (95% CI)	Gene
rs75917352*	2	145,370,978	T/G	Imp	0.99	3.54E-07	0.63 (0.53-0.76)	LOC100131409(54.4k)
rs6683622*	1	186,886,134	A/G	Imp	1.00	3.58E-06	1.34 (1.18-1.53)	PLA2G4A
rs13030122*	2	38,759,473	A/G	Imp	0.98	4.47E-06	0.69 (0.60-0.81)	HNRPLL(30.9k)
rs12904369*	15	97,335,674	A/C	Geno	1.00	4.70E-06	0.73 (0.64-0.83)	SPATA8(6.8k)
8-2706322	8	2,706,322	C/T	Imp	0.92	6.83E-06	1.67 (1.31-2.15)	CSMD1(86.6k)
rs17394336	22	26,665,807	A/G	Imp	0.98	6.85E-06	1.51 (1.25-1.82)	#SEZ6L#
rs9948832*	18	25,857,441	A/C	Imp	1.00	6.93E-06	1.37 (1.20-1.57)	CDH2(100.0k)
rs1859161	4	106,042,692	T/C	Geno	1.00	7.70E-06	0.63 (0.51-0.77)	TET2(25.2k)

CHR: Chromosome, POS: Physical position NCBI built 37.3 hg19, Allele A1: Minor allele, Allele A2: Major allele, Imp: SNP imputed, Geno: SNP genotyped, Info: SNPtest info value, Gene: Gene nearby the associated SNP, distance in brackets. The top-ranked SNP was represented in regions with several single marker association signals exceeding LMM P < 1.0×10^{-5} (*).

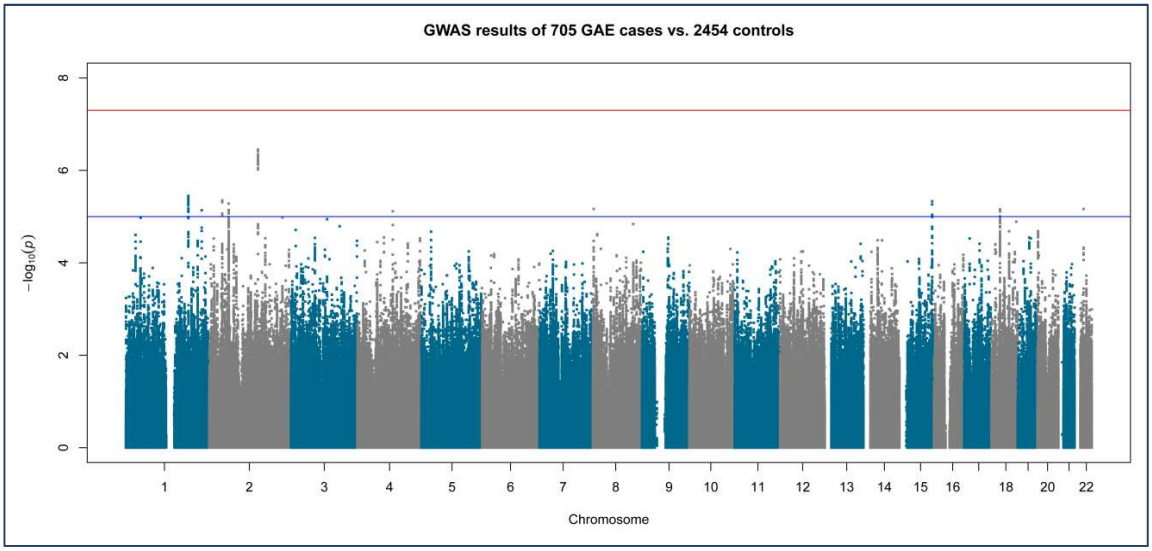


Figure 3-5 | Manhattan plot of genome-wide association in 705 GAE cases versus 2,454 controls. The x–axis shows the genomic position of SNP markers on chromosomes 1 to 22 (NCBI built 37.3, hg19), the y-axis shows the negative log10 of the LMM P-value per SNP. The blue line represents the threshold for a suggestive association with P ≤ 1.0×10^{-5}, the red line for significant association with p ≤ 5.0×10^{-8}.

Family-based GWAS of GAE

The family-based GWAS included 317 affected trios and 5,098,822 SNPs in the GAE subgroup (Table 3-5, Figure 3-6). No SNP with significant association was identified, but 19 SNPs showed suggestive association. The most prominent result was detected on chromosome 2q14.1 (rs6721192, chr2:116,483,984, TDT P = 1.42×10^{-6}, OR[T] = 0.56, 95% CI 0.44 – 0.71), in the genomic sequence of the gene DPP10. All P-values were corrected for genomic inflation (λ = 1.011, see QQ-plot in Suppl. 6.3.2).

Table 3-5 | Genome-wide association results of TDT P < 10^{-5} in GAE

SNP	CHR	POS	Allele A1/A2	Source	Info	TDT P	OR (95% CI)	Gene
rs6721192*	2	116,483,984	T/C	Imp	0.99	1.42E-06	0.56 (0.44-0.71)	DPP10
rs12449765	17	32,840,267	G/A	Geno	1.00	1.80E-06	2.28 (1.61-3.23)	C17orf102(60.9k)
rs114488716	5	160,281,684	G/A	Imp	0.99	3.03E-06	0.28 (0.16-0.49)	ATP10B(2.5k)
rs37277*	7	41,345,488	G/A	Geno	1.00	3.77E-06	0.46 (0.33-0.65)	
rs6735372*	2	60,538,864	A/G	Imp	0.99	4.37E-06	2.00 (1.48-2.70)	
rs1417251	1	102,920,751	T/G	Imp	0.99	5.19E-06	3.00 (1.83-4.92)	
rs10484968*	6	120,396,111	T/C	Geno	1.00	5.67E-06	7.75 (2.74-21.95)	
rs4903677	14	78,425,068	G/C	Geno	1.00	7.93E-06	2.10 (1.51-2.92)	ADCK1(24.8k)
rs72813644*	17	14,708,060	A/T	Imp	0.99	8.48E-06	0.38 (0.25-0.59)	FLJ45831(24.6k)
rs60782598	5	27,150,498	A/G	Impd	0.99	9.85E-06	0.55 (0.42-0.72)	

CHR: Chromosome, POS: Physical position NCBI built 37.3 hg19, Allele A1: Minor allele, Allele A2: Major allele, Imp: SNP imputed, Geno: SNP genotyped, Info: SNPtest info value, Gene: Gene nearby the associated SNP, distance in brackets. The top-ranked SNP was represented in regions with several single marker association signals exceeding TDT P < 1.0×10^{-5} (*).

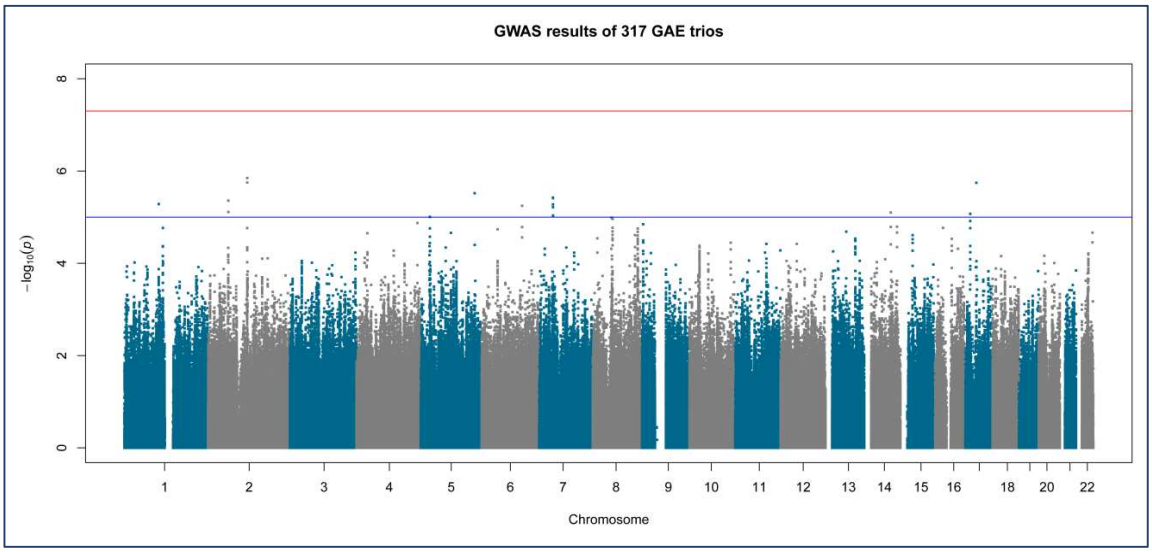


Figure 3-6 | Manhattan plot of genome-wide association in 317 GAE trios. The x–axis shows the genomic position of SNP markers on chromosomes 1 to 22 (NCBI built 37.3, hg19), the y-axis shows the negative log10 of the single marker P-value of the TDT association statistic. The blue line represents the threshold for a suggestive association with $P < 1.0 \times 10^{-5}$, the red line for significant association with $P < 5.0 \times 10^{-8}$.

GWAS meta-analysis of GAE

Results of meta-analysis of 705 IAE cases, 2,454 controls and 317 IAE trios, and 4,224,597 overlapping SNPs are shown in Table 3-6, Figure 3-7. No SNP reached threshold for genome-wide significance, but 64 SNPs reached the threshold for suggestive association. The highest peak was observed on chromosome 2q22.3 (rs75917352, chr2:145,370,978, meta $P = 1.41 \times 10^{-7}$, OR[T] = 0.67, 95% CI 0.58 – 0.79), the closest gene ZEB2 93.1 kb away. All P-values were corrected for genomic inflation ($\lambda = 1.009$, see QQ-plot in Suppl. 6.3.2).

Table 3-6 | Genome-wide association meta-analysis results of P < 10^{-5} in GAE

SNP	CHR	POS	Allele A1/A2	LMM P	TDT P	Meta P	OR_{meta} (95% CI)	Gene
rs75917352*	2	145,370,978	T/G	3.54E-07	0.12	1.41E-07	0.67 (0.58-0.79)	ZEB2(93.1k)
rs11585765*	1	226,704,825	T/C	7.31E-06	0.12	2.60E-06	0.75 (0.64-0.84)	C1orf95(31.7k)
rs9953249	18	74,380,851	A/G	1.29E-05	0.12	4.36E-06	0.70 (1.17-1.47)	FLJ44881(21.3k)
rs77859246*	11	122,976,828	A/T	2.36E-04	3.12E-03	5.58E-06	0.70 (0.60-0.82)	ASAM
rs7519192*	1	186,862,419	G/A	4.08E-06	0.40	5.85E-06	0.75 (1.15-1.45)	PLA2G4A
rs11682175*	2	57,987,593	C/T	3.43E-05	0.06	5.93E-06	0.75 (0.69-0.85)	
rs1418029	20	2,060,151	A/G	2.07E-05	0.12	6.77E-06	0.75 (1.31-1.98)	STK35(23.4k)
rs1519309	4	187,015,089	T/C	5.64E-05	0.04	6.79E-06	0.75 (1.14-1.43)	TLR3(8.8k)
rs57074415*	6	47,661,859	T/C	1.24E-03	1.66E-04	8.00E-06	0.72 (0.65-0.84)	GPR115(4.4k)
rs37277	7	41,345,488	G/A	4.85E-03	3.77E-03	9.06E-06	0.67 (0.61-0.82)	
rs17394336	22	26,665,807	A/G	6.85E-06	0.40	9.32E-06	0.75 (1.21-1.70)	SEZ6L
rs2215503*	18	25,890,637	G/A	1.62E-04	0.01	9.62E-06	1.35 (0.70-0.87)	

CHR: Chromosome, POS: Physical position NCBI built 37.3 hg19, Allele A1: Minor allele, Allele A2: Major allele, Gene: Gene nearby the associated SNP, distance in brackets. The top-ranked SNP was represented in regions with several single marker association signals exceeding Meta P < 1.0 × 10^{-5} (*).

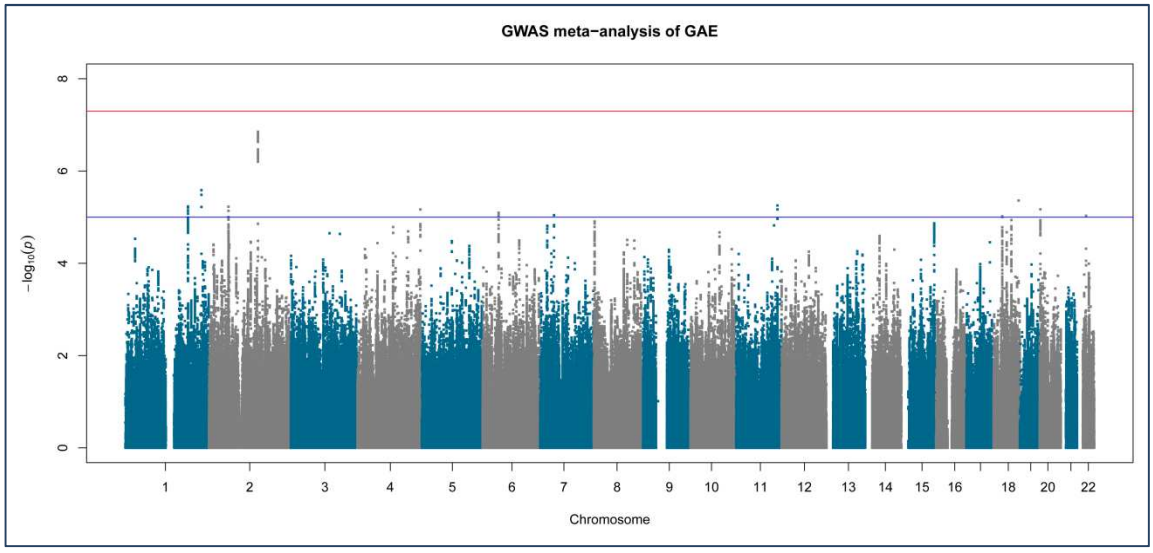


Figure 3-7 | Manhattan plot of genome-wide association meta-analysis in 705 GAE cases, 2,454 controls and 317 GAE trios; the x–axis shows the genomic position of SNP markers on chromosomes 1 to 22 (NCBI built 37.3, hg19), the y-axis shows the negative log10 of the meta P-value per SNP. The blue line represents the threshold for a suggestive association with P ≤ 1.0 × 10^{-5}, the red line for significant association with p ≤ 5.0 × 10^{-8}.

3.1.3 Genome-wide association and meta-analysis of JME patients

Genome-wide association was tested in the JME subgroup with 579 European JME cases, 2,454 German controls and 157 European parent-offspring trios affected by JME. The results of both association studies were combined into a genome-wide meta-analysis. The case-control data set in our study is equivalent to the Stage-1 data cohort of a previously published GWAS (EPICURE Consortium et al., 2012a).

Case-control GWAS of JME

The case-control based genome-wide association analysis in the JME subgroup was performed using 579 JME patients and 2,454 controls among 5,563,947 SNPs. No SNP reached the threshold for genome-wide significance, 64 SNPs reached the threshold for suggestive association (Table 3-7, Figure 3-8). The lowest P-value was observed at chromosome 1q43 (rs12059546, chr1:239970097, LMM $P = 1.97 \times 10^{-7}$, OR[G] = 1.53, 95% CI 1.32 – 1.79), within gene CHRM3. Genomic inflation factor was < 1 and therefore no genomic correction applied ($\lambda = 0.999$, see QQ-plot in Suppl. 6.3.3).

Table 3-7 | Genome-wide association results of LMM P < 10^{-5} in JME

SNP	CHR	POS	Allele A1/A2	Source	Info	LMM P	OR (95% CI)	Gene
rs12059546	1	239,970,097	G/A	Geno	1.00	1.97E-07	1.53 (1.32-1.79)	CHRM3
rs25833	5	66,175,175	T/C	Imp	0.97	3.00E-07	1.38 (1.20-1.58)	MAST4
rs3125293	1	212,763,530	G/T	Imp	0.97	2.48E-06	1.35 (1.18-1.55)	ATF3
rs34591495	3	48,644,202	T/C	Imp	0.95	2.91E-06	1.49 (1.24-1.79)	UQCRC1
rs13264787	8	128,631,805	G/T	Imp	0.94	3.28E-06	0.67 (0.56-0.80)	
rs17669194	13	34,267,584	T/C	Geno	1.00	4.35E-06	1.49 (1.26-1.77)	
18-20199084	18	20,199,084	C/A	Imp	0.92	6.07E-06	1.31 (1.15-1.49)	
rs6913416	6	157,454,046	C/T	Imp	0.98	6.28E-06	1.47 (1.21-1.79)	ARID1B
rs74621069	20	41,823,592	T/C	Imp	0.99	7.21E-06	2.02 (1.54-2.65)	PTPRT(5.0k)
rs11151793	18	70,422,507	C/T	Imp	1.00	7.25E-06	1.46 (1.26-1.69)	NETO1
rs12963674	18	55,978,595	T/G	Geno	1.00	7.47E-06	1.87 (1.46-2.40)	NEDD4L
rs79608766	20	41,82,3630	T/C	Imp	0.98	7.50E-06	2.02 (1.54-2.66)	PTPRT(5.1k)
rs6932561	6	14,076,716	T/C	Imp	1.00	9.36E-06	0.76 (0.66-0.86)	CD83(41.1k)
rs6039250	20	8,729,333	C/T	Imp	0.99	9.78E-06	0.67 (0.54-0.81)	PLCB1

CHR: Chromosome, POS: Physical position NCBI built 37.3 hg19, Allele A1: Minor allele, Allele A2: Major allele, Imp: SNP imputed, Geno: SNP genotyped, Info: SNPtest info value, Gene: Gene nearby the associated SNP, distance in brackets. The top-ranked SNP was represented in regions with several single marker association signals exceeding LMM P < 1.0 × 10^{-5} (*).

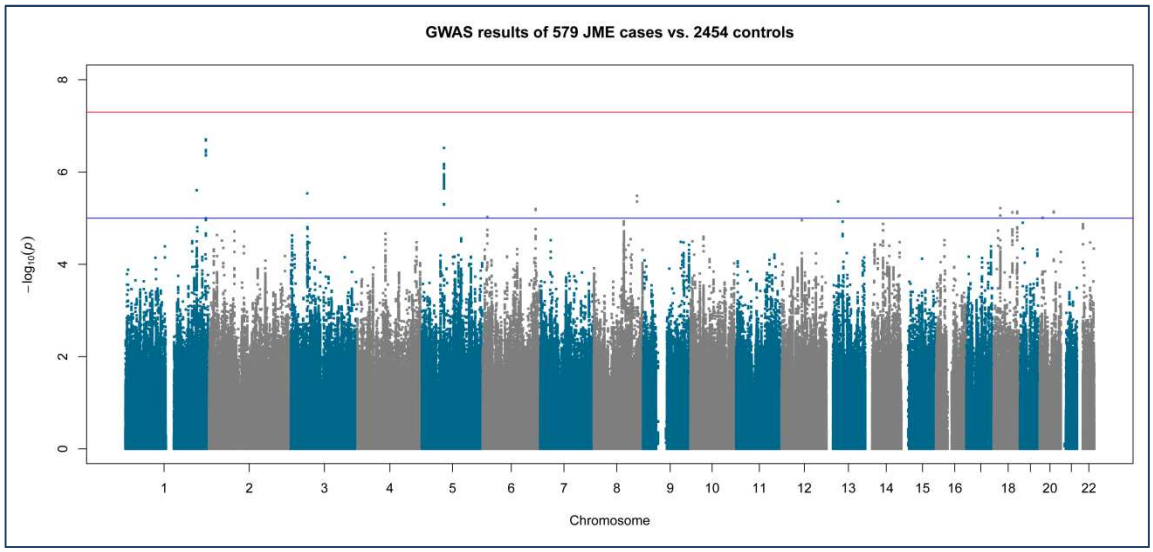


Figure 3-8 | Manhattan plot of genome-wide association in 579 JME cases versus 2,454 controls; the x–axis shows the genomic position of SNP markers on chromosomes 1 to 22 (NCBI built 37.3, hg19), the y-axis shows the negative log10 of the single marker P-value of the LMM association statistic. The blue line represents the threshold for a suggestive association with P < 1.0 ×10^{-5}, the red line for significant association with P < 5.0 × 10^{-8}.

Family-based GWAS of JME

Genome-wide association analysis in JME families included 157 parent-offspring trios affected by JME, using 5,098,752 SNPs. Only one SNP reached the suggestive threshold (top five results shown in Table 3-8, Figure 3-9). The suggestive p-value was reported for chromosome 4q28.1 (rs313047, chr7:86192385, TDT P = 6.61×10^{-6}, OR[T] = 2.26, 95% CI 1.58 – 3.23) in a non-coding region, with no gene as close as 100 kb. Correction for genomic inflation was required and applied to all p-values ($\lambda = 1.026$, see QQ-plot in Suppl. 6.3.3).

Table 3-8 | Top-ranked GWAS TDT results for JME

SNP	CHR	POS	Allele A1/A2	Source	Info	TDT P	OR (95% CI)	Gene
rs313047*	4	127,626,494	T/C	Imp	0.98	6.61E-06	2.26 (1.58-3.23)	
rs12130711*	1	247,607,642	T/C	Imp	0.98	1.02E-05	0.46 (0.32-0.65)	NLRP3
rs66978760*	13	61,849,275	T/A	Imp	0.92	1.26E-05	5.14 (2.29-11.56)	
rs131024*	22	49,174,763	T/G	Imp	0.96	1.27E-05	0.45 (0.31-0.64)	FAM19A5(27.0k)
rs17719479	16	78,491,051	G/C	Imp	0.95	1.82E-05	0.30 (0.17-0.53)	WWOX

CHR: Chromosome, POS: Physical position NCBI built 37.3 hg19, Allele A1: Minor allele, Allele A2: Major allele, Imp: SNP imputed, Geno: SNP genotyped, Info: SNPtest info value, Gene: Gene nearby the associated SNP, distance in brackets. The top-ranked SNP was represented in regions with several single marker association signals exceeding TDT P < 1.0×10^{-5} (*).

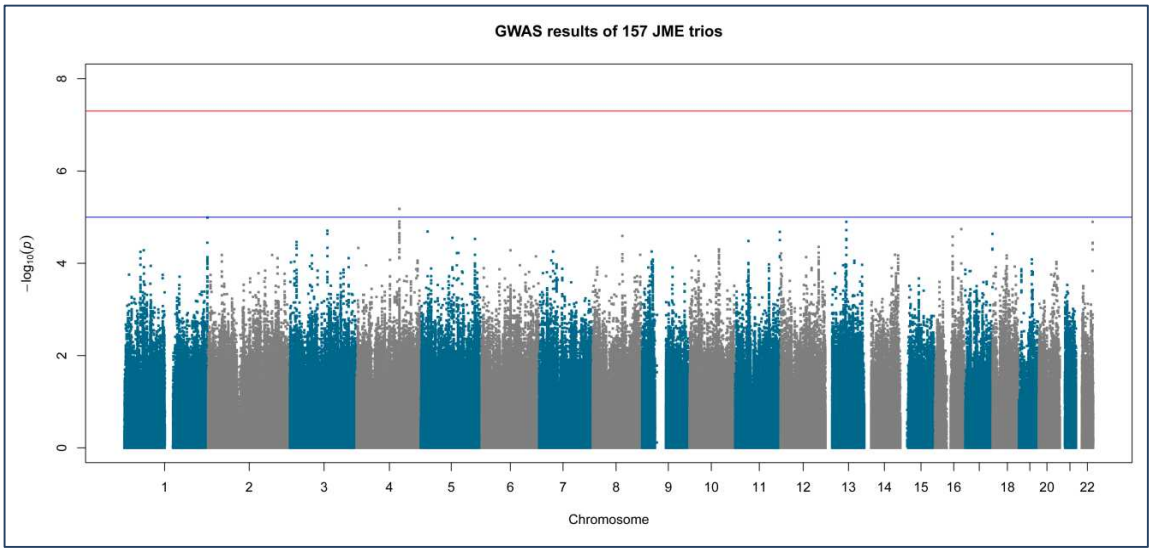


Figure 3-9 | Manhattan plot of genome-wide association in 157 JME trios. The x–axis shows the genomic position of SNP markers on chromosomes 1 to 22 (NCBI built 37.3, hg19), the y-axis shows the negative log10 of the single marker P-value of the TDT association statistic. The blue line represents the threshold for a suggestive association with P < 1.0×10^{-5}, the red line for significant association with P < 5.0×10^{-8}.

GWAS meta-analysis of JME

GWAS meta-analysis of 579 JME cases, 2,454 controls and 157 JME trios, was performed using 4,220,447 SNPs (Table 3-9, Figure 3-10). Five SNPs showed genome-wide significance, seven additional SNPs reached the threshold for suggestive association. The significant peak was observed on chromosome 1q43 (rs12059546, chr1:239970097, meta-analysis P = 2.27×10^{-08}, OR[G] = 1.53, 95% CI 1.33 – 1.78), within the gene *CHRM3*. All P-values were corrected for genomic inflation ($\lambda = 1.010$, see QQ-plot in Suppl. 6.3.3).

Table 3-9 | Genome-wide association meta-analysis results of $P < 10^{-5}$ in JME

			Allele				OR_{meta}	
SNP	**CHR**	**POS**	**A1/A2**	**LMM P**	**TDT P**	**Meta P**	**(95% CI)**	**Gene**
rs12059546*	1	239,970,097	G/A	1.97E-07	0.03	2.27E-08	1.53 (1.33-1.77)	CHRM3
rs72750824	9	112,380,046	G/C	3.26E-05	0.01	2.66E-06	1.66 (1.34-2.05)	PALM2(23.0k)
rs17669194	13	34,267,584	T/C	4.35E-06	0.49	5.03E-06	1.45 (1.23-1.70)	
rs845742	5	109,002,085	A/C	6.31E-05	0.02	7.53E-06	0.71 (0.62-0.80)	MAN2A1(23.1k)

CHR: Chromosome, POS: Physical position NCBI built 37.3 hg19, Allele A1: Minor allele, Allele A2: Major allele, Gene: Gene nearby the associated SNP, distance in brackets. The top-ranked SNP was represented in regions with several single marker association signals exceeding Meta $P < 1.0 \times 10^{-5}$ (*).

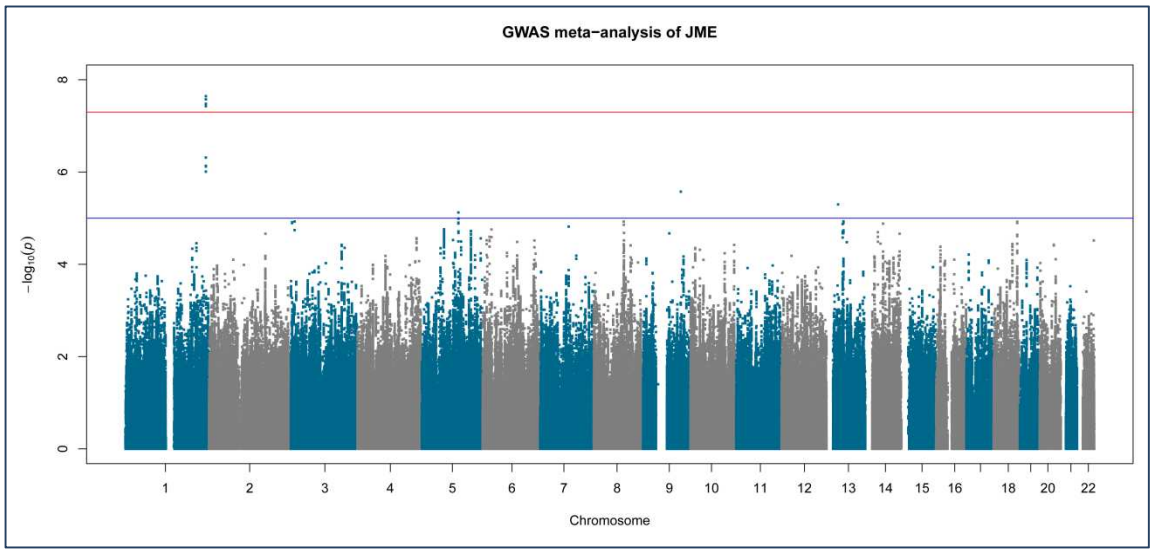


Figure 3-10 | Manhattan plot of genome-wide association meta-analysis in 579 JME cases, 2,454 controls and 157 JME trios. The x–axis shows the genomic position of SNP markers on chromosomes 1 to 22 (NCBI built 37.3, hg19), the y-axis shows the negative log10 of the meta P-value per SNP. The blue line represents the threshold for a suggestive association with $P \leq 1.0 \times 10^{-5}$, the red line for significant association with $p \leq 5.0 \times 10^{-8}$.

SNP markers achieving genome-wide significance were plotted with LocusZoom for regional view and to show LD of neighboring SNPs. The top-ranked SNP rs12059546 on chromosome 1q43 (Figure 3-11) was in tight LD with neighboring SNPs reaching significant and suggestive P-values. Several regions with increased recombination rate are close to the locus showing significance.

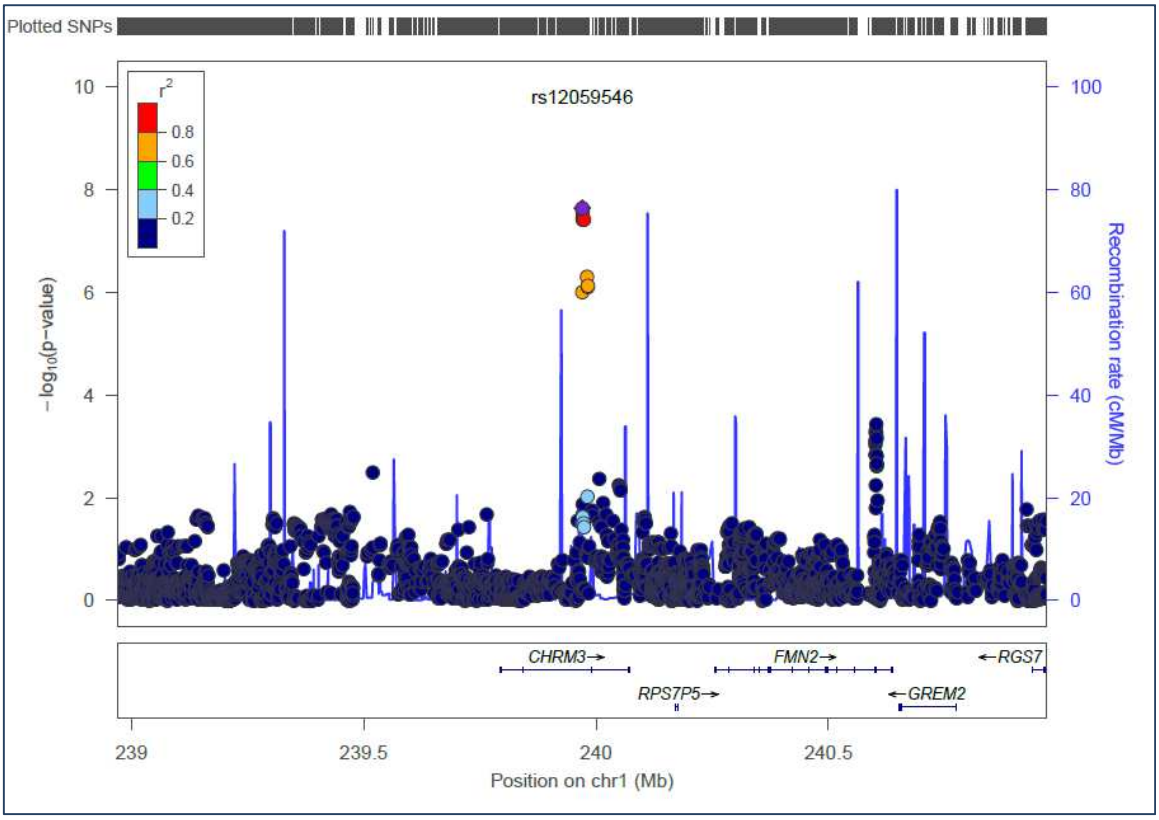


Figure 3-11 | Regional detail plot of the $-\log_{10}$ of LMM P-value of JME subgroup for SNP rs12059546 on chromosome 1q43, presenting genes and LD of neighboring SNPs close to detected association. SNP color code is complied with r^2 value referring to rs35577149, physical position in NCBI built 37.3, hg19.

3.2 Analysis of recurrent microdeletions associated with GGE

Six recurrent microdeletions that have previously been reported in neuropsychiatric disorders including autism, intellectual disability and schizophrenia, were evaluated for contributing to risk of GGE (Table 2-2). In total, a case-control sample cohort of 1,495 European GGE patients and 5,374 German controls in addition to 566 European parent-offspring trios affected by GGE were selected for microdeletion analysis. Parts of this study have previously been published (Helbig et al., 2009; de Kovel et al., 2010).

Recurrent microdeletions in GGE

In total, we detected deletions at the six candidate loci (1q21.1, 15q11.2, 15q13, 16p11.2, 16p13.11 and 22q11.2) in 56 (2.7%) out of 2,061 GGE cases and 22 (0.4%) out of 5,940 controls (OR = 7.51, 95% CI 4.50 – 12.95, χ^2 = 84.88, 1 degree of freedom (df), $p < 2.20 \times 10^{-16}$) (Figure 3-12, Table 3-10). Significant associations (nominal $P < 0.0083$) with GGE were found for the recurrent microdeletions at 15q11.2, 15q13, 16p11.2, 16p13.11 and 22q11.2 (Table 3-10).

Microdeletions at 1q21.1 were observed in 3 (0.1%) out of 2,061 GGE cases and in 2 (< 0.1%) out of 5,940 controls (OR = 4.33, 95% CI 0.50 – 51.82, Fisher's exact p = 0.11) (Table 3-10, Figure 3-12). An association of GGE with microdeletions at 15q11.2 was detected, which was found in 21 (1.0%) GGE cases and in 17 (0.3%) controls (OR = 3.59, 95% CI 1.80 – 7.25, χ^2 = 15.86, 1 df, $p = 1.12 \times 10^{-4}$) (Table 3-6, Figure 3-13). The strongest association of microdeletions with GGE was identified at site 15q13.3, with deletions in 14 (0.7%) GGE cases and none in controls (Fisher's exact $P = 5.48 \times 10^{-9}$) (Table 3-6, Figure 3-14). Likewise, association with GGE was observed at 16p13.11 with microdeletions in 11 (0.5%) GGE cases and in 2 (< 0.1%) controls (OR = 15.93, 95% CI 3.47 – 147.89, Fisher's exact $p = 1.48 \times 10^{-5}$) (Table 3-6, Figure 3-15). Microdeletions in 16p11.2 were observed in 5 (0.2%) GGE cases and in 1 (< 0.1%) control (OR = 11.54, 95% CI 1.41 – 567.06, Fisher's exact p = 0.02) (Table 3-6, Figure 3-16). At 22q11.2, microdeletions were identified in 3 (0.1%) out of 2,061 GGE cases and in none of 5,940 controls (Fisher's exact p = 0.02) (Table 3-6, Figure 3-17).

Table 3-10 | Recurrent microdeletions in 2,061 GGE patients and 5,940 controls

Chromosome region	GGE patients n = 1,495	Controls n = 5,374	GGE trio offsprings n=566	Trio pseudo-controls n=566	Combined GGE patients n = 2,061	Combined Controls n = 5,940	P-value, deKovel 2010 (1,234 cases, 3,022 controls)	P-value
1q21	2	1	1	1	3	2	-	0.11
15q11	13	16	8	1	21	17	4.2E-4	6.81×10^{-5}*
15q13	11	0	3	0	14	0	1.4E-5	5.48×10^{-9}
16p13	8	2	3	0	11	2	9.4E-3	1.48×10^{-5}
16p11	1	1	3	0	4	1	-	0.02
22q11	3	0	0	0	3	0	-	0.02
Total	38	20	18	2	56	22	1.1E-11	$< 2.2 \times 10^{-16}$*

Fisher's exact P-values two-sided; * χ^2 P-values with 1 degree of freedom

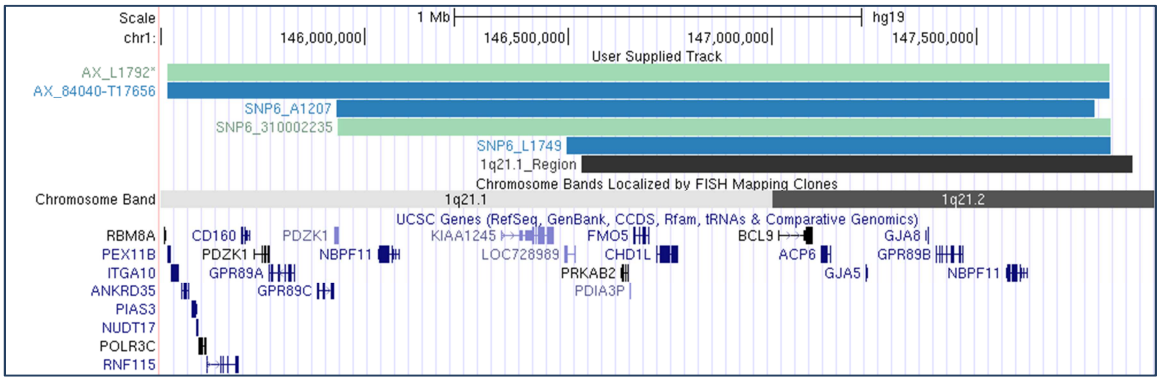


Figure 3-12 | Recurrent microdeletions at 1q21.1. Blue bars: Deletions in GGE patients, green bars: Deletions in controls and untransmitted deletions in trio pseudocontrols (*), grey bar: reported target region of the 1q21.1 microdeletion, SNP6: Sample genotyped with Affymetrix SNP6.0 array, AX: Sample genotyped with Affymetrix Axiom array. Produced with the University of California, Santa Cruz Genome Browser , positions in NCBI built 37.3, hg19 (http://www.genome.ucsc.edu).

Figure 3-13 | Recurrent microdeletions at 15q11.2. Blue bars: Deletions in GGE patients, green bars: Deletions in controls and untransmitted deletions in trio pseudocontrols (*), grey bar: reported target region of the 15q11.2 microdeletion, SNP6: Sample genotyped with Affymetrix SNP6.0 array, AX: Sample genotyped with Affymetrix Axiom array. Produced with the University of California, Santa Cruz Genome Browser, positions in NCBI built 37.3, hg19 (http://www.genome.ucsc.edu).

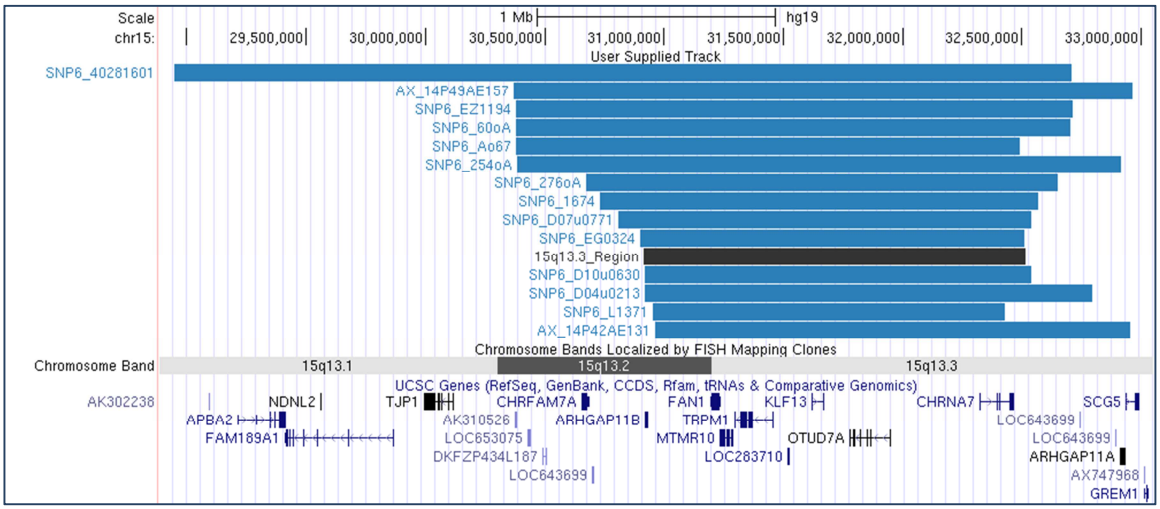


Figure 3-14 | Recurrent microdeletions at 15q13.3. Blue bars: Deletions in GGE patients, grey bar: reported target region of the 15q13.3 microdeletion, SNP6: Sample genotyped with Affymetrix SNP6.0 array, AX: Sample genotyped with Affymetrix Axiom array. Produced with the University of California, Santa Cruz Genome Browser, positions in NCBI built 37.3, hg19 (http://www.genome.ucsc.edu).

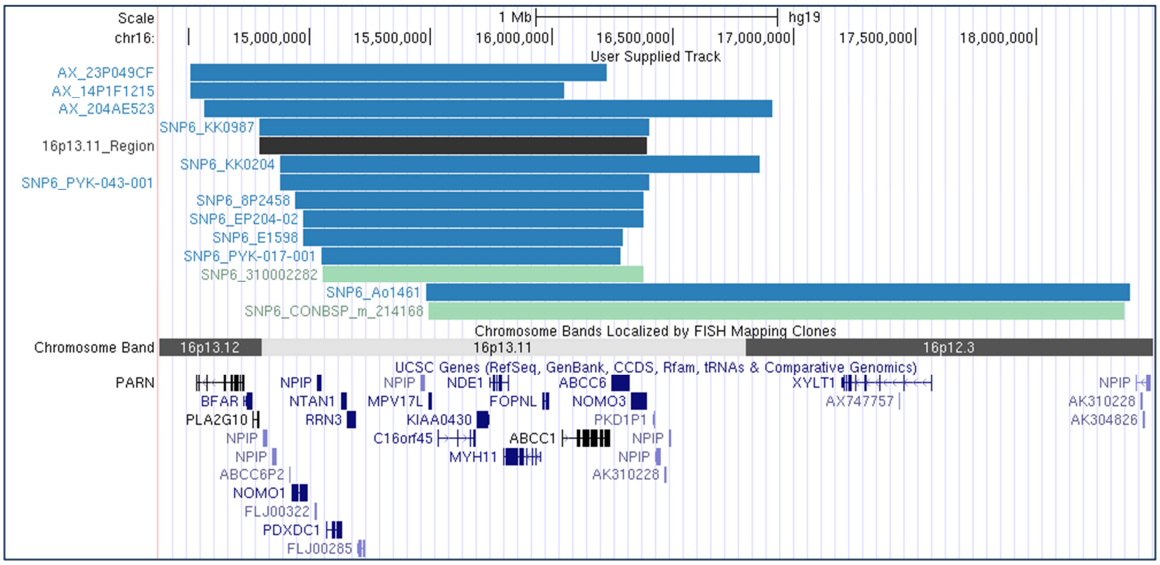


Figure 3-15 | Recurrent microdeletions at 16p13.11. Blue bars: Deletions in GGE patients, green bars: Deletions in controls and untransmitted deletions in trio pseudocontrols (*), grey bar: reported target region of the 16p13.11 microdeletion, SNP6: Sample genotyped with Affymetrix SNP6.0 array, AX: Sample genotyped with Affymetrix Axiom array. Produced with the University of California, Santa Cruz Genome Browser, positions in NCBI built 37.3, hg19 (http://www.genome.ucsc.edu).

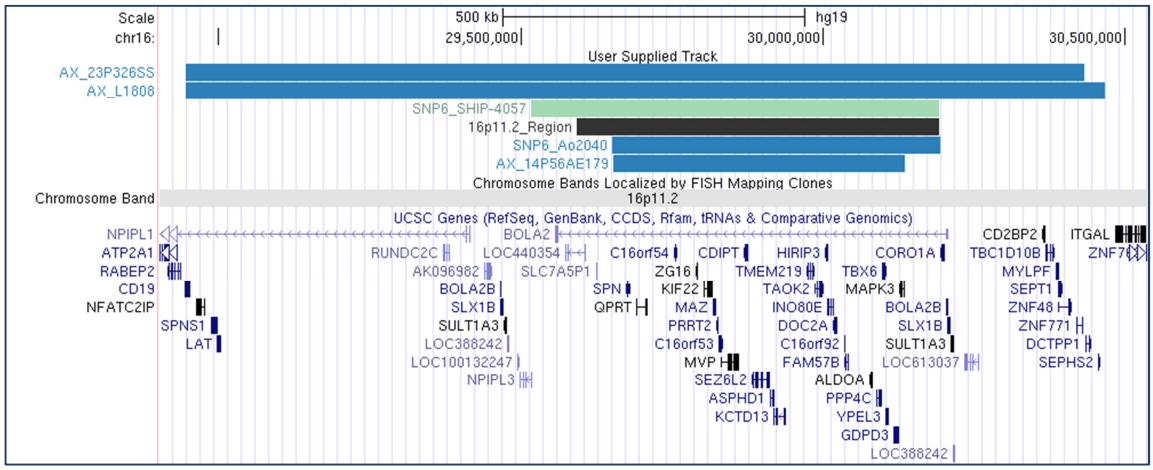


Figure 3-16 | Recurrent microdeletions at 16p11.2. Blue bars: Deletions in GGE patients, green bars: Deletions in controls and untransmitted deletions in trio pseudocontrols (*), grey bar: reported target region of the 16p1.2 microdeletion, SNP6: Sample genotyped with Affymetrix SNP6.0 array, AX: Sample genotyped with Affymetrix Axiom array. Produced with the University of California, Santa Cruz Genome Browser, positions in NCBI built 37.3, hg19 (http://www.genome.ucsc.edu).

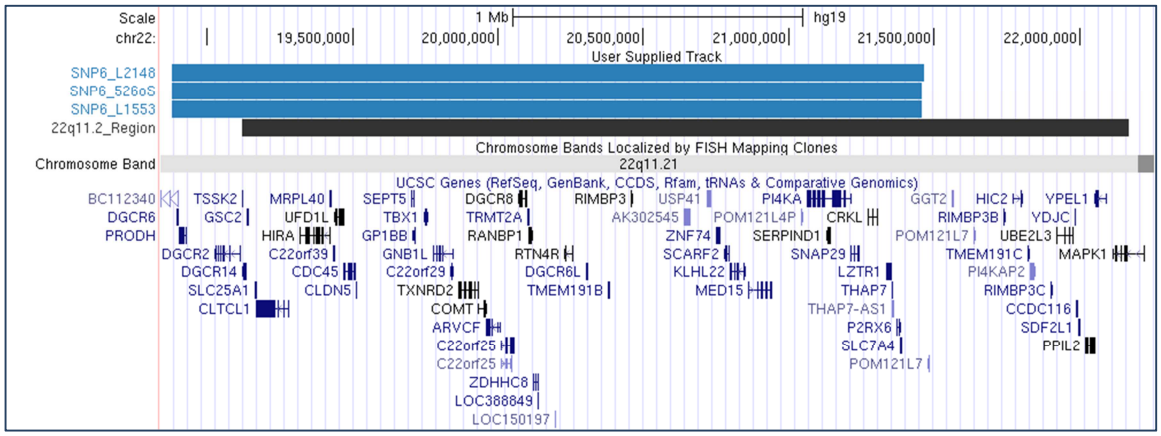


Figure 3-17 | Recurrent microdeletions at 22q11.2. Blue bars: Deletions in GGE patients, grey bar: reported target region of the 22q11.2 microdeletion, SNP6: Sample genotyped with Affymetrix SNP6.0 array, AX: Sample genotyped with Affymetrix Axiom array. Produced with the University of California, Santa Cruz Genome Browser, positions in NCBI built 37.3, hg19 (http://www.genome.ucsc.edu).

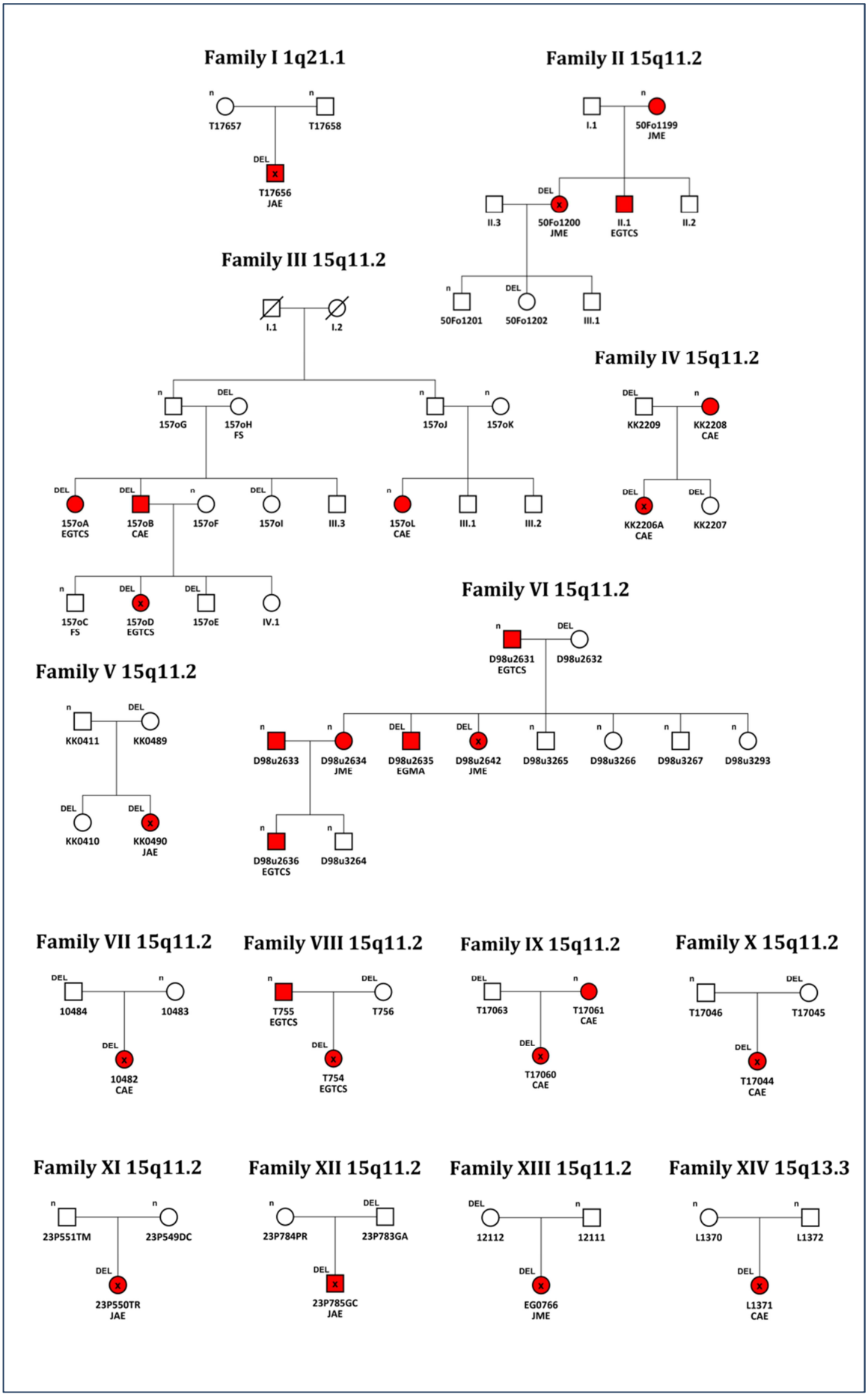
Familial segregation analysis

DNA samples from both parents were available from 31 out of 56 patients with identified microdeletions (Figure 3-20). For segregation analysis, all available family members were typed by quantitative PCR and/or aCGH. In summary, 20 out of 31 microdeletions were inherited, with 12 maternal and 9 paternal transmissions. Ten *de novo* deletions were identified in GGE patients (Figure 3-18, Suppl. 6-4). For 1q21.1 microdeletions, one family with a *de novo* microdeletion was available for segregation analysis (Figure 3-18). DNA from both parents was available for 12 out

of 22 GGE patients carrying a 15q11.2 microdeletion, with one *de novo*, maternal and paternal inheritance in five patients each (Figure 3-18). Two trios were observed with unaffected parents carrying 15q11.2 microdeletions (one mother and one father) that were not transmitted to the child affected with GGE. In 7 out of 14 patients with 15q13.3 microdeletions parental DNA was available. Paternal inheritance and maternal inheritances where found each in two out of eight transmissions, while *de novo* deletion events occurred in the remaining three families (Figure 3-18). DNA samples from both parents were available for 7 out of 11 families carrying 16p13.11 microdeletions. *De novo* microdeletions were observed in two GGE patients. Maternal transmission was identified in three and paternal transmission two out of eight patients with inherited microdeletions. Parental DNA for segregation analysis of 16p11.2 microdeletions was available from three families, showing two *de novo* events and one maternal transmission. The family of 22q11.2 microdeletion carrier was available for segregation analysis, displaying a *de novo* deletion in the affected child.

In summary, recurrent microdeletions are inherited at approximately equal rates (11 maternal and 8 paternal transmissions). Ten microdeletions arose *de novo*. Notably, microdeletions at 15q11.2 showed a lower rate of *de novo* deletions (five maternal and paternal transmissions, two *de novo* events). For six families with *de novo* mutations, sufficient genotype data was available to differentiate the original paternal and maternal segment of the *de novo* deletion. In three families the maternal strand was deleted, three families showed deletion of the paternal strand.

Overall families, 28 GGE deletion carriers shared their deletions with 4 affected and 24 unaffected first-degree relatives. Eight first-degree relatives did not carry the deletion, although affected by GGE (Figure 3-20). These eight affected relatives without a deletion were all found in families exhibiting the 15q11.2 deletion. In 18 out of 32 families, the transmitting parents were clinically unaffected (15q11.2 n = 9; 15q13.3: n = 3; 16p13.11: n = 5; 16p11.2: n= 1). One father carrying a 15q13.3 deletion was affected by GGE. Two fathers with 15q11.2 microdeletions were affected by CAE, one in addition to FS. For one sample with 16p13.11 microdeletion (204AE523, family XXIX), data of two additional siblings were available. All three are affected with GGE and inherited the microdeletion from their unaffected mother.


Family I 1q21.1
T17657
T17658
T17656
JAE
Family II 15q11.2
I.1
50Fo1199
JME
II.3
50Fo1200
JME
II.1
EGTCS
II.2
50Fo1201
50Fo1202
III.1
Family III 15q11.2
I.1
I.2
157oG
157oH
FS
157oJ
157oK
157oA
EGTCS
157oB
CAE
157oF
157oI
III.3
157oL
CAE
III.1
III.2
157oC
FS
157oD
EGTCS
157oE
IV.1
Family IV 15q11.2
KK2209
KK2208
CAE
KK2206A
CAE
KK2207
Family V 15q11.2
KK0411
KK0489
KK0410
KK0490
JAE
Family VI 15q11.2
D98u2631
EGTCS
D98u2632
D98u2633
D98u2634
JME
D98u2635
EGMA
D98u2642
JME
D98u3265
D98u3266
D98u3267
D98u3293
D98u2636
EGTCS
D98u3264
Family VII 15q11.2
10484
10483
10482
CAE
Family VIII 15q11.2
T755
EGTCS
T756
T754
EGTCS
Family IX 15q11.2
T17063
T17061
CAE
T17060
CAE
Family X 15q11.2
T17046
T17045
T17044
CAE
Family XI 15q11.2
23P551TM
23P549DC
23P550TR
JAE
Family XII 15q11.2
23P784PR
23P783GA
23P785GC
JAE
Family XIII 15q11.2
12112
12111
EG0766
JME
Family XIV 15q13.3
L1370
L1372
L1371
CAE

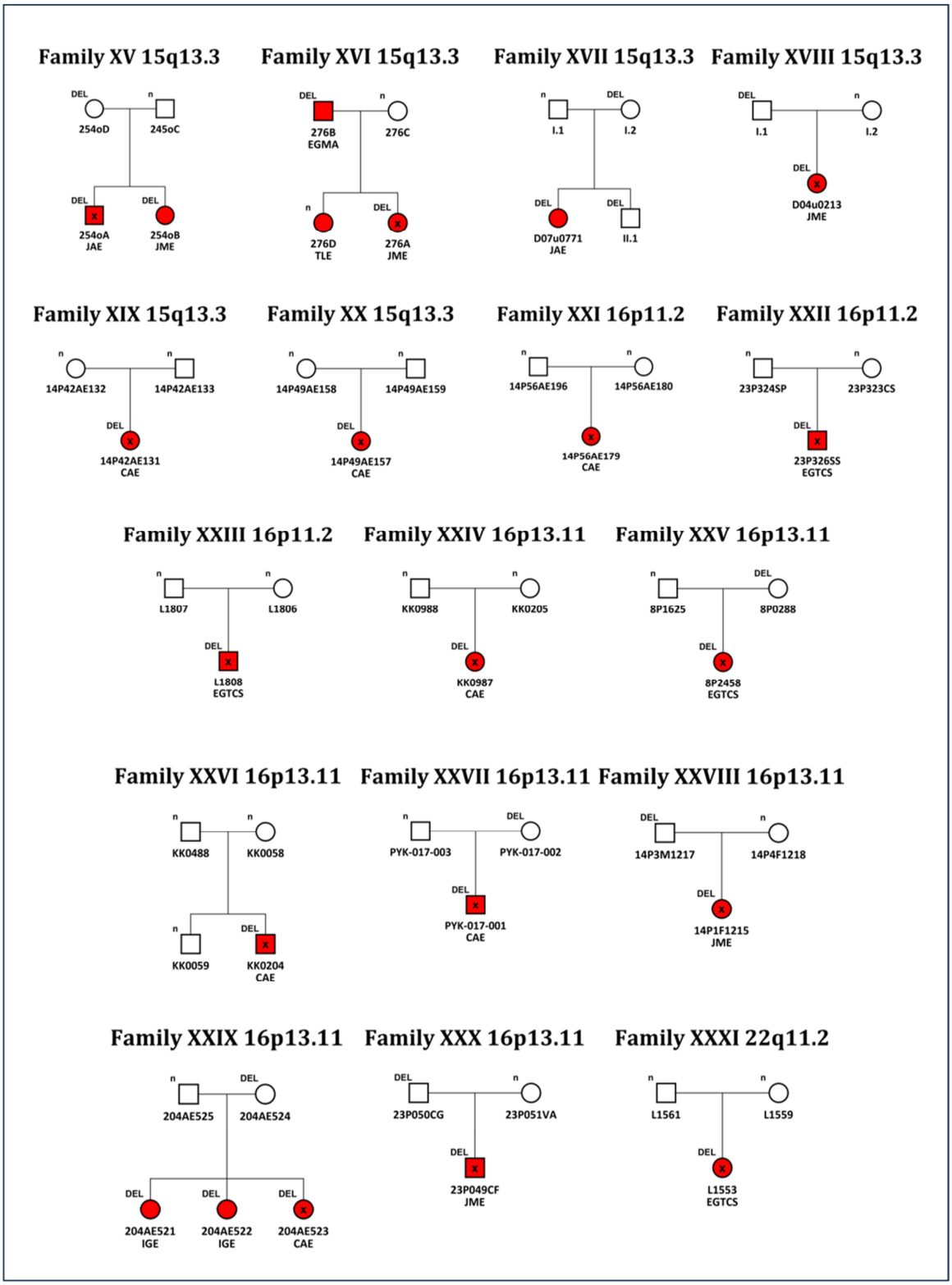


Figure 3-18 | Familial segregation of the microdeletions at 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.11 and 22q11.2. Red symbols: individuals affected by IGE, a "x" denote the index-IGE patient. FS: febrile seizure, CAE: childhood absence epilepsy, JAE: juvenile absence epilepsy, JME: juvenile myoclonic epilepsy, EGTCS: epilepsy with generalized tonic–clonic seizures alone, EGMA= epilepsy with generalized tonic–clonic seizures on awakening, TLE: temporal lobe epilepsy, copy number state: n: normal/two copies, DEL: deletion carrier.

In general, cosegregation of the microdeletions was incomplete and did not follow an autosomal dominant pattern of inheritance, notably in the three largest families with 15q11.2 microdeletions (Figure 3-20; families II, III and VI).

Penetrance estimations

Penetrance calculations were performed as described before in CNVs associated with schizophrenia (Vassos et al., 2010). Confidence Intervals were calculated by the Clopper–Pearson exact tail area method (Rosenfeld et al., 2012). In our study, no deletions in controls were detected for microdeletions at 15q13.3 and 22q11.2. Incidence of 15q13.3 microdeletions in controls was acquired from combined data previously published with GGE (Dibbens et al., 2009), while no data for microdeletions in controls was available for 22q11.2 microdeletions. Penetrance estimations for the analyzed microdeletions ranged from 1.06% (95% CI 0.003 – 0.038) at 15q11.2, to 12.49% (95% CI 0.070 – 0.201) at 15q13.3 (Table 3-11).

Table 3-11 | Penetrance estimations of recurrent microdeletions

Chr. Region	Position	Cases	Controls	Penetrance %	95% CI
1q21.1	146.5-147.9	3/2061	2/5940	1.28	0.003-0.038
15q11.2	22.75-23.2	21/2061	17/5940	1.06	0.007-0.017
15q13.3	30.9-32.5	14/2061	8/50115*	12.49	0.070-0.201
16p11.2	29.6-30.2	4/2061	1/5940	3.34	0.009-0.082
16p13.11	14.8-16.4	12/2061	2/5940	4.93	0.0256-0.084
22q11.2	19.2-22.2	1/2061	0/5940	-	-

* 15q13.3 microdeletions in controls from Dibbens et al., 2009; Position in NCBI built 37.3, hg 19.

Breakpoint validation

The described recurrent microdeletions are reported to be flanked by highly homologous segmental duplications, promoting NAHR and therefore genomic rearrangements between reputed breakpoints. As breakpoint-estimation is difficult in regions with highly repetitive sequences, may vary between array platforms and due to limitations of calling algorithms, we used a customized oligonucleotide microarray to refine the breakpoints of the identified microdeletions (Suppl. 6-4). Breakpoints and microdeletion sizes estimated with aCGH were consistent with those maintained with the Affymetrix SNP6.0 array.

Breakpoint positions from aCGH analysis were available for two samples genotyped with the Affymetrix Axiom® array, 276A and 157oB. Start and end point positions of detected microdeletions and LOH screening estimates differed from approximately 250 kb to 670 kb. The LOH region at 15q11.2 in 157oB was about 400 kb smaller than the microdeletion detected with aCGH, the LOH region at 15q13.3 in 276A about 950kb smaller. Consequently, LOH screening as a prediction method for microdeletions is suitable for a defined core region, but lacks for precision to estimate limits of large microdeletions.

Distribution of phenotypes among deletion carriers

The distribution of microdeletions among GGE syndromes and sex of deletion carriers (CAE/GAE 52.5%, JME 28.8%, GTCS 18.6%, males 32.2%, females 67.8%) was equivalent to the distribution of sex and phenotypes in the complete sample cohort (CAE/GAE 48.1%, JME 35.8%, GTCS 16.1%, males 38.2%, females 61.8%). We observed no other feature, neither the dominance of a particular seizure type, nor the shift in age-of-onset, nor other neuropsychiatric disorders in family members carrying a deletion.

3.3 Analysis of exon-ablating microdeletions associated with GGE

3.3.1 Exon-ablating microdeletions in *NRXN1*

In this candidate gene study, 1,569 unrelated GGE cases of European ancestry and 6,201 German controls were analyzed for exon-disrupting/removing microdeletions in *NRXN1*. Results of this study has previously been published (Møller et al., 2013).

A significant association of exon-ablating microdeletions was observed in 5 (0.3%) of 1,569 patients with GGE versus 2 (0.03%) of 6,201 control samples (P = 0.0049; OR = 9.91, 95% CI 1.92 – 51.12) (Figure 3-19, Suppl. 6.5). All five microdeletions in patients occurred in the 5´-terminal region of the genomic sequence of *NRXN1*; one including the promoter region and exon one, two including the promoter region and exons one and two, two microdeletions only including exon 2. Both exonic microdeletions in controls only included exon 2. Two patients with exonic microdeletions were affected by CAE; one by GAE, one by JME and one by EGTCS alone.

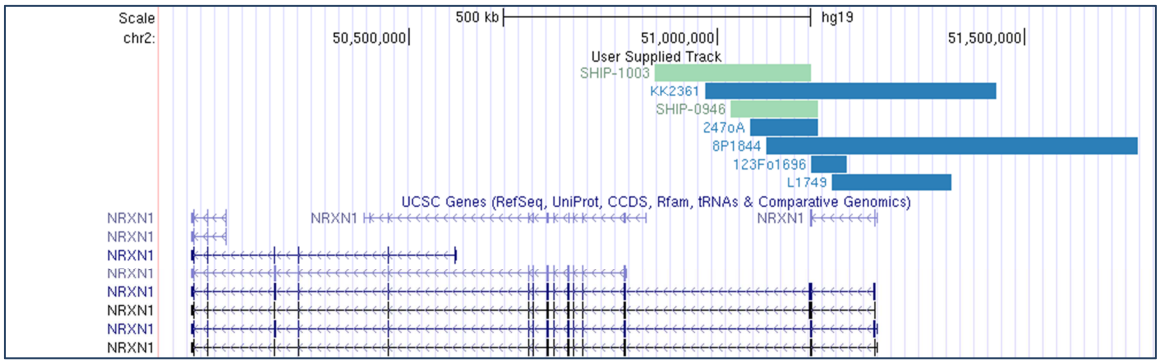


Figure 3-19 | Genomic positions of the seven exon-ablating *NRXN1* deletions. Deletions in cases are represented by blue bars, deletions in controls by green bars. The figure was produced with the UCSC Genome Browser (http://www.genome.ucsc.edu), built hg19 based on the genomic positions from the Affymetrix Genome-Wide Human SNP Array 6.0

Family I was of German ancestry with three affected family members carrying a 200 kb *NRXN1* deletion (chr2:51,185,310-51,379,597, hg19) and the unaffected father (Figure 3-20). Both children (index patient L1748 and his sister) had CAE with FS. Additionally, the male patient showed moderate learning disability. The mother experienced episodes of short unresponsiveness in her childhood. Albeit no further diagnostics and no therapy were applied, we assess those events as a possible CAE.

Family II was of German ancestry and consisted of three healthy and two children affected by GGE (Figure 3-20). Their mother was healthy and the father had type-1 diabetes. A 480 kb *NRXN1* deletion (chr2:50,979,977-51,453,231, hg19) was detected in all three affected children and in the father unaffected with neurologic disease. The index patient (KK2361) and her older sister were affected with CAE. Furthermore, the sister and one brother showed moderate ID, although the brother did not show any seizures.

Family III was of Danish ancestry and consisted of a female with IAE (8P1844) and her healthy parents and sister; DNA of an additional unaffected brother was not available (Figure 3-20). A 600 kb *NRXN1* deletion (chr2:51,080,429-51,682,854, hg19) was transmitted from the healthy father to the affected patient but not to her unaffected sister. A paternal half-sister had schizophrenia and Attention Deficit/Hyperactivity Disorder (ADHD), but was not available for testing.

Family IV was a multiplex GGE family of German ancestry including three affected individuals, two sisters and their father (Figure 3-20). The mother had no history of epileptic seizures and her EEG was normal. One clinically unaffected brother and

one unaffected sister were not examined. The index patient (1696) was carrying a 60-kb deletion (chr2:51,152,019-51,209,823, hg19) that occurred *de novo* in the proband and is affected by JME. Her sister is exhibiting JME as well. The father had partial epilepsy starting at the age of 57 years.

Family V consisted of two affected siblings and their unaffected parents of Turkish origin (Figure 3-20). The proband (247oA) had an EGTCS with also photosensitive seizures and is carrying a *de novo* 110kb *NRXN1* deletion (chr2:51,054,002-51,163,990, hg19). His affected brother also had EGTCS, but without microdeletions in *NRXN1*.

Segregation analysis was performed in all families of the five index cases carrying an exonic microdeletion in *NRXN1* (Figure 3-20). We observed four multiplex families with more than one affected individual with GGE. In total, we observed eleven deletion carriers within all five families, seven of those with GGE. Deletions have

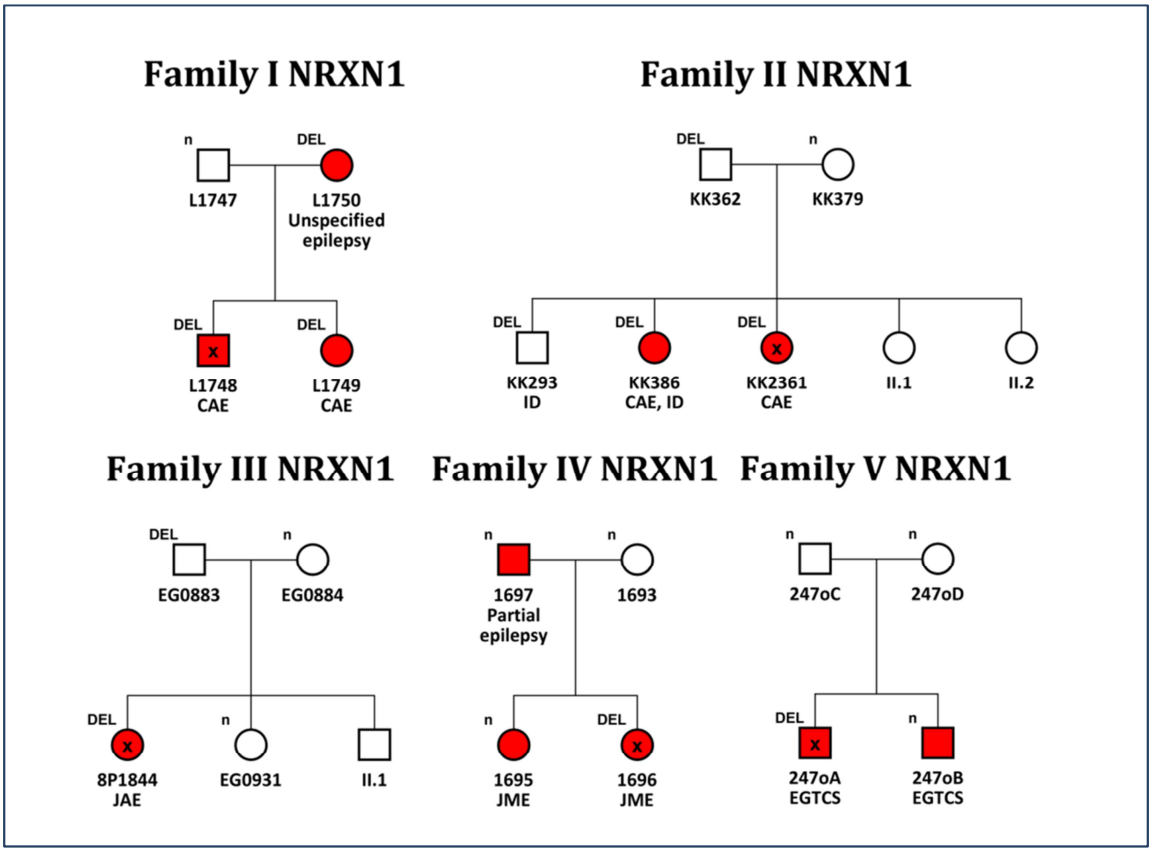


Figure 3-20 | Familial segregation of exon-ablating *NRXN1* deletions. Red symbols indicate persons affected by epilepsy, a "x" denote the index-IGE patient. CAE: childhood absences epilepsy; JAE: juvenile absence epilepsy; JME: juvenile myoclonic epilepsy; EGTCS: epilepsy with generalized tonic–clonic seizures alone; ID, intellectual disabilities; copy number state: n = normal/two copies; DEL = deletion carrier.

occurred *de novo* in two patients (families IV and V). One was inherited from an affected mother (family I), and two index patients inherited the deletion from an unaffected father (families II and III).

In summary, a *NRXN1* deletion was found in families I, II, and III in all affected family members, including one member with ID only. Since deletions were also present in two unaffected family members, penetrance is incomplete with an estimated probability of 78%. In families IV and V, the deletions are *de novo* mutations. Notably, affected family members without microdeletions in *NRXN1* are present in those families. CNV penetrance based on Vassos et al., 2012 was estimated at 2.9 % (95% CI 0.001 – 0.007).

3.3.2 Exon-ablating microdeletions in *RBFOX1*

In this candidate gene study, 1,408 unrelated GGE cases of European ancestry and 2,256 German controls were analyzed for deletions spanning or disrupting the genomic sequence of *RBFOX1*. Results of this study has previously been published (Lal et al., 2013).

Exonic microdeletions were present in 5 (0.4%) of 1,408 GGE cases and none in 2,256 controls (Fisher's exact P = 0.0083). No exonic deletions were detected in controls. Deletions were 68 to 896 kb in size and all located in the 5'-terminal region of *RBFOX1*, encompassing the untranslated exons 1 - 4 (Figure 3-21, Suppl. 6.6). All by array identified CNVs in GGE patients were hemizygous and successfully validated by qPCR, DNA samples of the control subjects were not available for validation.

From four out of five patients, DNA of family members was available for segregation analysis (Figure 3-22). In total, copy number state of the *RBFOX1* was assessed in 20 family members by qPCR. Twelve family members carried an exonic *RBFOX1* microdeletion (6 females, 6 males). All *RBFOX1* deletions identified in the GGE index patients were inherited, no *de novo* events occurred. The deletions were transmitted five times maternally and one time paternally. Seven out of 12 deletion carriers were affected by GGE; five deletion carriers were clinically unaffected. Seven out of nine GGE patients investigated carried an exonic *RBFOX1* deletion. Five of the seven

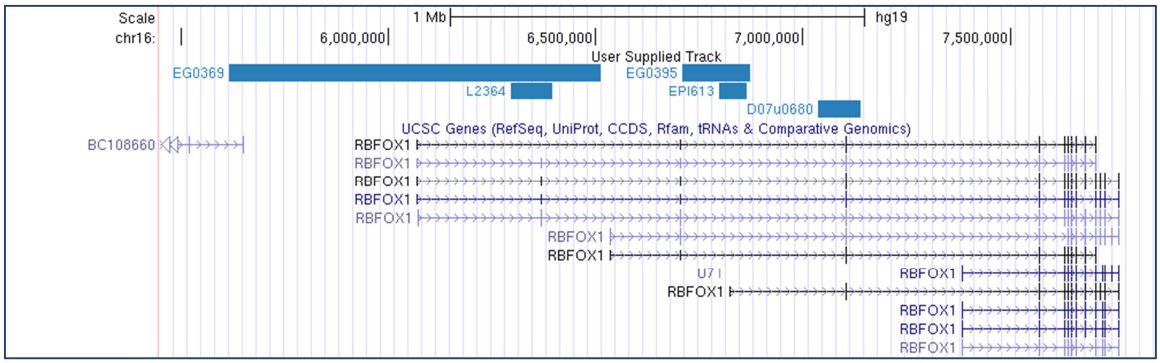


Figure 3-21 | Genomic positions of exon-ablating deletions in *RBFOX1*. Deletions in cases are represented by blue bars, deletions in controls by green bars. The figure was produced with the UCSC Genome Browser (http://www.genome.ucsc.edu), built hg19 based on the genomic positions from the Affymetrix Genome-Wide Human SNP Array 6.0.

deletion carriers were affected by typical absence epilepsies, and two carriers were affected by JME.

In family I, exonic microdeletions in *RBFOX1* were detected in all investigated family members with GGE. The index patient EG0369 was affected by CAE, delayed speech and learning disability. The deletion was transmitted maternally over two generations (mother with CAE and grandmother unaffected). DNA of an affected grand-grandfather was not available. The uncle of the index patient (EG0340) was affected by JME with absence seizures and GTCS, and was struck by unexpected sudden death at the age of 28.

The exonic *RBFOX1* microdeletions in family II were detected in the affected mother, the index patient and his unaffected brother. The father was unaffected and did not carry the deletion. Both mother and son are affected by IAE, while the index patients experienced FS and GTCS in addition.

Family III consisted of an affected mother and two affected sons; the father and one additional sister are unaffected. The *RBFOX1* deletion was transmitted by the unaffected father to the index patient (D07u680), affected by JME with GTCS. Magnetic resonance imaging scan showed structural abnormalities of the brain for bifrontal lesions in the patient due to a traumatic brain contusion occurring 16 years after the onset of the GGE. The mother and his brother were both affected by EGTCS but none of them carried a *RBFOX1* microdeletion. The brother showed a pervasive developmental disorder, which is part of the diagnostic group of autism spectrum

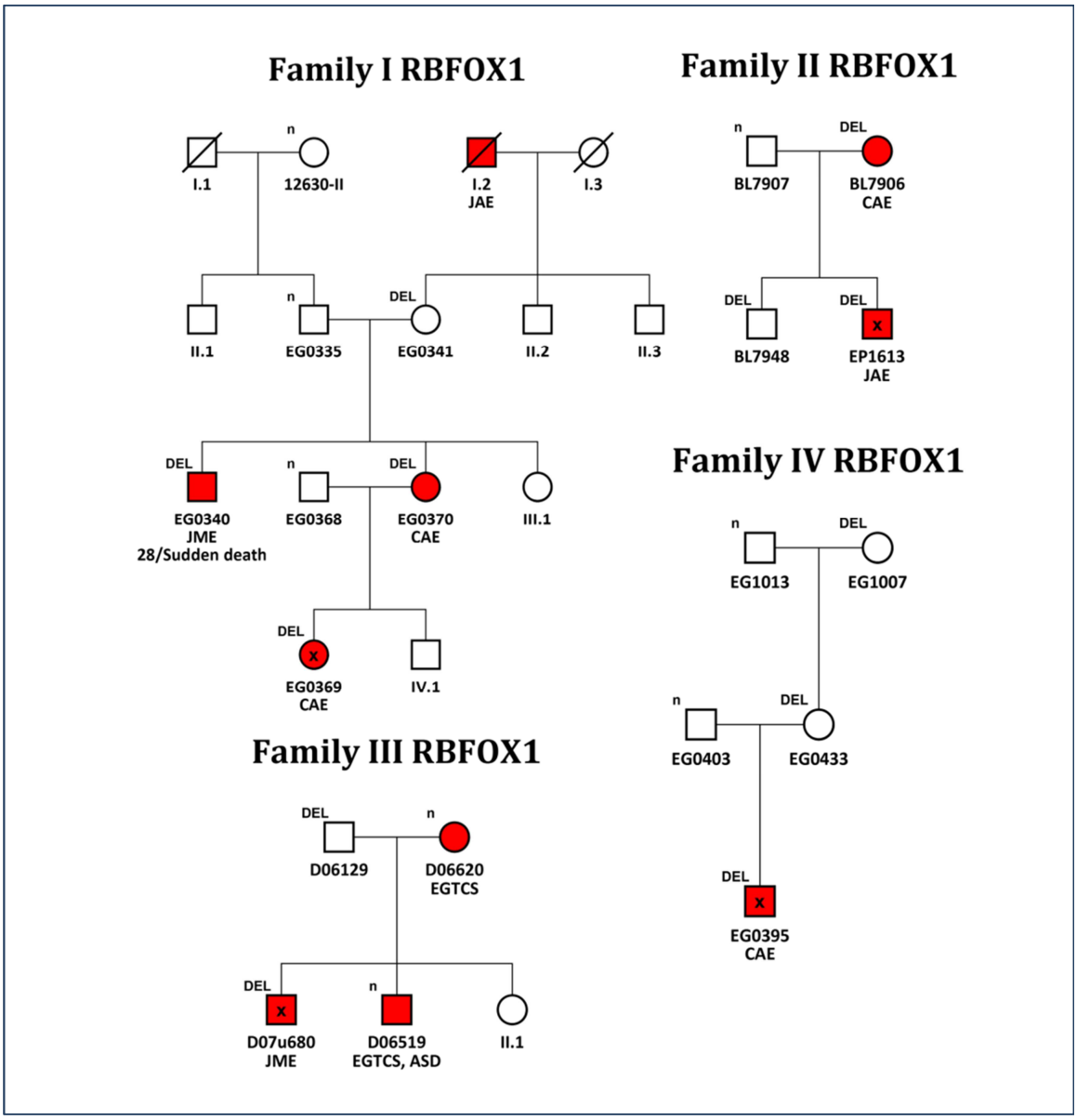


Figure 3-22 | Familial segregation of the exon-ablating *RBFOX1* deletions. Red symbols: individuals affected by GGE, a "x" denote the index-IGE patient. Crossed individuals: Deceased, CAE: childhood absence epilepsy; JME: juvenile myoclonic epilepsy, EGTCS: epilepsy with generalized tonic–clonic seizures alone, FS: febrile seizures, ASD: autism spectrum disorder, copy number state: n = normal/two copies, DEL: deletion carrier.

disorders (ASD). DNA of unaffected the sister was not available for testing. Vision impairment due to a strong myopia was present in the mother and all three siblings.

In family IV, an exonic *RBFOX1* microdeletion was transmitted maternally over three generations, with the male index patient (EG0395) as the only affected family member with CAE plus GTCS.

Exonic microdeletions in *RBFOX1* were found in seven affected and five unaffected family members, resulting in an estimated familial penetrance of 58%. Penetrance

estimation of exonic microdeletions based on Vassos et al., 2012 was not possible, because of insufficient control data.

In summary, the present candidate gene CNV study of the neuron-specific splicing regulator gene *RBFOX1* suggests that microdeletions affecting the untranslated 5'-terminal *RBFOX1* exons increase risk of common GGE syndromes. The present findings warrant further studies to replicate an involvement of *RBFOX1* in the genetic predisposition of GGE syndromes and other common neurodevelopmental disorders and to elucidate the pathogenic mechanisms of epileptogenesis resulting from *RBFOX1*-mediated alterations of the splicing process of neuronal genes.

4 Discussion

4.1 Genome-wide association and meta-analysis

The present GWAS meta-analysis displayed significant association at 2p16.1 with GGE and at 1q43 with JME (figures 3-3, 3-10). Despite a respectable size of the study sample, power calculations estimated a power of only 80% for meta-analysis to detect risk alleles with OR ≥ 1.34 in GGE (1,523 cases, 2545 controls and 566 trios), OR ≥ 1.47 in GAE (705 cases, 2545 controls and 317 trios) and with OR ≥ 1.56 in JME (579 cases, 2545 controls and 157 trios). Although power calculations for case-control GWAS alone showed almost similar results (GGE OR ≥ 1.40, GAE OR ≥ 1.57, and JME OR ≥ 1.64), the detectable effect sizes of risk factors in trio-based analyses only were considerably higher (OR ≥ 1.87 in GGE, OR ≥ 2.35 in GAE, and OR ≥ 3.60 in JME trios). In comparison to the case-control sample cohort, the trio sample cohort was relatively small, contributing only 566 cases (for GGE; 317 for GAE and 157 for JME respectively) to the GWAS meta-analysis. Thus, statistical power was only increased marginally by combining both data sets, especially for the JME subgroup.

We successfully strengthened the statistical evidence of previously reported associations at 1q43, 2p16.1, and 2q22.3 (EPICURE Consortium et al., 2012a), supporting their role as susceptibility factors for GGE. Further replication is required for novel association findings at 8q12.2, particularly with regard to the lack of a plausible candidate gene for epileptogenesis at this site. A locus at 17q21.32 previously reported to be associated with GGE as well as a locus at 2q24.3 with former suggestive evidence for association with GGE (EPICURE Consortium et al., 2012a) could not be confirmed in our meta-analysis. This demonstrates the dilemma of high variability in the outcome of GWAS, even if a part of the study cohort is overlapping between analyses.

4.1.1 Significant association with GGE at 2p16.1

Significant association was detected for GGE in the chromosomal region 2p16.1 (Figure 3-3). The closest genes are *VRK2* (vaccinia related kinase 2; ~340 kb away) and *FANCL* (Fanconi anemia, complementation group L; ~ 455 kb away). We therefore strengthened recently published results including our case-control cohort,

reporting association at 2p16.1 with GGE (rs13026414, P = 2.5 × 10^{-9}, OR = 0.81) (EPICURE Consortium et al., 2012a). Previously, SNP rs2312147 close to *VRK2* was associated with schizophrenia in a European (P = 1.9 × 10^{-9}, OR = 1.09) (Steinberg et al., 2011) and an Asian-European meta-analysis (P = 2.21 × 10^{-8}, OR = 1.11) (Li et al., 2012). Regarding our own significant association signal at rs35577149, SNP rs2312147 is approximately 300 kb away and shows modest LD (r^2 = 0.195, D' = 0.660). *VRK2* lies within the region of the 2p16.1-p15 deletion syndrome (OMIM #612513), causing developmental delay and/or mental retardation, autistic features, dysmorphic features, microcephaly, febrile seizures and intractable seizures with neonatal onset (Chandler et al., 2006; Rajcan-Separovic et al., 2007; Chabchoub et al., 2008; Félix et al., 2010). Consequently, genetic variation at the *VRK2* locus may contribute to risk for GGE.

4.1.2 Significant association with JME at 1q43

Significant association was observed at chromosomal region 1q43, within the intronic region of the gene *CHRM3* (cholinergic receptor, muscarinic 3; Figure 3-10). Accordingly, association has been reported previously at 1q43 (rs12059546, P = 2.3 × 10^{-7}, OR = 1.41) (EPICURE Consortium et al., 2012a). Despite mutations of genes encoding neuronal nicotinic acetylcholine receptor subunits are known to cause autosomal dominant nocturnal frontal lobe epilepsy (*CHRNA4, CHRNB2*) (Steinlein and Bertrand, 2010), no associations were found of genes encoding muscarinic acetylcholine receptors with epilepsy. *CHRM3* confers differential cholinergic modulation to neurochemically distinct hippocampal basket cell subtypes (Rio et al., 2010). After surgically removing sclerotic hippocampi of patients with refractory mesial temporal lobe epilepsy, hippocampal CA3 transcriptome signature revealed a cell-type specific expression of the M3 muscarinic acetylcholine receptor in distinct subtypes of hippocampal interneurons, providing a molecular mechanism for a differential cholinergic modulation of hippocampal circuitry (Ding et al., 2010; Bando et al., 2011). This may have an influence on synchronization and excitability of thalamocortical circuits and thereby seizure susceptibility, although M3 receptor knockout mice did not show an increase in pilocarpine-induced seizure activity (Bymaster et al., 2003). Most recently, a case study about a patient with ASD and a 473 kb microdeletion affecting the *CHRM3* gene has been reported (Petersen et al.,

2013). Similarly, a microdeletion at 1q43 in a patient with mental retardation and short stature with three genes including *CHRM3* was reported (Perrone et al., 2012). Those studies strongly suggest a possible link between M3-muscarinic receptor pathophysiology and a neurodevelopmental phenotype.

4.1.3 Suggestive association with GGE at 8q12.2

GWAS meta-analysis in GGE revealed evidence for suggestive association at chromosomal region 8q21.2. To our knowledge, this locus has not been reported to be associated with GGE or other epilepsies before. Notably, this locus is the only top-hit from TDT analyses reaching suggestive level after meta-analysis. The association signal lies within the intronic region of *FAM110B* (family with sequence similarity 110, member b). FAM110 proteins have been localized to centrosomes and accumulate at the microtubule organization center in interphase and at spindle poles in mitosis. Moreover, ectopic expression of FAM110B and FAM110C proteins impaired cell cycle progression in G1 phase (Hauge et al., 2007). Given the fact that the trio has a low power to detect low-risk effects and despite of potent candidate genes, this result requires further consideration.

4.1.4 Suggestive association with GAE at 2q22.3

Suggestive evidence for association with GAE was found in an intergenic region at 2q22.2 (Figure 3-6), a locus that showed significant association in a previous study (rs10496964, $P = 9.1 \times 10^{-9}$, OR = 0.68) (EPICURE Consortium et al., 2012a). The suggestive association signal is approximately 93 kb away from *ZEB2* (zinc finger e box-binding homeobox 2). Mutations in *ZEB2* are responsible for the Mowat–Wilson syndrome (OMIM #235730), characterized by typical face, moderate-to-severe mental retardation, Hirschsprung disease, multiple congenital anomalies and a prevalence for epilepsy of about 70-75% (Garavelli et al., 2009; Cordelli et al., 2013). Recent studies in mice have revealed, that *ZEB2* expression is required for generation and migration of cortical interneurons (McKinsey et al., 2013; van den Berghe et al., 2013). These interneurons are involved in phasing and synchronizing of neuronal activity (Cobb et al., 1995; Markram et al., 2004), suggesting a pathogenic mechanism that may operate in both epilepsy and Mowat–Wilson syndrome.

4.2 Recurrent microdeletions in GGE

In the present study, we investigated the impact of six large recurrent microdeletions (at 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.11 and 22q11.2), previously associated with neuropsychiatric disorders, to risk for common GGE syndromes (Ullmann et al., 2007; Sebat et al., 2007; Bassett et al., 2008; Brunetti-Pierri et al., 2008; Sharp et al., 2008; Stefansson et al., 2008; Weiss et al., 2008; Kumar et al., 2008; Marshall et al., 2008; Mefford et al., 2008; Schizophrenia Consortium, 2008; Hannes et al., 2009; Kirov et al., 2009; Need et al., 2009). Combined analysis of all six recurrent microdeletions showed a significant enrichment of microdeletions in GGE patients on the contrary to unaffected controls and trio parents ($P < 2.20 \times 10^{-16}$). Microdeletions at 15q11.2, 15q13.3 and 16p13.11 are significantly associated with GGE (15q11.2: $P = 6.81 \times 10^{-5}$, 15q13.3: $P = 5.48 \times 10^{-9}$, 16p13.11 = 4.42×10^{-6}). Nominal P-values of recurrent microdeletions at 16p11.2 and 22q11.2 (both P-values = 0.02) would present further association, but after correction for multiple testing, significance level for association is not achieved ($P < 0.0083$). Still, our findings strengthen previous associations findings (de Kovel et al., 2010).

The reported microdeletions comprise a number of promising candidate genes (Table 2-2). The gene *HYDIN2* (chr1:146,310,551-146,334,209; hydrocephalus-inducing, mouse, homolog of, 2) in the region of the 1q21.1 microdeletion is a homologue of *HYDIN*, which is expressed in the mouse brain in the developing choroid plexus. A frameshift mutation was reported to cause hydrocephalus in mice (Davy and Robinson, 2003). Accordingly, Brunetti-Pierri et al. proposed *HYDIN2* as the most potent candidate gene in 1q21.2 microdeletions and microduplications (Brunetti-Pierri et al., 2008). The 15q11.2 microdeletion region includes the gene *CYFIP1* (chr15: 15:22,892,683-23,003,602;cytoplasmic *FMRP*-interacting protein 1), that interacts with fragile X mental retardation protein (*FMRP*) as well as with the Rho GTPase *Rac1*, which is involved in regulating axonal and dendritic outgrowth and the development and maintenance of neuronal structures (Kobayashi et al., 1998; Stefansson et al., 2008). Furthermore, haploinsufficiency of *Cyfip1* was reported to produce fragile X-like phenotypes in mice. The gene *CHRNA7* (15:32,322,685-32,462,383; cholinergic receptor, neuronal nicotinic, alpha

polypeptide 7) in the 15q13.3 microdeletion is a prominent candidate gene in 15q13.3 microdeletions associated with schizophrenia (Schizophrenia Consortium, 2008; Stefansson et al., 2008) and epilepsy (Dibbens et al., 2009; Helbig et al., 2009; Mefford et al., 2010). The nicotinic acetylcholine receptors (nAChRs) are members of a superfamily of ligand-gated ion channels that mediate fast signal transmission at synapses and showed previous linkage with schizophrenia (Leonard and Freedman, 2006). Recent findings of small, atypical 15q13.3 deletions encompassing only the *CHRNA7* gene support the evidence, that *CHRNA7* is responsible for the pathogenic effect of the 15q13.3 microdeletion (Liao et al., 2011; Hoppman-Chaney et al., 2013). Notably, homozygous microdeletions at 15q13.3 deletions result in a more severe phenotype including severe ID, epileptic encephalopathy and hypotonia (Endris et al., 2010; Spielmann et al., 2011).

In the 16p11.2 microdeletion region, the gene *PRRT2* (proline-rich transmembrane protein 2) was recently identified as a major cause for epilepsy with benign familial infantile seizures (Heron et al., 2012; Scheffer et al., 2012; Schubert et al., 2012). The gene *NDE1* (16:15,737,123-15,820,209; nudE nuclear distribution E homolog 1) at chromosomal region 16p13.3 is strongly expressed in apical precursors in the ventricular zone and in the newborn neuronal population of the human embryonic brain. Mutations in NDE1 are believed to cause both severe failure of neurogenesis and deficiency in cortical lamination, resulting in lissencephaly and/or microcephaly (Bakircioglu et al., 2011). Lying in the microdeletion region 22q11.1, gene *COMT* (chr22:19,929,262 - 19,957,497; catechol-o-methyltransferase) has been associated with schizophrenia (Palmatier et al., 2004; Lee et al., 2005). COMT is one of the major mammalian enzymes involved in the metabolic degradation of catecholamines, catalyzing the transfer of a methyl group from S-adenosyl-methionine (SAM) to a hydroxyl group on a catechol nucleus (e.g., dopamine, norepinephrine, or catechol estrogen) (Chen et al., 2004). An additional candidate gene at 22q11.1 is the SNP29 gene (chr22:21,213,292-21,245,502; synaptosomal-associated protein, 29kDa) encoding a protein involved in multiple membrane trafficking steps. The *SNAP29* protein binds to multiple syntaxins and is localized in intracellular membrane structures. Mutations in *SNPA29* were found to cause the CEDNIK syndrome (OMIM #609528) (Sprecher et al., 2005; Fuchs-Telem et al.,

2011), that shows some overlap of clinical features with 22q11.1 microdeletion carriers (Carvill and Mefford, 2013).

Microdeletion carriers with GGE display a typical distribution of clinical features of GGE syndromes regarding seizure type and age of onset compared to all analyzed cases. The investigated microdeletions showed considerable variability in familial segregation and magnitude of epileptogenic effect. The microdeletion 15q13.3 represents the most prominent risk factor, with an estimated OR > 50 based on an estimated frequency of about 0.02 % in the general population (Malhotra and Sebat, 2012). According to more extensive analyses of CNVs in 517 individuals with various idiopathic, non-lesional epilepsies, 15q13.3 microdeletions are only reported in epilepsies with GGE syndromes exclusively, besides finding in other neurodevelopmental disorders (Mefford et al., 2010).

Both 15q11.2 and 16p13.11 are more frequent in controls (0.29 % and 0.03 % respectively), thus showing lower OR (15q11.2 OR = 3.59, 16p13.3 =17.39). Due to the high frequency in controls of 15q11.2 microdeletions, 16p13.11 is the second strongest risk factor for IGE. Despite their lack of significance, high ORs are present in microdeletions at 16p11.2 (OR = 11.54) and 22q11.2 (OR > 50) as well. In particular, microdeletions in 22q11.2 are relatively rare with a frequency of < 0.2 % in GGE and 0.05 – 0.61 % in other neuropsychiatric disorders, not a single 22q11.2 microdeletion was found in any of over 70,000 combined controls (Malhotra and Sebat, 2012). Consequently, microdeletions at 22q11.2 represent the CNV with the highest effect magnitude in neuropsychiatric disorders to date.

Familial segregation of microdeletions was investigated where additional family members were available. Microdeletions at 15q11.2 in particular did not cosegregate with GGE in three large families (Figure 3-20). No other large families were available for segregation analysis. Consistent with two small families in which 15q13.3 deletions segregated with unaffected family members in the present study, Dibbens and colleagues (Dibbens et al., 2009) found incomplete penetrance of the 15q13.3 microdeletion in four out of seven pedigrees and three pedigrees included family members with GGE lacking the 15q13.3 deletion. Hence, 15q13.3 microdeletions are not sufficient to express a disease phenotype on their own, which might also vary considerably depending on the genetic background and possible

environmental effects. *De novo* mutations were detected in all six recurrent microdeletions. Microdeletions at 15q11.2 showed a lower rate of *de novo* microdeletions compared to inherited events (ratio 1:5), while inheritance of the other microdeletions was distributed more equally (15q13 ratio 3:4, 16p11.2 2:1, 16p13.11 3:5). In six families with *de novo* events, sufficient genotype data was available to identify the deleted strand, resulting in three maternal and three paternal mediated *de novo* microdeletions. The presence of *de novo* deletion events in conjunction with low population frequencies implicates purifying selection and thus may suggest a strong influence on the disease phenotype.

Despite recent findings of a male-biased autosomal effect of 16p13.11 CNVs in neurodevelopmental disorders (Tropeano et al., 2013), the frequency of 16p13.11 microdeletions in our study showed no male-biased effect. In contrast, 16p13.11 microdeletions were twice as frequent in female cases, than in male cases (eight deletions in female GGE cases, four in male GGE patients). In controls, 16p13.11 microdeletions were distributed equally.

Remarkable phenotypic variability is observed in studies published for carriers of the six recurrent microdeletions (Table 2-2). Range of the phenotypic spectrum varies from apparently unaffected carriers to individuals with severe cognitive deficits, dysmorphisms and various neuropsychiatric features. The present epilepsy sample was ascertained by the GGE phenotype excluding those patients affected by major psychiatric and mental disorders. Moreover, carriers of microdeletions were re-evaluated for the presence of intellectual disability or other neuropsychiatric disorders. Notably, six patients show learning disability, legasthenia, mild developmental delay and intellectual disability (Suppl. 6.4), although it is unclear if those features are comorbidities of, or caused by GGE. However, it is unlikely that the excess of microdeletions found in our study is caused by unobserved comorbidity of neuropsychiatric disorders and GGE.

Penetrance calculations of the reported microdeletions again show broad range of variability among our findings. In all analyzed microdeletions, penetrance is incomplete (except from 22q22.1, where no control data was available), ranging from 1.06 % (95% CI 0.01 – 0.02) for 15q11.2 to 12.49 % (95% CI 0.01 – 0.20) for 15q13.3. The incomplete penetrance suggests that the impact of recurrent

microdeletions is modified by other genetic loci and/or environmental factors (Vassos et al., 2010), supporting an oligogenic model for epilepsy. Specifically, penetrance estimations may be useful for genetic counseling in families carrying those microdeletions, accounting more information to the understanding of developing epilepsy (Vassos et al., 2010; Rosenfeld et al., 2012).

The possible mechanisms by which microdeletions mediate their pathogenic effects has recently been reviewed (Carvill and Mefford, 2013) but still remain in the very most cases. Haploinsufficiency of the deleted segment was considered the most likely mechanism for a long time (Itsara et al., 2009; Sharp, 2009; Heinzen et al., 2010). Other genetic mechanisms such as imprinting, unmasking of different recessive allelic mutations on the intact homologous chromosomal segment and background genomic variation may contribute to the highly variable phenotypic expression (Sharp, 2009), despite other acquired or environmental factors. A "two-hit" model of additional pathogenic CNVs contributing to a phenotype is conceivable as well (Girirajan et al., 2010, 2012). Recurrent microdeletions may confer a pleiotropic effect underlying various neuropsychiatric disorders. The complex interaction with additional factors might determine the specific phenotype.

Although the microdeletions investigated are individually rare (<1%) in patients with GGE, they collectively account for a significant fraction of the genetic variance of common GGE syndromes. By identifying recurrent microdeletions at 15q11.2, 15q13.3 and 16p13.11 as collectively significant genetic risk factors for GGE, our study provides new insights into the complex genetic predisposition of common epilepsies. Our present family study revealed a high percentage (>70%) of apparently unaffected parents transmitting the microdeletion to the affected child (Figure 3-20), suggesting that the microdeletion alone is not sufficient to cause an epilepsy phenotype in most cases. Likewise, unprecedented phenotypic heterogeneity has been found for seemingly identical microdeletions at 1q21.1, 16p11.2, 15q13.3 and 22q11.2, ranging from severe genomic syndromes (e.g. 22q11.2 microdeletion: DiGeorge syndrome, velocardiofacial syndrome) to a wide range of neuropsychiatric disorders (e.g. schizophrenia, intellectual disability and autism spectrum disorder), as well as in apparently unaffected individuals (for review see Mefford and Eichler, 2009). Together, our findings suggest the role of

neurodevelopmental processes in epileptogenesis. Given the frequency of recurrent microdeletions in various neuropsychiatric and neurodevelopmental disorders, identification of genetic and non-genetic factors determining phenotype–genotype relationship will be a major focus of future research.

4.3 Exon-ablating microdeletions in *NRXN1* and *RBFOX1*

Candidate gene studies are well suited for detecting association in genes with common and more complex diseases, where the risk associated with any given candidate gene is expected to be relatively small (Risch and Merikangas, 1996; Collins et al., 1997). However, this approach is largely limited by its reliance on the *a priori* knowledge about the physiological, biochemical or functional aspects of possible candidates.

The candidate genes *NRXN1* and *RBFOX1* were selected to investigate the contribution of exon-ablating microdeletions as risk factors to common GGE syndromes (see 1.5). Both studies underwent slightly different procedures for sample selection, due to a different starting point of the analysis, an accession of experience in handling large CNV data sets, different agreements with cooperation partners, and finally in different principle investigators who contributed diverging study approaches. The *NRXN1* study sample underwent little sample selection, resulting in a large dataset of 1,569 GGE cases and 6,201 controls. Samples were only filtered by SNP QC criteria and no stringent PCA was applied. In contrast, selection of samples for the *RBFOX1* study was relatively strict, resulting in 1,408 GGE cases and 2,256 controls. Stringent PCA was performed and all samples with an excess of > 50 microdeletions were excluded from analysis. An additional control cohort was not available for screening. Nevertheless, results of both studies are still valid: All detected microdeletions were confirmed with at least one additional method. Procedures for CNV validation had an important influence for consideration of QC criteria: While validation of *NRXN1* microdeletions were less complex, explanatory screenings showed that it was necessary to maintain a substantial amount of qPCR assays for validation of microdeletions in *RBFOX1*. Thus, different CNVs show a high variability in appearance, requiring careful and individual adjustment of the experimental design.

4.3.1 Exon-ablating microdeletions in *NRXN1*

We investigated the influence of exon-disrupting/removing microdeletions of *NRXN1* to increase the risk of common GGE syndromes. We detected significant association (p = 0.0049; OR = 9.91, 95% CI 1.92 – 51.12) with microdeletions of the promoter region, exon 1 and/or exon 2 of *NRXN1* in 5 of 1,569 individuals with GGE (0.3), and in 2 of 6,201 controls (0.03%). Taking into account that our population-based association analysis compared the frequency of *NRXN1* deletions in GGE patients of European origin with that observed in German population controls (EPICURE Consortium et al., 2012a), it might be possible that confounding by population stratification might affect the present association result. However, this potential bias is unlikely to play a substantial role, considering that similar frequencies of exon-disrupting *NRXN1* microdeletions have been reported in three cohorts of mainly Caucasian/European population controls (total: 3 of 7,700 = 0.039%; 1/3,181 (Schizophrenia Consortium, 2008); 0 of 2,493 (Itsara et al., 2009); 2 of 2,026 (Shaikh et al., 2009)).

A remarkable phenotypic variability for individuals with *NRXN1* deletions are reported, ranging from apparently unaffected carriers to individuals with severe cognitive deficits, ASD, and schizophrenia. The present epilepsy cohort was ascertained by the GGE phenotype, excluding those patients affected by severe ID and major psychiatric disorder such as ASD and schizophrenia (see 2.1). None of the index patients had any psychiatric disorder, and four of the five index patients had normal intelligence, whereas one had borderline intelligence. It is therefore unlikely that the excess of exonic *NRXN1* deletions found in the present study is caused by unobserved comorbidity of IGE and psychiatric disorders or severe intellectual disability.

The human neurexin gene family consists of the three genes *NRXN1*, *NRXN2*, and *NRXN3.* All three are subject to alternative promoter usage and extensive alternative splicing (Rowen et al., 2002). Two isoforms are known for each neurexin gene: A longer α-form that is transcribed from a promoter upstream of exon 1, and a short β-form transcribed from an intragenic promoter. Expression of α-neurexins is essential for functional organization of synapses. Knockout mice lacking the α-

neurexins die shortly after birth (Missler et al., 2003). All five exonic microdeletions identified in GGE patients lead to a disruption of the α-isoform of the gene.

Segregation analysis of exonic *NRXN1* microdeletions was analyzed in all five IGE families. The deletion occurred *de novo* in two patients, two were inherited from unaffected parents, and one was inherited from an affected parent.

Several family members affected with epilepsy without and two unaffected individuals with the microdeletion are present. This suggests that exonic *NRXN1* microdeletions, like other microdeletions associated with GGE, are susceptibility variants, rather than highly penetrant mutations in these families.

The exonic microdeletions in *NRXN1* showed an incomplete penetrance of 78% in our three families with an inherited microdeletion. Compared to the previous reports of *NRXN1* deletions in ID families, the penetrance in our families seems to be slightly higher (67% penetrance (Ching et al., 2010) and 50% penetrance (Gregor et al., 2011)). CNV penetrance estimations showed an incomplete penetrance of 2.9 % (95% CI 0.001-0.007). In comparison to our findings in recurrent microdeletions, this is in the middle range of our previous results, supporting the role of *NRXN1* as a risk factor with moderate effect to GGE.

In conclusion, we observed a significant excess of GGE patients carrying an exonic microdeletion of *NRXN1* compared to controls. The deletion also acts in concert with other factors to modify neurologic phenotypes. Systematic screening for additional pathologic variants through array CGH studies or massive parallel sequencing studies may reveal additional modifying variants that explain the high variability of phenotypes.

4.3.2 Exon-ablating microdeletions in *RBFOX1*

The screening for exonic microdeletions in *RBFOX1* revealed a significant excess of exonic deletions in patients with GGE compared with population controls. Microdeletions disrupting the exonic sequence of *RBFOX1* were present in 5 (0.4%) of 1,408 GGE cases and none in 2,256 controls (Fisher's exact P = 0.0083). In four of five microdeletion carriers, additional family members were available for segregation analysis. All of them were inherited, no *de novo* event was observed in families. Notably, none of the previously identified microdeletions associated with

GGE at 15q11.2, 15q13.3, and 16p13.11 was found in the index patients carrying a *RBFOX1* microdeletion.

Microdeletions showed some variability in size, ranging from 68 to 896 kb. All were located in the 5'-terminal region of *RBFOX1*, encompassing the untranslated exons one to four (Figure 3-23). The *RBFOX1* 5'-terminal exons represent highly conserved genomic sequences and are predominantly expressed in brain, suggesting that the 5'-terminal *RBFOX1* region contains important regulatory elements (Damianov and Black, 2010). Accordingly to our study, structural genomic variations disrupted the 5'-terminal exons of *RBFOX1* have been previously reported in three single patients with neurodevelopmental disorders and epilepsy (Bhalla et al., 2004; Martin et al., 2007b; Gallant et al., 2011). A significantly reduced *RBFOX1* mRNA expression in lymphocytes was reported in a female with ASD, carrying a deletion of *RBFOX1* exon 1 due to a *de novo* translocation t(15p;16p) (Martin et al., 2007b). Equally, reduction in *RBFOX1* mRNA expression might be expected in the members of family I in our study, carrying a large 896 kb microdeletion deleting *RBFOX1* exons 1–2. This family is of particular interest because of the consistent cosegregation of the GGE trait with the *RBFOX1* microdeletion (Figure 3-23).

The functional mechanism of the four smaller microdeletions involving the *RBFOX1* 5'-terminal exons 2–4, exon 1B and exon 4 remain elusive. Specifically, family III does not show cosegregation of GGE with the 163 kb spanning microdeletion affecting exon 4. This is consistent with heterogeneous cosegregation patterns, incomplete penetrance, and variable phenotypic expressivity observed for the recurrent microdeletions at 15q11.2, 15q13.3 and 16p13.11.

Furthermore, we observed familial comorbidity with other neurodevelopmental disorders, such as learning disability and autism spectrum disorder, in *RBFOX1* families I and III. As subjects with severe intellectual disability or predominant neuropsychiatric disorders others than GGE were excluded from this study, comorbidity of generalized seizures with other neurodevelopmental disorders is expected to be more common among microdeletion carriers.

Familial penetrance of exonic microdeletions in *RBFOX1* showed an incomplete penetrance of 58%, being inherited in seven affected and five unaffected family

members, which is lower than in families with exonic *NRXN1* microdeletions (78%). Similarly, the P-value of association with GGE is lower in *NRXN1* (P = 0.0048) than in *RBFOX1* (P = 0.0083). In summary, *NRXN1* seems to have greater effect as a risk factor for GGE, although no exonic microdeletions of *RBFOX1* were found in the control cohort.

Considering the key role of the splicing regulator *RBFOX1* in the control of neuronal excitation and seizure susceptibility (Gehman et al., 2011), the present findings suggest that rare microdeletions affecting the untranslated 5'-terminal *RBFOX1* exons increase risk of common GGE syndromes. Variable expressivity, incomplete penetrance, and heterogeneous cosegregation patterns suggest that *RBFOX1* deletions act as susceptibility factor in a genetically complex etiology where heterogeneous combinations of genetic factors determine the disease phenotype.

4.4 Outlook

The present study strengthens the evidence for three previously associated loci at 1q43, 2p16.1, and 2q22.3 and yielded a novel locus with suggestive evidence for association at 8q12.2. This may serve as a starting point for the identification of new genes and/or factors involved in epileptogenesis. Definitely, further validation of the new 8q21.2 locus is required in an independent sample cohort to replicate our association finding.

The same is true for the associated microdeletions at 15q11.2, 15q13.3, and 16p13.11 and the exon-ablating microdeletions in *NRXN1* and *RBFOX1*. The challenge for all association findings will be the identification of the causal variants/deletion targets and the revelation of the underlying mechanisms contributing to elevated risk of GGE. While approaches to the role of *NRXN1* and *RBFOX1* might suggest further functional analyses, the task for large recurrent microdeletions is to identify mutations in candidate genes and/or modifying factors contributing to GGE. A comprehensive CNV analysis of CNV burden and enrichment of possible disease-associated genes in CNVs of GGE patients may extend our understanding of CNV patterns in GGE syndromes.

Next generation sequencing (NGS) techniques provide a promising tool for mutation screening in both candidate genes and regions at a large scale. Still, a careful

experimental design is necessary for this approach, given the high requirements in cost and DNA resources required for NGS up to date. Massively parallel sequencing of exomes or even whole genomes gives us the opportunity to obtain a massive amount of information from a single assay, which is as well the greatest pitfall of this development. The handling and utilization of datasets of fast growing size will be a challenge not less important than the validation and functional analysis of variations detected by NGS methods.

With recent increase in CPU power and decrease of HDD memory costs, computationally intensive approaches like large scale meta-analysis on an international level and extensive SNP × SNP interaction studies are getting accomplishable. For example, a large multi-center meta-analysis of 9 GWAS of breast cancer identified 41 new susceptibility loci (Michailidou et al., 2013). Despite the promising reports, large-scale meta-analyses require the combination of several heterogeneous studies, and for conscientious data evaluation to assure valid results.

Beside the identification of new variants and mutations, methylation and gene expression analysis might contribute additional important insights to epileptogenesis. A major problem in epilepsy is the lack of appropriate material, as brain tissue of epilepsy patients can only obtained from TLE patients that underwent resection surgery. In addition, DNA acquired from brain tissue showed a lower quality than other DNAs from blood samples used in an exploratory attempt. Especially calling of CNVs resulted in low quality data. To some degree, this accounts for inappropriate handling of tissue samples during and after surgery. However, brain tissue sample from TLE patients remains a rare source. A more promising approach is the reprogramming of somatic cells to create induced pluripotent stem cells (iPSC), which can be transformed into neurons. The performance of this method has been recently improved by advances in the understanding of the underlying mechanisms (Zhao et al., 2013). Analysis of neurons generated of iPSCs from patients with schizophrenia showed altered gene expression patterns of several genes and pathways, and a decrease in neuronal connectivity and neurite number (Brennand et al., 2011).

Integrative analyses of data from GWAS and CNV analyses together with methylation and gene expression patterns may result in comprehensive findings on the basic

understanding of common GGE syndromes. All the obtained data may be put together into a systems biology approach that is more focused on underlying pathways and interactions of genetic and molecular factors than on the individual findings themselves. Systems biology is a promising trend in bioscience and epilepsy research that might result in a better understanding of disease and the development of new, maybe even multi-targeted drugs (Loeb, 2011; Margineanu, 2012).

Animal models, especially on mice and rats, have been very important to study epilepsy. Until today, they give important insights in the mechanisms underlying epilepsy (Reid et al., 2012; Yalçın, 2012b). In addition, they provide the toolkit for the research and development of new therapy approaches. A recent example is the effect of apripiprazole, well-known for the treatment of e.g. schizophrenia, in a genetic rat model with absence epilepsy and mild-depression comorbidity (Russo et al., 2013). In the last few years, zebrafish (Danio rerio) are emerging as a promising model organism to study various brain disorders (reviewed in Stewart et al., 2012). Limitations of the zebrafish model are its size, more primitive behavior patterns and the evolutionary distance to humans. But as rodent models are more expensive to maintain and more difficult to genetically modify, and invertebrates lack a complex nervous system, zebrafish offer a reasonable compromise. In addition, both larvae and adult fish are available for analysis. Zebrafish may play an important role in detecting conserved mechanism and pathways in the etiology of GGE, as well as a powerful model for new therapy approaches.

An integrative approach to identify novel epilepsy-associated genes and factors contributing to disease risk and phenotypic variability will increase our knowledge of the underlying mechanisms and the genetic architecture of common GGE syndromes. These findings may improve genetic counseling and prognosis for individuals affected by GGE and their families and may contribute to improve therapy of GGE patients.

5 References

Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.

Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Peltonen, L., et al. (2010). Integrating common and rare genetic variation in diverse human populations. Nature *467*, 52–58.

Auweter, S.D., Fasan, R., Reymond, L., Underwood, J.G., Black, D.L., Pitsch, S., and Allain, F.H.-T. (2006). Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. EMBO J. *25*, 163–173.

Bakircioglu, M., Carvalho, O.P., Khurshid, M., Cox, J.J., Tuysuz, B., Barak, T., Yilmaz, S., Caglayan, O., Dincer, A., Nicholas, A.K., et al. (2011). The essential role of centrosomal NDE1 in human cerebral cortex neurogenesis. Am. J. Hum. Genet. *88*, 523–535.

Bando, S.Y., Alegro, M.C., Amaro, E., Silva, A.V., Castro, L.H.M., Wen, H.-T., Lima, L. de A., Brentani, H., and Moreira-Filho, C.A. (2011). Hippocampal CA3 Transcriptome Signature Correlates with Initial Precipitating Injury in Refractory Mesial Temporal Lobe Epilepsy. PLoS ONE *6*, e26268.

Barnby, G., Abbott, A., Sykes, N., Morris, A., Weeks, D.E., Mott, R., Lamb, J., Bailey, A.J., and Monaco, A.P. (2005). Candidate-gene screening and association analysis at the autism-susceptibility locus on chromosome 16p: evidence of association at GRIN2A and ABAT. Am. J. Hum. Genet. *76*, 950–966.

Bassett, A.S., Marshall, C.R., Lionel, A.C., Chow, E.W.C., and Scherer, S.W. (2008). Copy number variations and risk for schizophrenia in 22q11.2 deletion syndrome. Hum. Mol. Genet. *17*, 4045–4053.

Baulac, S., Huberfeld, G., Gourfinkel-An, I., Mitropoulou, G., Beranger, A., Prud'homme, J.F., Baulac, M., Brice, A., Bruzzone, R., and LeGuern, E. (2001). First genetic evidence of GABA(A) receptor dysfunction in epilepsy: a mutation in the gamma2-subunit gene. Nat. Genet. *28*, 46–48.

Berg, A.T., Shinnar, S., Levy, S.R., and Testa, F.M. (1999). Newly diagnosed epilepsy in children: presentation at diagnosis. Epilepsia *40*, 445–452.

Berg, A.T., Berkovic, S.F., Brodie, M.J., Buchhalter, J., Cross, J.H., van Emde Boas, W., Engel, J., French, J., Glauser, T.A., Mathern, G.W., et al. (2010). Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. Epilepsia *51*, 676–685.

Van den Berghe, V., Stappers, E., Vandesande, B., Dimidschstein, J., Kroes, R., Francis, A., Conidi, A., Lesage, F., Dries, R., Cazzola, S., et al. (2013). Directed migration of cortical interneurons depends on the cell-autonomous action of Sip1. Neuron *77*, 70–82.

Bergren, S.K., Chen, S., Galecki, A., and Kearney, J.A. (2005). Genetic modifiers affecting severity of epilepsy caused by mutation of sodium channel Scn2a. Mamm. Genome *16*, 683–690.

Berkovic, S.F., Howell, R.A., Hay, D.A., and Hopper, J.L. (1998). Epilepsies in twins: genetics of the major epilepsy syndromes. Ann. Neurol. *43*, 435–445.

Berkovic, S.F., Mulley, J.C., Scheffer, I.E., and Petrou, S. (2006). Human epilepsies: interaction of genetic and acquired factors. Trends Neurosci. *29*, 391–397.

Bhalla, K., Phillips, H.A., Crawford, J., McKenzie, O.L.D., Mulley, J.C., Eyre, H., Gardner, A.E., Kremmidiotis, G., and Callen, D.F. (2004). The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene. J. Hum. Genet. *49*, 308–311.

Bianchi, A., Viaggi, S., and Chiossi, E. (2003). Family study of epilepsy in first degree relatives: data from the Italian Episcreen Study. Seizure *12*, 203–210.

Bland, J.M., and Altman, D.G. (1995). Multiple significance tests: the Bonferroni method. BMJ *310*, 170.

Brennand, K.J., Simone, A., Jou, J., Gelboin-Burkhart, C., Tran, N., Sangar, S., Li, Y., Mu, Y., Chen, G., Yu, D., et al. (2011). Modelling schizophrenia using human induced pluripotent stem cells. Nature *473*, 221–225.

Brunetti-Pierri, N., Berg, J.S., Scaglia, F., Belmont, J., Bacino, C.A., Sahoo, T., Lalani, S.R., Graham, B., Lee, B., Shinawi, M., et al. (2008). Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. Nat. Genet. *40*, 1466–1471.

Bundey, S., Hardy, C., Vickers, S., Kilpatrick, M.W., and Corbett, J.A. (1994). Duplication of the 15q11-13 region in a patient with autism, epilepsy and ataxia. Dev Med Child Neurol *36*, 736–742.

Butler, M.G., Meaney, F.J., and Palmer, C.G. (1986). Clinical and cytogenetic survey of 39 individuals with Prader-Labhart-Willi syndrome. Am. J. Med. Genet. *23*, 793–809.

Bymaster, F.P., Carter, P.A., Yamada, M., Gomeza, J., Wess, J., Hamilton, S.E., Nathanson, N.M., McKinzie, D.L., and Felder, C.C. (2003). Role of specific muscarinic receptor subtypes in cholinergic parasympathomimetic responses, in vivo phosphoinositide hydrolysis, and pilocarpine-induced seizure activity. European Journal of Neuroscience *17*, 1403–1410.

Callenbach, P.M., Geerts, A.T., Arts, W.F., van Donselaar, C.A., Peters, A.C., Stroink, H., and Brouwer, O.F. (1998). Familial occurrence of epilepsy in children with newly diagnosed multiple seizures: Dutch Study of Epilepsy in Childhood. Epilepsia *39*, 331–336.

Capovilla, G., Berg, A.T., Cross, J.H., Moshe, S.L., Vigevano, F., Wolf, P., and Avanzini, G. (2009). Conceptual dichotomies in classifying epilepsies: Partial versus generalized and idiopathic versus symptomatic (April 18-20, 2008, Monreale, Italy). Epilepsia.

Carvill, G.L., and Mefford, H.C. (2013). Microdeletion syndromes. Curr. Opin. Genet. Dev.

Cavalleri, G.L., Weale, M.E., Shianna, K.V., Singh, R., Lynch, J.M., Grinton, B., Szoeke, C., Murphy, K., Kinirons, P., O'Rourke, D., et al. (2007). Multicentre search for genetic susceptibility loci in sporadic epilepsy syndrome and seizure types: a case-control study. Lancet Neurol *6*, 970–980.

Chabchoub, E., Vermeesch, J.R., de Ravel, T., de Cock, P., and Fryns, J.-P. (2008). The facial dysmorphy in the newly recognised microdeletion 2p15-p16.1 refined to a 570 kb region in 2p15. J. Med. Genet. *45*, 189–192.

Chandler, K.E., Del Rio, A., Rakshi, K., Springell, K., Williams, D.K., Stoodley, N., Woods, C.G., and Pilz, D.T. (2006). Leucodysplasia, microcephaly, cerebral malformation (LMC): a novel recessive disorder linked to 2p16. Brain *129*, 272–277.

Chen, J., Lipska, B.K., Halim, N., Ma, Q.D., Matsumoto, M., Melhem, S., Kolachana, B.S., Hyde, T.M., Herman, M.M., Apud, J., et al. (2004). Functional analysis of genetic variation in catechol-O-methyltransferase (COMT): effects on mRNA, protein, and enzyme activity in postmortem human brain. Am. J. Hum. Genet. *75*, 807–821.

Ching, M.S.L., Shen, Y., Tan, W.-H., Jeste, S.S., Morrow, E.M., Chen, X., Mukaddes, N.M., Yoo, S.-Y., Hanson, E., Hundley, R., et al. (2010). Deletions of NRXN1 (neurexin-1) predispose to a wide spectrum of developmental disorders. Am. J. Med. Genet. B Neuropsychiatr. Genet. *153B*, 937–947.

Chioza, B., Wilkie, H., Nashef, L., Blower, J., McCormick, D., Sham, P., Asherson, P., and Makoff, A.J. (2001). Association between the alpha(1a) calcium channel gene CACNA1A and idiopathic generalized epilepsy. Neurology *56*, 1245–1246.

Claes, L., Del-Favero, J., Ceulemans, B., Lagae, L., Van Broeckhoven, C., and De Jonghe, P. (2001). De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy. Am. J. Hum. Genet. *68*, 1327–1332.

Cobb, S.R., Buhl, E.H., Halasy, K., Paulsen, O., and Somogyi, P. (1995). Synchronization of neuronal activity in hippocampus by individual GABAergic interneurons. Nature *378*, 75–78.

Colella, S., Yau, C., Taylor, J.M., Mirza, G., Butler, H., Clouston, P., Bassett, A.S., Seller, A., Holmes, C.C., and Ragoussis, J. (2007). QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. Nucleic Acids Res. *35*, 2013–2025.

Collins, F.S., Guyer, M.S., and Charkravarti, A. (1997). Variations on a theme: cataloging human DNA sequence variation. Science *278*, 1580–1581.

Cook, E.H., Jr, and Scherer, S.W. (2008). Copy-number variations associated with neuropsychiatric conditions. Nature *455*, 919–923.

Cordelli, D.M., Garavelli, L., Savasta, S., Guerra, A., Pellicciari, A., Giordano, L., Bonetti, S., Cecconi, I., Wischmeijer, A., Seri, M., et al. (2013). Epilepsy in Mowat-

Wilson syndrome: delineation of the electroclinical phenotype. Am. J. Med. Genet. A *161A*, 273–284.

Corradini, I., Verderio, C., Sala, M., Wilson, M.C., and Matteoli, M. (2009). SNAP-25 in neuropsychiatric disorders. Ann. N. Y. Acad. Sci. *1152*, 93–99.

Cossette, P., Liu, L., Brisebois, K., Dong, H., Lortie, A., Vanasse, M., Saint-Hilaire, J.-M., Carmant, L., Verner, A., Lu, W.-Y., et al. (2002). Mutation of GABRA1 in an autosomal dominant form of juvenile myoclonic epilepsy. Nat. Genet. *31*, 184–189.

Crunelli, V., and Leresche, N. (2002). Childhood absence epilepsy: genes, channels, neurons and networks. Nat. Rev. Neurosci. *3*, 371–382.

Damianov, A., and Black, D.L. (2010). Autoregulation of Fox protein expression to produce dominant negative splicing factors. RNA *16*, 405–416.

Davy, B.E., and Robinson, M.L. (2003). Congenital hydrocephalus in hy3 mice is caused by a frameshift mutation in Hydin, a large novel gene. Hum. Mol. Genet. *12*, 1163–1170.

Devlin, B., and Roeder, K. (1999). Genomic control for association studies. Biometrics *55*, 997–1004.

Devlin, B., Bacanu, S.-A., and Roeder, K. (2004). Genomic Control to the extreme. Nat. Genet. *36*, 1129–1130; author reply 1131.

Dibbens, L.M., Tarpey, P.S., Hynes, K., Bayly, M.A., Scheffer, I.E., Smith, R., Bomar, J., Sutton, E., Vandeleur, L., Shoubridge, C., et al. (2008). X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. Nat. Genet. *40*, 776–781.

Dibbens, L.M., Mullen, S., Helbig, I., Mefford, H.C., Bayly, M.A., Bellows, S., Leu, C., Trucks, H., Obermeier, T., Wittig, M., et al. (2009). Familial and sporadic 15q13.3 microdeletions in idiopathic generalized epilepsy: precedent for disorders with complex inheritance. Hum. Mol. Genet. *18*, 3626–3631.

Ding, J.B., Guzman, J.N., Peterson, J.D., Goldberg, J.A., and Surmeier, D.J. (2010). Thalamic Gating of Corticostriatal Signaling by Cholinergic Interneurons. Neuron *67*, 294–307.

Duncan, J.S., Sander, J.W., Sisodiya, S.M., and Walker, M.C. (2006). Adult epilepsy. Lancet *367*, 1087–1100.

Duong, L., Klitten, L.L., Møller, R.S., Ingason, A., Jakobsen, K.D., Skjødt, C., Didriksen, M., Hjalgrim, H., Werge, T., and Tommerup, N. (2012). Mutations in NRXN1 in a family multiply affected with brain disorders: NRXN1 mutations and brain disorders. Am. J. Med. Genet. B Neuropsychiatr. Genet. *159B*, 354–358.

Durner, M., Keddache, M.A., Tomasini, L., Shinnar, S., Resor, S.R., Cohen, J., Harden, C., Moshe, S.L., Rosenbaum, D., Kang, H., et al. (2001). Genome scan of idiopathic generalized epilepsy: evidence for major susceptibility gene and modifying genes influencing the seizure type. Ann. Neurol. *49*, 328–335.

Elia, J., Glessner, J.T., Wang, K., Takahashi, N., Shtir, C.J., Hadley, D., Sleiman, P.M.A., Zhang, H., Kim, C.E., Robison, R., et al. (2012). Genome-wide copy number variation study associates metabotropic glutamate receptor gene networks with attention deficit hyperactivity disorder. Nat. Genet. *44*, 78–84.

Endris, V., Hackmann, K., Neuhann, T.M., Grasshoff, U., Bonin, M., Haug, U., Hahn, G., Schallner, J.C., Schröck, E., Tinschert, S., et al. (2010). Homozygous loss of CHRNA7 on chromosome 15q13.3 causes severe encephalopathy with seizures and hypotonia. Am. J. Med. Genet. A *152A*, 2908–2911.

Engel, J., Jr (2001). A proposed diagnostic scheme for people with epileptic seizures and with epilepsy: report of the ILAE Task Force on Classification and Terminology. Epilepsia *42*, 796–803.

Engel, J., Jr (2006). Report of the ILAE classification core group. Epilepsia *47*, 1558–1568.

EPICURE Consortium, EMINet Consortium, Steffens, M., Leu, C., Ruppert, A.-K., Zara, F., Striano, P., Robbiano, A., Capovilla, G., Tinuper, P., et al. (2012a). Genome-wide association analysis of genetic generalized epilepsies implicates susceptibility loci at 1q43, 2p16.1, 2q22.3 and 17q21.32. Hum. Mol. Genet. *21*, 5359–5372.

EPICURE Consortium, Leu, C., de Kovel, C.G.F., Zara, F., Striano, P., Pezzella, M., Robbiano, A., Bianchi, A., Bisulli, F., Coppola, A., et al. (2012b). Genome-wide linkage meta-analysis identifies susceptibility loci at 2q34 and 13q31.3 for genetic generalized epilepsies. Epilepsia *53*, 308–318.

Escayg, A., De Waard, M., Lee, D.D., Bichet, D., Wolf, P., Mayer, T., Johnston, J., Baloh, R., Sander, T., and Meisler, M.H. (2000). Coding and noncoding variation of the human calcium-channel beta4-subunit gene CACNB4 in patients with idiopathic generalized epilepsy and episodic ataxia. Am. J. Hum. Genet. *66*, 1531–1539.

Ewens, W.J., and Spielman, R.S. (1995). The transmission/disequilibrium test: history, subdivision, and admixture. Am. J. Hum. Genet. *57*, 455–464.

Félix, T.M., Petrin, A.L., Sanseverino, M.T.V., and Murray, J.C. (2010). Further characterization of microdeletion syndrome involving 2p15-p16.1. Am. J. Med. Genet. A *152A*, 2604–2608.

Feuk, L., Carson, A.R., and Scherer, S.W. (2006). Structural variation in the human genome. Nat. Rev. Genet. *7*, 85–97.

Fisher, R.S., van Emde Boas, W., Blume, W., Elger, C., Genton, P., Lee, P., and Engel, J., Jr (2005). Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Epilepsia *46*, 470–472.

Fogel, B.L., Wexler, E., Wahnich, A., Friedrich, T., Vijayendran, C., Gao, F., Parikshak, N., Konopka, G., and Geschwind, D.H. (2012). RBFOX1 regulates both splicing and transcriptional networks in human neuronal development. Hum. Mol. Genet. *21*, 4171–4186.

Freitag, C.M., May, T.W., Pfäfflin, M., König, S., and Rating, D. (2001). Incidence of epilepsies and epileptic syndromes in children and adolescents: a population-based prospective study in Germany. Epilepsia *42*, 979–985.

Fuchs-Telem, D., Stewart, H., Rapaport, D., Nousbeck, J., Gat, A., Gini, M., Lugassy, Y., Emmert, S., Eckl, K., Hennies, H.C., et al. (2011). CEDNIK syndrome results from loss-of-function mutations in SNAP29. Br. J. Dermatol. *164*, 610–616.

Gallant, N.M., Baldwin, E., Salamon, N., Dipple, K.M., and Quintero-Rivera, F. (2011). Pontocerebellar hypoplasia in association with de novo 19p13.11p13.12 microdeletion. Am. J. Med. Genet. A *155A*, 2871–2878.

Garavelli, L., Zollino, M., Mainardi, P.C., Gurrieri, F., Rivieri, F., Soli, F., Verri, R., Albertini, E., Favaron, E., Zignani, M., et al. (2009). Mowat-Wilson syndrome: facial phenotype changing with age: study of 19 Italian patients and review of the literature. Am. J. Med. Genet. A *149A*, 417–426.

Gardiner, M. (2005). Genetics of idiopathic generalized epilepsies. Epilepsia *46 Suppl 9*, 15–20.

Gastaut, H. (1964). A Proposed International Classification of Epileptic Seizures. Epilepsia *5*, 297–306.

Gastaut, H. (1969). Classification of the epilepsies. Proposal for an international classification. Epilepsia *10*, Suppl:14–21.

Gehman, L.T., Stoilov, P., Maguire, J., Damianov, A., Lin, C.-H., Shiue, L., Ares, M., Jr, Mody, I., and Black, D.L. (2011). The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. Nat. Genet. *43*, 706–711.

Girirajan, S., Rosenfeld, J.A., Cooper, G.M., Antonacci, F., Siswara, P., Itsara, A., Vives, L., Walsh, T., McCarthy, S.E., Baker, C., et al. (2010). A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. Nat. Genet. *42*, 203–209.

Girirajan, S., Rosenfeld, J.A., Coe, B.P., Parikh, S., Friedman, N., Goldstein, A., Filipink, R.A., McConnell, J.S., Angle, B., Meschino, W.S., et al. (2012). Phenotypic heterogeneity of genomic disorders and rare copy-number variants. N. Engl. J. Med. *367*, 1321–1331.

Glasscock, E., Qian, J., Yoo, J.W., and Noebels, J.L. (2007). Masking epilepsy by combining two epilepsy genes. Nat. Neurosci. *10*, 1554–1558.

Greenberg, D.A., and Subaran, R. (2011). Blinders, phenotype, and fashionable genetic analysis: a critical examination of the current state of epilepsy genetic studies. Epilepsia *52*, 1–9.

Greenberg, D.A., Durner, M., Keddache, M., Shinnar, S., Resor, S.R., Moshe, S.L., Rosenbaum, D., Cohen, J., Harden, C., Kang, H., et al. (2000). Reproducibility and complications in gene searches: linkage on chromosome 6, heterogeneity, association, and maternal inheritance in juvenile myoclonic epilepsy. Am. J. Hum. Genet. *66*, 508–516.

Gregor, A., Albrecht, B., Bader, I., Bijlsma, E.K., Ekici, A.B., Engels, H., Hackmann, K., Horn, D., Hoyer, J., Klapecki, J., et al. (2011). Expanding the clinical spectrum associated with defects in CNTNAP2 and NRXN1. BMC Med. Genet. *12*, 106.

Gu, W., and Lupski, J.R. (2009). CNV and nervous system diseases - what's new? Cytogenet Genome Res *123*, 54–64.

Gu, W., Zhang, F., and Lupski, J.R. (2008). Mechanisms for human genomic rearrangements. Pathogenetics *1*, 4.

Guo, Y., Baum, L.W., Sham, P.C., Wong, V., Ng, P.W., Lui, C.H.T., Sin, N.C., Tsoi, T.H., Tang, C.S.M., Kwan, J.S.H., et al. (2012). Two-stage genome-wide association study identifies variants in CAMSAP1L1 as susceptibility loci for epilepsy in Chinese. Hum. Mol. Genet. *21*, 1184–1189.

Gurnett, C.A., and Hedera, P. (2007). New ideas in epilepsy genetics: novel epilepsy genes, copy number alterations, and gene regulation. Arch. Neurol. *64*, 324–328.

Hannes, F.D., Sharp, A.J., Mefford, H.C., de Ravel, T., Ruivenkamp, C.A., Breuning, M.H., Fryns, J.-P., Devriendt, K., Van Buggenhout, G., Vogels, A., et al. (2009). Recurrent reciprocal deletions and duplications of 16p13.11: the deletion is a risk factor for MR/MCA while the duplication may be a rare benign variant. J. Med. Genet. *46*, 223–232.

Harkin, L.A., McMahon, J.M., Iona, X., Dibbens, L., Pelekanos, J.T., Zuberi, S.M., Sadleir, L.G., Andermann, E., Gill, D., Farrell, K., et al. (2007). The spectrum of SCN1A-related infantile epileptic encephalopathies. Brain *130*, 843–852.

Harrison, V., Connell, L., Hayesmoore, J., McParland, J., Pike, M.G., and Blair, E. (2011). Compound heterozygous deletion of NRXN1 causing severe developmental delay with early onset epilepsy in two sisters. American Journal of Medical Genetics Part A *155*, 2826–2831.

Hauge, H., Patzke, S., and Aasheim, H.-C. (2007). Characterization of the FAM110 gene family. Genomics *90*, 14–27.

Hauser, W.A., Annegers, J.F., and Kurland, L.T. (1993). Incidence of epilepsy and unprovoked seizures in Rochester, Minnesota: 1935-1984. Epilepsia *34*, 453–468.

Heinzen, E.L., Radtke, R.A., Urban, T.J., Cavalleri, G.L., Depondt, C., Need, A.C., Walley, N.M., Nicoletti, P., Ge, D., Catarino, C.B., et al. (2010). Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. Am. J. Hum. Genet. *86*, 707–718.

Heinzen, E.L., Depondt, C., Cavalleri, G.L., Ruzzo, E.K., Walley, N.M., Need, A.C., Ge, D., He, M., Cirulli, E.T., Zhao, Q., et al. (2012). Exome sequencing followed by large-scale genotyping fails to identify single rare variants of large effect in idiopathic generalized epilepsy. Am. J. Hum. Genet. *91*, 293–302.

Helbig, I., and Lowenstein, D.H. (2013). Genetics of the epilepsies: where are we and where are we going? Curr. Opin. Neurol. *26*, 179–185.

Helbig, I., Scheffer, I.E., Mulley, J.C., and Berkovic, S.F. (2008). Navigating the channels and beyond: unravelling the genetics of the epilepsies. Lancet Neurol *7*, 231–245.

Helbig, I., Mefford, H.C., Sharp, A.J., Guipponi, M., Fichera, M., Franke, A., Muhle, H., de Kovel, C., Baker, C., von Spiczak, S., et al. (2009). 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. Nat. Genet. *41*, 160–162.

Hempelmann, A., Taylor, K.P., Heils, A., Lorenz, S., Prud'homme, J.-F., Nabbout, R., Dulac, O., Rudolf, G., Zara, F., Bianchi, A., et al. (2006). Exploration of the genetic architecture of idiopathic generalized epilepsies. Epilepsia *47*, 1682–1690.

Heron, S.E., Grinton, B.E., Kivity, S., Afawi, Z., Zuberi, S.M., Hughes, J.N., Pridmore, C., Hodgson, B.L., Iona, X., Sadleir, L.G., et al. (2012). PRRT2 mutations cause benign familial infantile epilepsy and infantile convulsions with choreoathetosis syndrome. Am. J. Hum. Genet. *90*, 152–160.

Hildebrand, M.S., Dahl, H.-H.M., Damiano, J.A., Smith, R.J.H., Scheffer, I.E., and Berkovic, S.F. (2013). Recent advances in the molecular genetics of epilepsy. J. Med. Genet.

Hoffmann, T.J., Kvale, M.N., Hesselson, S.E., Zhan, Y., Aquino, C., Cao, Y., Cawley, S., Chung, E., Connell, S., Eshragh, J., et al. (2011). Next generation genome-wide association tool: design and coverage of a high-throughput European-optimized SNP array. Genomics *98*, 79–89.

Hoppman-Chaney, N., Wain, K., Seger, P.R., Superneau, D.W., and Hodge, J.C. (2013). Identification of single gene deletions at 15q13.3: further evidence that CHRNA7 causes the 15q13.3 microdeletion syndrome phenotype. Clin. Genet. *83*, 345–351.

Howie, B.N., Donnelly, P., and Marchini, J. (2009). A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. PLoS Genet *5*, e1000529.

Iafrate, A.J., Feuk, L., Rivera, M.N., Listewnik, M.L., Donahoe, P.K., Qi, Y., Scherer, S.W., and Lee, C. (2004). Detection of large-scale variation in the human genome. Nat. Genet. *36*, 949–951.

ILAE (1981). Proposal for revised clinical and electroencephalographic classification of epileptic seizures. From the Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia *22*, 489–501.

ILAE (1985). Proposal for classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia *26*, 268–278.

ILAE (1989). Proposal for revised classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia *30*, 389–399.

Itsara, A., Cooper, G.M., Baker, C., Girirajan, S., Li, J., Absher, D., Krauss, R.M., Myers, R.M., Ridker, P.M., Chasman, D.I., et al. (2009). Population analysis of large copy number variants and hotspots of human genetic disease. Am. J. Hum. Genet. *84*, 148–161.

Jallon, P., Loiseau, P., and Loiseau, J. (2001). Newly diagnosed unprovoked epileptic seizures: presentation at diagnosis in CAROLE study. Coordination Active du Réseau Observatoire Longitudinal de l' Epilepsie. Epilepsia *42*, 464–475.

Janz, D. (1997). The Idiopathic Generalized Epilepsies of Adolescence with Childhood and Juvenile Age of Onset. Epilepsia *38*, 4–11.

Jensen, F.E. (2011). Epilepsy as a spectrum disorder: Implications from novel clinical and basic neuroscience. Epilepsia *52 Suppl 1*, 1–6.

Jin, Y., Suzuki, H., Maegawa, S., Endo, H., Sugano, S., Hashimoto, K., Yasuda, K., and Inoue, K. (2003). A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. EMBO J. *22*, 905–912.

Johnson, M.R. (2011). Comment on "Blinders, phenotype, and fashionable genetic analysis: a critical examination of the current state of epilepsy genetic studies."Epilepsia *52*, 190–191; dicussion 193–196.

Johnson, M.R., Milne, R.L., Torn-Broers, Y., Hopper, J.L., Scheffer, I.E., and Berkovic, S.F. (2003). A twin study of genetic influences on epilepsy outcome. Twin Res *6*, 140–146.

Jorge, B.S., Campbell, C.M., Miller, A.R., Rutter, E.D., Gurnett, C.A., Vanoye, C.G., George, A.L., Jr, and Kearney, J.A. (2011). Voltage-gated potassium channel KCNV2 (Kv8.2) contributes to epilepsy susceptibility. Proc. Natl. Acad. Sci. U.S.A. *108*, 5443–5448.

Jouvenceau, A., Eunson, L.H., Spauschus, A., Ramesh, V., Zuberi, S.M., Kullmann, D.M., and Hanna, M.G. (2001). Human epilepsy associated with dysfunction of the brain P/Q-type calcium channel. Lancet *358*, 801–807.

Kanner, A.M. (2013). Do psychiatric comorbidities have a negative impact on the course and treatment of seizure disorders? Curr. Opin. Neurol. *26*, 208–213.

Kasperaviciūte, D., Catarino, C.B., Heinzen, E.L., Depondt, C., Cavalleri, G.L., Caboclo, L.O., Tate, S.K., Jamnadas-Khoda, J., Chinthapalli, K., Clayton, L.M.S., et al. (2010). Common genetic variation and susceptibility to partial epilepsies: a genome-wide association study. Brain *133*, 2136–2147.

Kearney, J.A., Yang, Y., Beyer, B., Bergren, S.K., Claes, L., Dejonghe, P., and Frankel, W.N. (2006). Severe epilepsy resulting from genetic interaction between Scn2a and Kcnq2. Hum. Mol. Genet. *15*, 1043–1048.

Kennedy, G.C., Matsuzaki, H., Dong, S., Liu, W., Huang, J., Liu, G., Su, X., Cao, M., Chen, W., Zhang, J., et al. (2003). Large-scale genotyping of complex DNA. Nat. Biotechnol. *21*, 1233–1237.

Kim, D., Song, I., Keum, S., Lee, T., Jeong, M.J., Kim, S.S., McEnery, M.W., and Shin, H.S. (2001). Lack of the burst firing of thalamocortical relay neurons and resistance to absence seizures in mice lacking alpha(1G) T-type Ca(2+) channels. Neuron *31*, 35–45.

Kinirons, P., Rabinowitz, D., Gravel, M., Long, J., Winawer, M., Sénéchal, G., Ottman, R., and Cossette, P. (2008). Phenotypic concordance in 70 families with IGE-implications for genetic studies of epilepsy. Epilepsy Res. *82*, 21–28.

Kirov, G., Grozeva, D., Norton, N., Ivanov, D., Mantripragada, K.K., Holmans, P., Craddock, N., Owen, M.J., and O'Donovan, M.C. (2009). Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. Hum. Mol. Genet. *18*, 1497–1503.

Kjeldsen, M.J., Corey, L.A., Christensen, K., and Friis, M.L. (2003). Epileptic seizures and syndromes in twins: the importance of genetic factors. Epilepsy Res. *55*, 137–146.

Kobayashi, K., Kuroda, S., Fukata, M., Nakamura, T., Nagase, T., Nomura, N., Matsuura, Y., Yoshida-Kubomura, N., Iwamatsu, A., and Kaibuchi, K. (1998). p140Sra-1 (specifically Rac1-associated protein) is a novel specific target for Rac1 small GTPase. J. Biol. Chem. *273*, 291–295.

Korbel, J.O., Urban, A.E., Affourtit, J.P., Godwin, B., Grubert, F., Simons, J.F., Kim, P.M., Palejev, D., Carriero, N.J., Du, L., et al. (2007). Paired-End Mapping Reveals Extensive Structural Variation in the Human Genome. Science *318*, 420–426.

Korn, J.M., Kuruvilla, F.G., McCarroll, S.A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P.J., Darvishi, K., et al. (2008). Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat. Genet. *40*, 1253–1260.

De Kovel, C.G.F., Trucks, H., Helbig, I., Mefford, H.C., Baker, C., Leu, C., Kluck, C., Muhle, H., von Spiczak, S., Ostertag, P., et al. (2010). Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. Brain *133*, 23–32.

Krawczak, M., Nikolaus, S., von Eberstein, H., Croucher, P.J.P., El Mokhtari, N.E., and Schreiber, S. (2006). PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. Community Genet *9*, 55–61.

Kumar, R.A., KaraMohamed, S., Sudi, J., Conrad, D.F., Brune, C., Badner, J.A., Gilliam, T.C., Nowak, N.J., Cook, E.H., Jr, Dobyns, W.B., et al. (2008). Recurrent 16p11.2 microdeletions in autism. Hum. Mol. Genet. *17*, 628–638.

Lachance-Touchette, P., Brown, P., Meloche, C., Kinirons, P., Lapointe, L., Lacasse, H., Lortie, A., Carmant, L., Bedford, F., Bowie, D., et al. (2011). Novel α1 and γ2 GABAA receptor subunit mutations in families with idiopathic generalized epilepsy. Eur. J. Neurosci. *34*, 237–249.

Lal, D., Trucks, H., Møller, R.S., Hjalgrim, H., Koeleman, B.P.C., de Kovel, C.G.F., Visscher, F., Weber, Y.G., Lerche, H., Becker, F., et al. (2013). Rare exonic deletions of the RBFOX1 gene increase risk of idiopathic generalized epilepsy. Epilepsia *54*, 265–271.

Lee, J.A., Carvalho, C.M.B., and Lupski, J.R. (2007). A DNA Replication Mechanism for Generating Nonrecurrent Rearrangements Associated with Genomic Disorders. Cell *131*, 1235–1247.

Lee, S.-G., Joo, Y., Kim, B., Chung, S., Kim, H.-L., Lee, I., Choi, B., Kim, C., and Song, K. (2005). Association of Ala72Ser polymorphism with COMT enzyme activity and the risk of schizophrenia in Koreans. Hum. Genet. *116*, 319–328.

Léon, C., de Nijs, L., Chanas, G., Delgado-Escueta, A.V., Grisar, T., and Lakaye, B. (2010). Distribution of EFHC1 or Myoclonin 1 in mouse neural structures. Epilepsy Res. *88*, 196–207.

Leonard, S., and Freedman, R. (2006). Genetics of chromosome 15q13-q14 in schizophrenia. Biol. Psychiatry *60*, 115–122.

Li, M., Wang, Y., Zheng, X., Ikeda, M., Iwata, N., Luo, X., Chong, S.-A., Lee, J., Rietschel, M., Zhang, F., et al. (2012). Meta-analysis and brain imaging data support the involvement of VRK2 (rs2312147) in schizophrenia susceptibility. Schizophr. Res. *142*, 200–205.

Liang, J., Zhang, Y., Chen, Y., Wang, J., Pan, H., Wu, H., Xu, K., Liu, X., Jiang, Y., Shen, Y., et al. (2007). Common polymorphisms in the CACNA1H gene associated with childhood absence epilepsy in Chinese Han population. Ann. Hum. Genet. *71*, 325–335.

Liao, J., DeWard, S.J., Madan-Khetarpal, S., Surti, U., and Hu, J. (2011). A small homozygous microdeletion of 15q13.3 including the CHRNA7 gene in a girl with a spectrum of severe neurodevelopmental features. Am. J. Med. Genet. A *155A*, 2795–2800.

Lippert, C., Listgarten, J., Liu, Y., Kadie, C.M., Davidson, R.I., and Heckerman, D. (2011). FaST linear mixed models for genome-wide association studies. Nat Meth *8*, 833–835.

Loeb, J.A. (2011). Identifying targets for preventing epilepsy using systems biology. Neurosci. Lett. *497*, 205–212.

Lupski, J.R., and Stankiewicz, P. (2005). Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet. *1*, e49.

Macdonald, R.L., Kang, J.-Q., and Gallagher, M.J. (2010). Mutations in GABAA receptor subunits associated with genetic epilepsies. J. Physiol. (Lond.) *588*, 1861–1869.

Mägi, R., and Morris, A.P. (2010). GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics *11*, 288.

Malhotra, D., and Sebat, J. (2012). CNVs: harbingers of a rare variant revolution in psychiatric genetics. Cell *148*, 1223–1241.

Marchini, J., Howie, B., Myers, S., McVean, G., and Donnelly, P. (2007). A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet *39*, 906–913.

Margineanu, D.G. (2012). Systems biology impact on antiepileptic drug discovery. Epilepsy Res. *98*, 104–115.

Marini, C., Scheffer, I.E., Crossland, K.M., Grinton, B.E., Phillips, F.L., McMahon, J.M., Turner, S.J., Dean, J.T., Kivity, S., Mazarib, A., et al. (2004). Genetic architecture of idiopathic generalized epilepsy: clinical genetic analysis of 55 multiplex families. Epilepsia *45*, 467–478.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. Nat. Rev. Neurosci. *5*, 793–807.

Marsh, E., Fulp, C., Gomez, E., Nasrallah, I., Minarcik, J., Sudi, J., Christian, S.L., Mancini, G., Labosky, P., Dobyns, W., et al. (2009). Targeted loss of Arx results in a developmental epilepsy mouse model and recapitulates the human phenotype in heterozygous females. Brain *132*, 1563–1576.

Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y., et al. (2008). Structural variation of chromosomes in autism spectrum disorder. Am. J. Hum. Genet. *82*, 477–488.

Martin, C.L., Duvall, J.A., Ilkin, Y., Simon, J.S., Arreaza, M.G., Wilkes, K., Alvarez-Retuerto, A., Whichello, A., Powell, C.M., Rao, K., et al. (2007a). Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. Am. J. Med. Genet. B Neuropsychiatr. Genet. *144B*, 869–876.

Martin, M.S., Tang, B., Papale, L.A., Yu, F.H., Catterall, W.A., and Escayg, A. (2007c). The voltage-gated sodium channel Scn8a is a genetic modifier of severe myoclonic epilepsy of infancy. Hum. Mol. Genet. *16*, 2892–2899.

Martínez-Juárez, I.E., Alonso, M.E., Medina, M.T., Durón, R.M., Bailey, J.N., López-Ruiz, M., Ramos-Ramírez, R., León, L., Pineda, G., Castroviejo, I.P., et al. (2006). Juvenile myoclonic epilepsy subsyndromes: family studies and long-term follow-up. Brain *129*, 1269–1280.

Matsuzaki, H., Loi, H., Dong, S., Tsai, Y.-Y., Fang, J., Law, J., Di, X., Liu, W.-M., Yang, G., Liu, G., et al. (2004). Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. Genome Res. *14*, 414–425.

McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P.A., and Hirschhorn, J.N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat. Rev. Genet. *9*, 356–369.

McKinsey, G.L., Lindtner, S., Trzcinski, B., Visel, A., Pennacchio, L.A., Huylebroeck, D., Higashi, Y., and Rubenstein, J.L.R. (2013). Dlx1&2-dependent expression of Zfhx1b (Sip1, Zeb2) regulates the fate switch between cortical and striatal interneurons. Neuron *77*, 83–98.

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika *12*, 153–157.

Mefford, H.C., and Eichler, E.E. (2009). Duplication hotspots, rare genomic disorders, and common disease. Curr. Opin. Genet. Dev. *19*, 196–204.

Mefford, H.C., and Mulley, J.C. (2010). Genetically complex epilepsies, copy number variants and syndrome constellations. Genome Med *2*, 71.

Mefford, H.C., Sharp, A.J., Baker, C., Itsara, A., Jiang, Z., Buysse, K., Huang, S., Maloney, V.K., Crolla, J.A., Baralle, D., et al. (2008). Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. N. Engl. J. Med. *359*, 1685–1699.

Mefford, H.C., Muhle, H., Ostertag, P., von Spiczak, S., Buysse, K., Baker, C., Franke, A., Malafosse, A., Genton, P., Thomas, P., et al. (2010). Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. PLoS Genet. *6*, e1000962.

Meisler, M.H., Kearney, J., Ottman, R., and Escayg, A. (2001). Identification of epilepsy genes in human and mouse. Annu. Rev. Genet. *35*, 567–588.

Menashe, I., Rosenberg, P.S., and Chen, B.E. (2008). PGA: power calculator for case-control genetic association analyses. BMC Genet. *9*, 36.

Michailidou, K., Hall, P., Gonzalez-Neira, A., Ghoussaini, M., Dennis, J., Milne, R.L., Schmidt, M.K., Chang-Claude, J., Bojesen, S.E., Bolla, M.K., et al. (2013). Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat. Genet. *45*, 353–361, 361e1–2.

Miller, D.T., Shen, Y., Weiss, L.A., Korn, J., Anselm, I., Bridgemohan, C., Cox, G.F., Dickinson, H., Gentile, J., Harris, D.J., et al. (2009). Microdeletion/duplication at 15q13.2q13.3 among individuals with features of autism and other neuropsychiatric disorders. J. Med. Genet. *46*, 242–248.

Missler, M., Zhang, W., Rohlmann, A., Kattenstroth, G., Hammer, R.E., Gottmann, K., and Südhof, T.C. (2003). Alpha-neurexins couple Ca2+ channels to synaptic vesicle exocytosis. Nature *423*, 939–948.

Moeller, F., Maneshi, M., Pittau, F., Gholipour, T., Bellec, P., Dubeau, F., Grova, C., and Gotman, J. (2011). Functional connectivity in patients with idiopathic generalized epilepsy. Epilepsia *52*, 515–522.

Møller, R.S., Weber, Y.G., Klitten, L.L., Trucks, H., Muhle, H., Kunz, W.S., Mefford, H.C., Franke, A., Kautza, M., Wolf, P., et al. (2013). Exon-disrupting deletions of NRXN1 in idiopathic generalized epilepsy. Epilepsia *54*, 256–264.

Mullen, S.A., Marini, C., Suls, A., Mei, D., Della Giustina, E., Buti, D., Arsov, T., Damiano, J., Lawrence, K., De Jonghe, P., et al. (2011). Glucose transporter 1 deficiency as a treatable cause of myoclonic astatic epilepsy. Arch. Neurol. *68*, 1152–1155.

Mulley, J.C., Scheffer, I.E., Harkin, L.A., Berkovic, S.F., and Dibbens, L.M. (2005). Susceptibility genes for complex epilepsy. Hum. Mol. Genet. *14 Spec No. 2*, R243–249.

Mullins, G.M., O'sullivan, S.S., Neligan, A., McCarthy, A., McNamara, B., Galvin, R.J., and Sweeney, B.J. (2007). A study of idiopathic generalised epilepsy in an Irish population. Seizure *16*, 204–210.

Need, A.C., Ge, D., Weale, M.E., Maia, J., Feng, S., Heinzen, E.L., Shianna, K.V., Yoon, W., Kasperaviciūte, D., Gennarelli, M., et al. (2009). A genome-wide investigation of SNPs and CNVs in schizophrenia. PLoS Genet. *5*, e1000373.

De Nijs, L., Léon, C., Nguyen, L., Loturco, J.J., Delgado-Escueta, A.V., Grisar, T., and Lakaye, B. (2009). EFHC1 interacts with microtubules to regulate cell division and cortical development. Nat. Neurosci. *12*, 1266–1274.

De Nijs, L., Wolkoff, N., Coumans, B., Delgado-Escueta, A.V., Grisar, T., and Lakaye, B. (2012). Mutations of EFHC1, linked to juvenile myoclonic epilepsy, disrupt radial and tangential migrations during brain development. Hum. Mol. Genet. *21*, 5106–5117.

Nordli, D.R., Jr (2005). Idiopathic generalized epilepsies recognized by the International League Against Epilepsy. Epilepsia *46 Suppl 9*, 48–56.

Olsen, R.W., DeLorey, T.M., Gordey, M., and Kang, M.H. (1999). GABA receptor function and epilepsy. Adv Neurol *79*, 499–510.

Ottman, R. (2005). Analysis of genetically complex epilepsies. Epilepsia *46 Suppl 10*, 7–14.

Ottman, R., Annegers, J.F., Hauser, W.A., and Kurland, L.T. (1989). Seizure risk in offspring of parents with generalized versus partial epilepsy. Epilepsia *30*, 157–161.

Ottman, R., Lee, J.H., Hauser, W.A., and Risch, N. (1998). Are generalized and localization-related epilepsies genetically distinct? Arch. Neurol. *55*, 339–344.

Ottman, R., Hirose, S., Jain, S., Lerche, H., Lopes-Cendes, I., Noebels, J.L., Serratosa, J., Zara, F., and Scheffer, I.E. (2010). Genetic testing in the epilepsies--report of the ILAE Genetics Commission. Epilepsia *51*, 655–670.

Pagnamenta, A.T., Wing, K., Sadighi Akha, E., Knight, S.J.L., Bölte, S., Schmötzer, G., Duketis, E., Poustka, F., Klauck, S.M., Poustka, A., et al. (2009). A 15q13.3 microdeletion segregating with autism. Eur. J. Hum. Genet. *17*, 687–692.

Pal, D.K., Evgrafov, O.V., Tabares, P., Zhang, F., Durner, M., and Greenberg, D.A. (2003). BRD2 (RING3) is a probable major susceptibility gene for common juvenile myoclonic epilepsy. Am. J. Hum. Genet. *73*, 261–270.

Palmatier, M.A., Pakstis, A.J., Speed, W., Paschou, P., Goldman, D., Odunsi, A., Okonofua, F., Kajuna, S., Karoma, N., Kungulilo, S., et al. (2004). COMT haplotypes suggest P2 promoter region relevance for schizophrenia. Mol. Psychiatry *9*, 859–870.

Pandolfo, M. (2011). Genetics of epilepsy. Semin Neurol *31*, 506–518.

Papale, L.A., Beyer, B., Jones, J.M., Sharkey, L.M., Tufik, S., Epstein, M., Letts, V.A., Meisler, M.H., Frankel, W.N., and Escayg, A. (2009). Heterozygous mutations of the voltage-gated sodium channel SCN8A are associated with spike-wave discharges and absence epilepsy in mice. Hum. Mol. Genet. *18*, 1633–1641.

Patterson, N., Price, A.L., and Reich, D. (2006). Population Structure and Eigenanalysis. PLoS Genet *2*.

Perrone, M.D., Rocca, M.S., Bruno, I., Faletra, F., Pecile, V., and Gasparini, P. (2012). De novo 911 Kb interstitial deletion on chromosome 1q43 in a boy with mental retardation and short stature. European Journal of Medical Genetics *55*, 117–119.

Petersen, A.K., Ahmad, A., Shafiq, M., Brown-Kipphut, B., Fong, C.-T., and Anwar Iqbal, M. (2013). Deletion 1q43 encompassing only CHRM3 in a patient with autistic disorder. Eur J Med Genet *56*, 118–122.

Pinto, D., Westland, B., de Haan, G.-J., Rudolf, G., da Silva, B.M., Hirsch, E., Lindhout, D., Trenité, D.G.A.K.-N., and Koeleman, B.P.C. (2005). Genome-wide linkage scan of epilepsy-related photoparoxysmal electroencephalographic response: evidence for linkage on chromosomes 7q32 and 16p13. Hum. Mol. Genet. *14*, 171–178.

Poduri, A., and Lowenstein, D. (2011a). Epilepsy genetics--past, present, and future. Curr. Opin. Genet. Dev. *21*, 325–332.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909.

Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R., and Willer, C.J. (2010). LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics *26*, 2336–2337.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

Rajcan-Separovic, E., Harvard, C., Liu, X., McGillivray, B., Hall, J.G., Qiao, Y., Hurlburt, J., Hildebrand, J., Mickelson, E.C.R., Holden, J.J.A., et al. (2007). Clinical and molecular cytogenetic characterisation of a newly recognised microdeletion syndrome involving 2p15-16.1. J. Med. Genet. *44*, 269–276.

Ratnapriya, R., Vijai, J., Kadandale, J.S., Iyer, R.S., Radhakrishnan, K., and Anand, A. (2010). A locus for juvenile myoclonic epilepsy maps to 2q33-q36. Hum. Genet. *128*, 123–130.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., et al. (2006). Global variation in copy number in the human genome. Nature *444*, 444–454.

Rees, M.I. (2010). The genetics of epilepsy--the past, the present and future. Seizure *19*, 680–683.

Reid, C.A., Berkovic, S.F., and Petrou, S. (2009). Mechanisms of human inherited epilepsies. Prog. Neurobiol. *87*, 41–57.

Reid, C.A., Phillips, A.M., and Petrou, S. (2012). HCN channelopathies: pathophysiology in genetic epilepsy and therapeutic implications. Br. J. Pharmacol. *165*, 49–56.

Rincon, G., Weber, K.L., Eenennaam, A.L.V., Golden, B.L., and Medrano, J.F. (2011). Hot topic: performance of bovine high-density genotyping platforms in Holsteins and Jerseys. J. Dairy Sci. *94*, 6116–6121.

Rio, C.A.C., Lawrence, J.J., Tricoire, L., Erdelyi, F., Szabo, G., and McBain, C.J. (2010). M3 Muscarinic Acetylcholine Receptor Expression Confers Differential Cholinergic Modulation to Neurochemically Distinct Hippocampal Basket Cell Subtypes. J. Neurosci. *30*, 6011–6024.

Risch, N., and Merikangas, K. (1996). The future of genetic studies of complex human diseases. Science *273*, 1516–1517.

Rosenfeld, J.A., Coe, B.P., Eichler, E.E., Cuckle, H., and Shaffer, L.G. (2012). Estimates of penetrance for recurrent pathogenic copy-number variations. Genet. Med.

Rowen, L., Young, J., Birditt, B., Kaur, A., Madan, A., Philipps, D.L., Qin, S., Minx, P., Wilson, R.K., Hood, L., et al. (2002). Analysis of the human neurexin genes: alternative splicing and the generation of protein diversity. Genomics *79*, 587–597.

Rujescu, D., Ingason, A., Cichon, S., Pietiläinen, O.P.H., Barnes, M.R., Toulopoulou, T., Picchioni, M., Vassos, E., Ettinger, U., Bramon, E., et al. (2009). Disruption of the neurexin 1 gene is associated with schizophrenia. Hum. Mol. Genet. *18*, 988–996.

Russo, E., Citraro, R., Davoli, A., Gallelli, L., Di Paola, E.D., and De Sarro, G. (2013). Ameliorating effects of aripiprazole on cognitive functions and depressive-like behavior in a genetic rat model of absence epilepsy and mild-depression comorbidity. Neuropharmacology *64*, 371–379.

Saitsu, H., Kato, M., Mizuguchi, T., Hamada, K., Osaka, H., Tohyama, J., Uruno, K., Kumada, S., Nishiyama, K., Nishimura, A., et al. (2008). De novo mutations in the gene encoding STXBP1 (MUNC18-1) cause early infantile epileptic encephalopathy. Nat. Genet. *40*, 782–788.

Sander, T. (1996). The genetics of idiopathic generalized epilepsy: implications for the understanding of its aetiology. Mol Med Today *2*, 173–180.

Sander, T., Schulz, H., Saar, K., Gennaro, E., Riggio, M.C., Bianchi, A., Zara, F., Luna, D., Bulteau, C., Kaminska, A., et al. (2000). Genome search for susceptibility loci of common idiopathic generalised epilepsies. Hum. Mol. Genet. *9*, 1465–1472.

Scheffer, I.E., and Berkovic, S.F. (1997). Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes. Brain *120 (Pt 3)*, 479–490.

Scheffer, I.E., Grinton, B.E., Heron, S.E., Kivity, S., Afawi, Z., Iona, X., Goldberg-Stern, H., Kinali, M., Andrews, I., Guerrini, R., et al. (2012). PRRT2 phenotypic spectrum includes sporadic and fever-related infantile seizures. Neurology *79*, 2104–2108.

Schizophrenia Consortium (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature *455*, 237–241.

Schubert, J., Paravidino, R., Becker, F., Berger, A., Bebek, N., Bianchi, A., Brockmann, K., Capovilla, G., Dalla Bernardina, B., Fukuyama, Y., et al. (2012). PRRT2 mutations are the major cause of benign familial infantile seizures. Hum. Mutat. *33*, 1439–1443.

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Månér, S., Massa, H., Walker, M., Chi, M., et al. (2004). Large-scale copy number polymorphism in the human genome. Science *305*, 525–528.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., et al. (2007). Strong association of de novo copy number mutations with autism. Science *316*, 445–449.

Sebro, R., and Rogus, J.J. (2010). The power of the Transmission Disequilibrium Test in the presence of population stratification. Eur. J. Hum. Genet. *18*, 1032–1038.

Shaikh, T.H., Gai, X., Perin, J.C., Glessner, J.T., Xie, H., Murphy, K., O'Hara, R., Casalunovo, T., Conlin, L.K., D'Arcy, M., et al. (2009). High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. Genome Res. *19*, 1682–1690.

Sharp, A.J. (2009). Emerging themes and new challenges in defining the role of structural variation in human disease. Human Mutation *30*, 135–144.

Sharp, A.J., Mefford, H.C., Li, K., Baker, C., Skinner, C., Stevenson, R.E., Schroer, R.J., Novara, F., De Gregori, M., Ciccone, R., et al. (2008). A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. Nat. Genet. *40*, 322–328.

Shen, J., Gilmore, E.C., Marshall, C.A., Haddadin, M., Reynolds, J.J., Eyaid, W., Bodell, A., Barry, B., Gleason, D., Allen, K., et al. (2010). Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. Nat. Genet. *42*, 245–249.

Slavotinek, A.M. (2008). Novel microdeletion syndromes detected by chromosome microarrays. Hum. Genet. *124*, 1–17.

Spielman, R.S., McGinnis, R.E., and Ewens, W.J. (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am. J. Hum. Genet. *52*, 506–516.

Spielmann, M., Reichelt, G., Hertzberg, C., Trimborn, M., Mundlos, S., Horn, D., and Klopocki, E. (2011). Homozygous deletion of chromosome 15q13.3 including CHRNA7 causes severe mental retardation, seizures, muscular hypotonia, and the loss of KLF13 and TRPM1 potentially cause macrocytosis and congenital retinal dysfunction in siblings. Eur J Med Genet *54*, e441–445.

Sprecher, E., Ishida-Yamamoto, A., Mizrahi-Koren, M., Rapaport, D., Goldsher, D., Indelman, M., Topaz, O., Chefetz, I., Keren, H., O'brien, T.J., et al. (2005). A mutation in SNAP29, coding for a SNARE protein involved in intracellular trafficking, causes a novel neurocutaneous syndrome characterized by cerebral dysgenesis, neuropathy, ichthyosis, and palmoplantar keratoderma. Am. J. Hum. Genet. *77*, 242–251.

Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O.P.H., Ingason, A., Steinberg, S., Fossdal, R., Sigurdsson, E., Sigmundsson, T., Buizer-Voskamp, J.E., et al. (2008). Large recurrent microdeletions associated with schizophrenia. Nature *455*, 232–236.

Steinberg, S., de Jong, S., Andreassen, O.A., Werge, T., Børglum, A.D., Mors, O., Mortensen, P.B., Gustafsson, O., Costas, J., Pietiläinen, O.P.H., et al. (2011). Common variants at VRK2 and TCF4 conferring risk of schizophrenia. Hum. Mol. Genet. *20*, 4076–4081.

Steinlein, O.K., and Bertrand, D. (2010). Nicotinic receptor channelopathies and epilepsy. Pflugers Arch. *460*, 495–503.

Stewart, A.M., Desmond, D., Kyzar, E., Gaikwad, S., Roth, A., Riehl, R., Collins, C., Monnig, L., Green, J., and Kalueff, A.V. (2012). Perspectives of zebrafish models of epilepsy: what, how and where next? Brain Res. Bull. *87*, 135–143.

Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., Grassi, A. de, Lee, C., et al. (2007). Relative Impact of Nucleotide and Copy Number Variation on Gene Expression Phenotypes. Science *315*, 848–853.

Südhof, T.C. (2008). Neuroligins and neurexins link synaptic function to cognitive disease. Nature *455*, 903–911.

Suls, A., Mullen, S.A., Weber, Y.G., Verhaert, K., Ceulemans, B., Guerrini, R., Wuttke, T.V., Salvo-Vargas, A., Deprez, L., Claes, L.R.F., et al. (2009). Early-onset absence epilepsy caused by mutations in the glucose transporter GLUT1. Ann. Neurol. *66*, 415–419.

Suzuki, T., Delgado-Escueta, A.V., Aguan, K., Alonso, M.E., Shi, J., Hara, Y., Nishida, M., Numata, T., Medina, M.T., Takeuchi, T., et al. (2004). Mutations in EFHC1 cause juvenile myoclonic epilepsy. Nat. Genet. *36*, 842–849.

Tan, N.C.K., and Berkovic, S.F. (2010). The Epilepsy Genetic Association Database (epiGAD): analysis of 165 genetic association studies, 1996-2008. Epilepsia *51*, 686–689.

Tan, H.O., Reid, C.A., Chiu, C., Jones, M.V., and Petrou, S. (2008). Increased thalamic inhibition in the absence seizure prone DBA/2J mouse. Epilepsia *49*, 921–925.

Tan, N.C.K., Mulley, J.C., and Berkovic, S.F. (2004). Genetic association studies in epilepsy: “the truth is out there.”Epilepsia *45*, 1429–1442.

Tauer, U., Lorenz, S., Lenzen, K.P., Heils, A., Muhle, H., Gresch, M., Neubauer, B.A., Waltz, S., Rudolf, G., Mattheisen, M., et al. (2005). Genetic dissection of photosensitivity and its relation to idiopathic generalized epilepsy. Ann. Neurol. *57*, 866–873.

Tellez-Zenteno, J.F., Patten, S.B., Jetté, N., Williams, J., and Wiebe, S. (2007). Psychiatric comorbidity in epilepsy: a population-based analysis. Epilepsia *48*, 2336–2344.

Thiele, H., and Nürnberg, P. (2005). HaploPainter: a tool for drawing pedigrees with complex haplotypes. Bioinformatics *21*, 1730–1732.

Thomas, D.C., and Witte, J.S. (2002). Point: Population Stratification: A Problem for Case-Control Studies of Candidate-Gene Associations? Cancer Epidemiol Biomarkers Prev *11*, 505–512.

Thomas, E.A., Reid, C.A., Berkovic, S.F., and Petrou, S. (2009). Prediction by modeling that epilepsy may be caused by very small functional changes in ion channels. Arch. Neurol. *66*, 1225–1232.

Tropeano, M., Ahn, J.W., Dobson, R.J.B., Breen, G., Rucker, J., Dixit, A., Pal, D.K., McGuffin, P., Farmer, A., White, P.S., et al. (2013). Male-biased autosomal effect of 16p13.11 copy number variation in neurodevelopmental disorders. PLoS ONE *8*, e61365.

Turnbull, J., Lohi, H., Kearney, J.A., Rouleau, G.A., Delgado-Escueta, A.V., Meisler, M.H., Cossette, P., and Minassian, B.A. (2005). Sacred disease secrets revealed: the genetics of human epilepsy. Hum. Mol. Genet. *14*, 2491–2500.

Tuzun, E., Sharp, A.J., Bailey, J.A., Kaul, R., Morrison, V.A., Pertz, L.M., Haugen, E., Hayden, H., Albertson, D., Pinkel, D., et al. (2005). Fine-scale structural variation of the human genome. Nat. Genet. *37*, 727–732.

Ullmann, R., Turner, G., Kirchhoff, M., Chen, W., Tonge, B., Rosenberg, C., Field, M., Vianna-Morgante, A.M., Christie, L., Krepischi-Santos, A.C., et al. (2007). Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. Hum. Mutat. *28*, 674–682.

Vassos, E., Collier, D.A., Holden, S., Patch, C., Rujescu, D., St Clair, D., and Lewis, C.M. (2010). Penetrance for copy number variants associated with schizophrenia. Hum. Mol. Genet. *19*, 3477–3481.

Veeramah, K.R., O'Brien, J.E., Meisler, M.H., Cheng, X., Dib-Hajj, S.D., Waxman, S.G., Talwar, D., Girirajan, S., Eichler, E.E., Restifo, L.L., et al. (2012). De novo pathogenic SCN8A mutation identified by whole-genome sequencing of a family quartet affected by infantile epileptic encephalopathy and SUDEP. Am. J. Hum. Genet. *90*, 502–510.

Viera, A.J. (2008). Odds Ratios and Risk Ratios: What's the Difference and Why Does It Matter? Southern Medical Journal *101*, 730–734.

Vitko, I., Chen, Y., Arias, J.M., Shen, Y., Wu, X.-R., and Perez-Reyes, E. (2005). Functional characterization and neuronal modeling of the effects of childhood absence epilepsy variants of CACNA1H, a T-type calcium channel. J. Neurosci. *25*, 4844–4855.

Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor, R.M., Blencowe, B.J., and Geschwind, D.H. (2011). Transcriptomic analysis of autistic brain reveals convergent molecular pathology. Nature *474*, 380–384.

Völzke, H., Alte, D., Schmidt, C.O., Radke, D., Lorbeer, R., Friedrich, N., Aumann, N., Lau, K., Piontek, M., Born, G., et al. (2011). Cohort profile: the study of health in Pomerania. Int J Epidemiol *40*, 294–307.

Vorstman, J.A.S., and Ophoff, R.A. (2013). Genetic causes of developmental disorders. Curr. Opin. Neurol. *26*, 128–136.

Wallace, R.H., Marini, C., Petrou, S., Harkin, L.A., Bowser, D.N., Panchal, R.G., Williams, D.A., Sutherland, G.R., Mulley, J.C., Scheffer, I.E., et al. (2001). Mutant GABA(A) receptor gamma2-subunit in childhood absence epilepsy and febrile seizures. Nat. Genet. *28*, 49–52.

Wang, D., Sun, Y., Stang, P., Berlin, J.A., Wilcox, M.A., and Li, Q. (2009). Comparison of methods for correcting population stratification in a genome-wide association study of rheumatoid arthritis: principal-component analysis versus multidimensional scaling. BMC Proc *3 Suppl 7*, S109.

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F.A., Hakonarson, H., and Bucan, M. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. *17*, 1665–1674.

Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M.A.R., Green, T., et al. (2008). Association between microdeletion and microduplication at 16p11.2 and autism. N. Engl. J. Med. *358*, 667–675.

Wichmann, H.-E., Gieger, C., and Illig, T. (2005). KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen *67 Suppl 1*, S26–30.

Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190–2191.

Winawer, M., Ottman, R., and Rabinowitz, D. (2002). Concordance of disease form in kindreds ascertained through affected individuals. Stat Med *21*, 1887–1897.

Winawer, M.R., Marini, C., Grinton, B.E., Rabinowitz, D., Berkovic, S.F., Scheffer, I.E., and Ottman, R. (2005). Familial clustering of seizure types within the idiopathic generalized epilepsies. Neurology *65*, 523–528.

Wirrell, E.C., Camfield, C.S., Camfield, P.R., Gordon, K.E., and Dooley, J.M. (1996). Long-term prognosis of typical childhood absence epilepsy: remission or progression to juvenile myoclonic epilepsy. Neurology *47*, 912–918.

Wong, M. (2010). Juvenile Myoclonic Epilepsy: Is It an Idiopathic Epilepsy Caused by a Malformation of Cortical Development? Epilepsy Curr *10*, 69–71.

Yalçın, O. (2012a). Genes and molecular mechanisms involved in the epileptogenesis of idiopathic absence epilepsies. Seizure *21*, 79–86.

Yu, F.H., Mantegazza, M., Westenbroek, R.E., Robbins, C.A., Kalume, F., Burton, K.A., Spain, W.J., McKnight, G.S., Scheuer, T., and Catterall, W.A. (2006). Reduced sodium current in GABAergic interneurons in a mouse model of severe myoclonic epilepsy in infancy. Nat. Neurosci. *9*, 1142–1149.

Zambrelli, E., and Canevini, M.P. (2011). Pre- and post-dormitum epilepsies: idiopathic generalized epilepsies. Sleep Med. *12 Suppl 2*, S17–21.

Zhao, W., Li, Q., Ayers, S., Gu, Y., Shi, Z., Zhu, Q., Chen, Y., Wang, H.Y., and Wang, R.-F. (2013). Jmjd3 Inhibits Reprogramming by Upregulating Expression of INK4a/Arf and Targeting PHF20 for Ubiquitination. Cell *152*, 1037–1050.

6 Supplementary material

6.1 Overview of seizure types

TABLE 6-1| Overview of seizure types

Self-limited epileptic seizures

- I. Generalized onset
  - A. Seizures with tonic and/or clonic manifestations
    - 1. Tonic-clonic seizures
    - 2. Clonic seizures
    - 3. Tonic seizures
  - B. Absences
    - 1. Typical absnces
    - 2. Atypical absences
    - 3. Myoclonic absences
  - C. Myoclonic seizure types
    - 1. Myoclonic seizures
    - 2. Myoclonic astatic seizures
    - 3. Eyelid myoclonia
  - D. Epileptic spasms
  - E. Atonic seizures
- II. Focal onset (partial)
  - A. Local
    - 1. Neocortical
      - a. Without local spread
        - i Focal clonic seizures
        - ii Focal myoclonic seizures
        - iii Inhibitory motor seizures
        - iv Focal sensory seizures with elementary symptoms
        - v Aphasic seizures
      - b. With local spread
        - i Jacksonian march seizures
        - ii Focal (asymmetrical) tonic seizures
        - iii Focal sensory seizures with experiential symptoms
    - 2. Hippocampal and parahippocampal
  - B. With ipsilateral propagation to:
    - 1. Neocortical areas (includes hemiclonic seizures)
    - 2. Limbic areas (includes gelastic seizures)
  - C. With contralateral spread to:
    - 1. Neocortical areas (hyperkinetic seizures)
    - 2. Limbic areas (dyscognitive seizures with or without automatisms [psychomotor])
  - D. Secondarily generalized
    - 1. Tonic-clonic seizures
    - 2. Absence seizures
    - 3. Epileptic spasms (unverified)
- III. Neonatal seizures

Table continued on page 105. Modified from Engel et al., 2006.

TABLE 6-1| Overview of seizure types (continued)

Status epilepticus

I. Epilepsia partialis continua (EPC)
- A. As occurs with Rasmussen syndrome
- B. As occurs with focal lesions
- C. As a component of inborn errors of metabolism

II. Supplementary motor area (SMA) status epilepticus

III. Aura continua

IV. Dyscognitive focal (psychomotor, complex partial) status epilepticus
- A. Mesial temporal
- B. Neocortical

V. Tonic-clonic status epilepticus

VI. Absence status epilepticus
- A. Typical and atypical absence status epilepticus
- B. Myoclonic absence status epilepticus

VII. Myoclonic status epilepticus

VIII. Tonic status epilepticus

IX. Subtle status epilepticus

Modified from Engel et al., 2006.

6.2 Genomic regions excluded from PCA

Table 6-2 | Genomic regions excluded from PCA

Chr. region	Start bp	End bp
1p33-p32.3	48,277,980	52,297,979
2p11.2-q11.2	86,078,342	101,051,482
2q21.2-q22.1	134,656,268	138,176,268
2q32.1-q32.2	183,164,494	190,184,494
3p21.31	47,514,996	50,034,996
3p12.2-p12.1	83,407,310	86,927,310
3p11.2-q11.2	88,907,310	96,027,310
3q29	195,387,785	197,396,290
5p12-q11.2	43,954,243	51,474,243
5q21.1	97,962,100	100,482,101
5q23.3-q31.1	128,962,101	131,982,101
5q31.2	135,462,101	138,482,101
6p22.2-p21.32	24,882,021	33,402,022
6p12.1-q12	56,882,041	63,952,041
6q24.1	139,948,307	14,246,8307
7p11.2-q11.22a	55,215,791	66,565,850
8p23.1	7,178,552	12,452,658
8p11.21-q11.21	42,870,843	49,847,447
8q23.2-q23.3	111,920,824	114,940,824
10p11.21-q11.21	36,949,994	43,689,994
11p11.2-q12.1	45,033,424	57,253,424
11q14.2-q14.3	87,850,352	90,870,352
12p11.21-q12	33,098,733	41,723,733
12q24.11-q24.13	111,027,280	113,547,280
15q13.2-q13.3	30,726,915	32,825,174
15q24.1-q24.2	74,354,360	75,579,130
17q12	34,804,328	36,329,039
17q21.31-q21.32	43,534,138	44,643,937
20q11.22-q11.23	32,526,339	35,076,586

Known regions with high long range LD excluded from PCA, collectively from Fellay et al., 2007; Price et al., 2008; Deng et al., 2008; Antonacci et al., 2009; Kasperaviciūte et al., 2010. Modified from [Molekulargenetische Exploration der idiopathisch generalisierten Epilepsien], Leu 2012. Physical positions in NCBI built 37.7, hg19.

6.3 QQ-plots of genome-wide associations studies

6.3.1 QQ-plots of GWAS in GGE

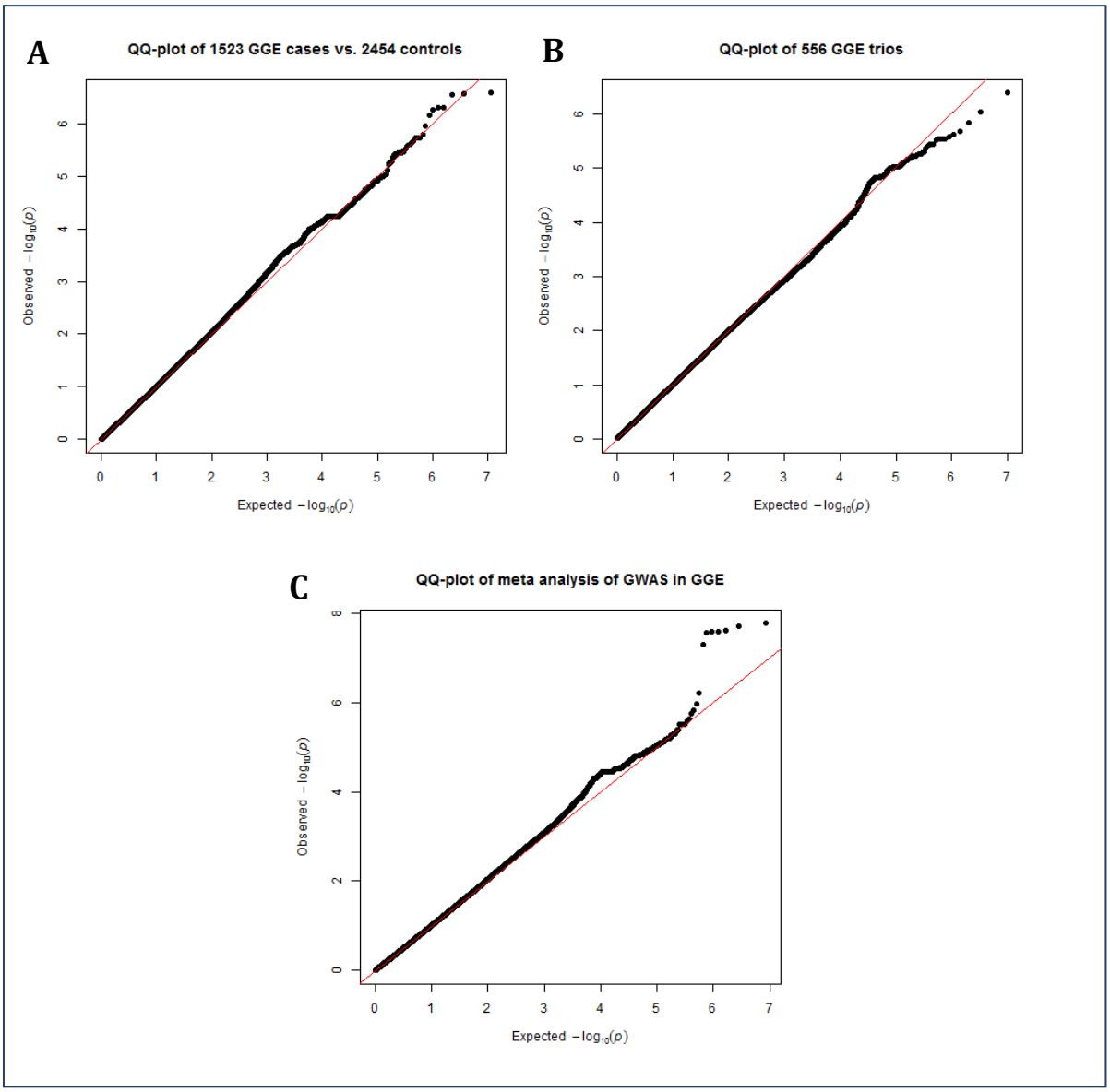


Figure 6-1 | QQ-plots of GWAS in GGE: A) LMM, 1,523 GGE cases vs. 2,454 controls; B) TDT, 566 GGE trios; C) Genome-wide meta-analysis of GGE.

6.3.2 QQ-plots of GWAS in GAE

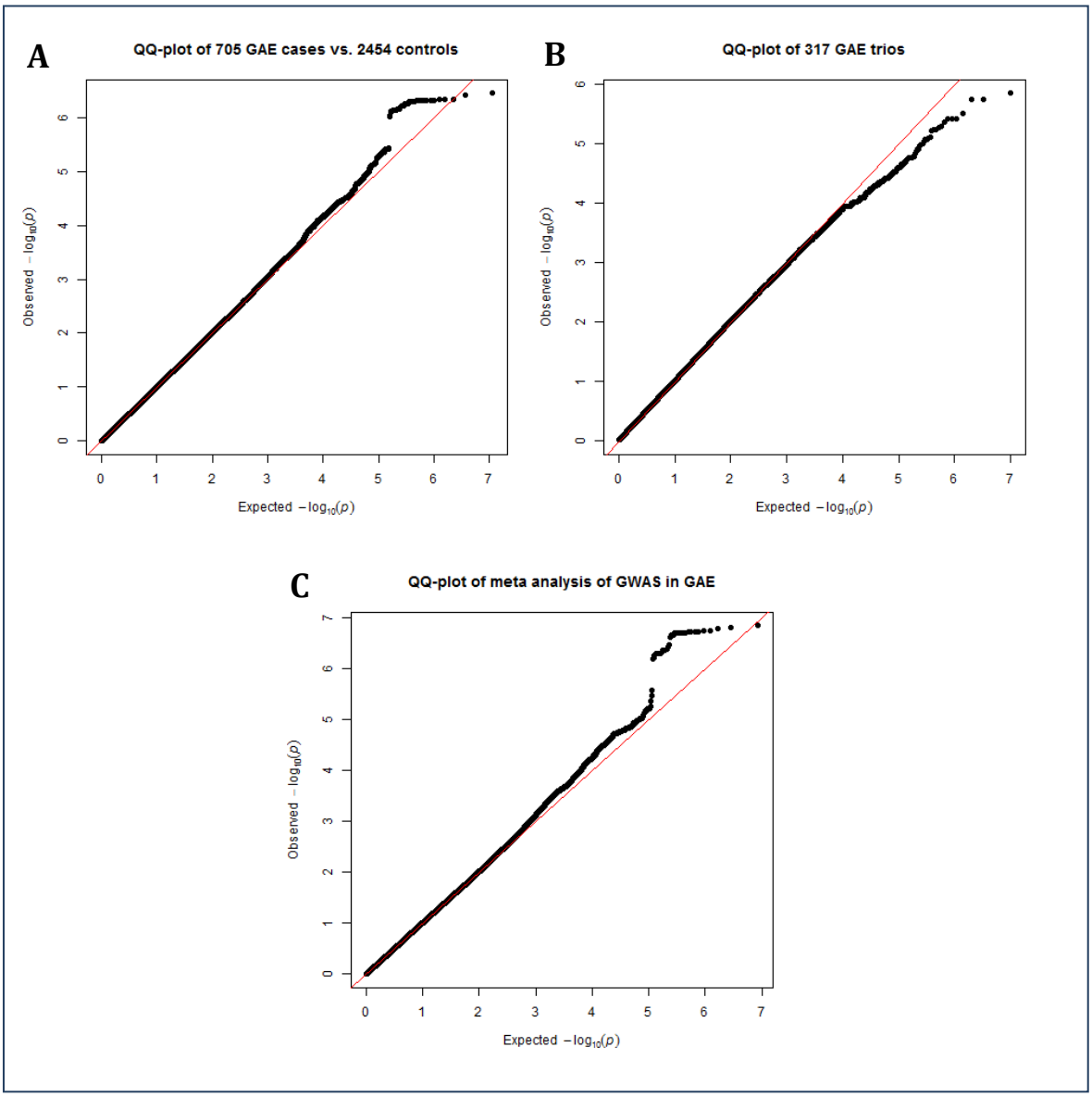


Figure 6-2 | QQ-plots of GWAS in GAE: A) LMM, 705 GGE cases vs. 2,454 controls; B) TDT, 317 GAE trios; C) Genome-wide meta-analysis of GGE.

6.3.3 QQ-plots of GWAS in JME

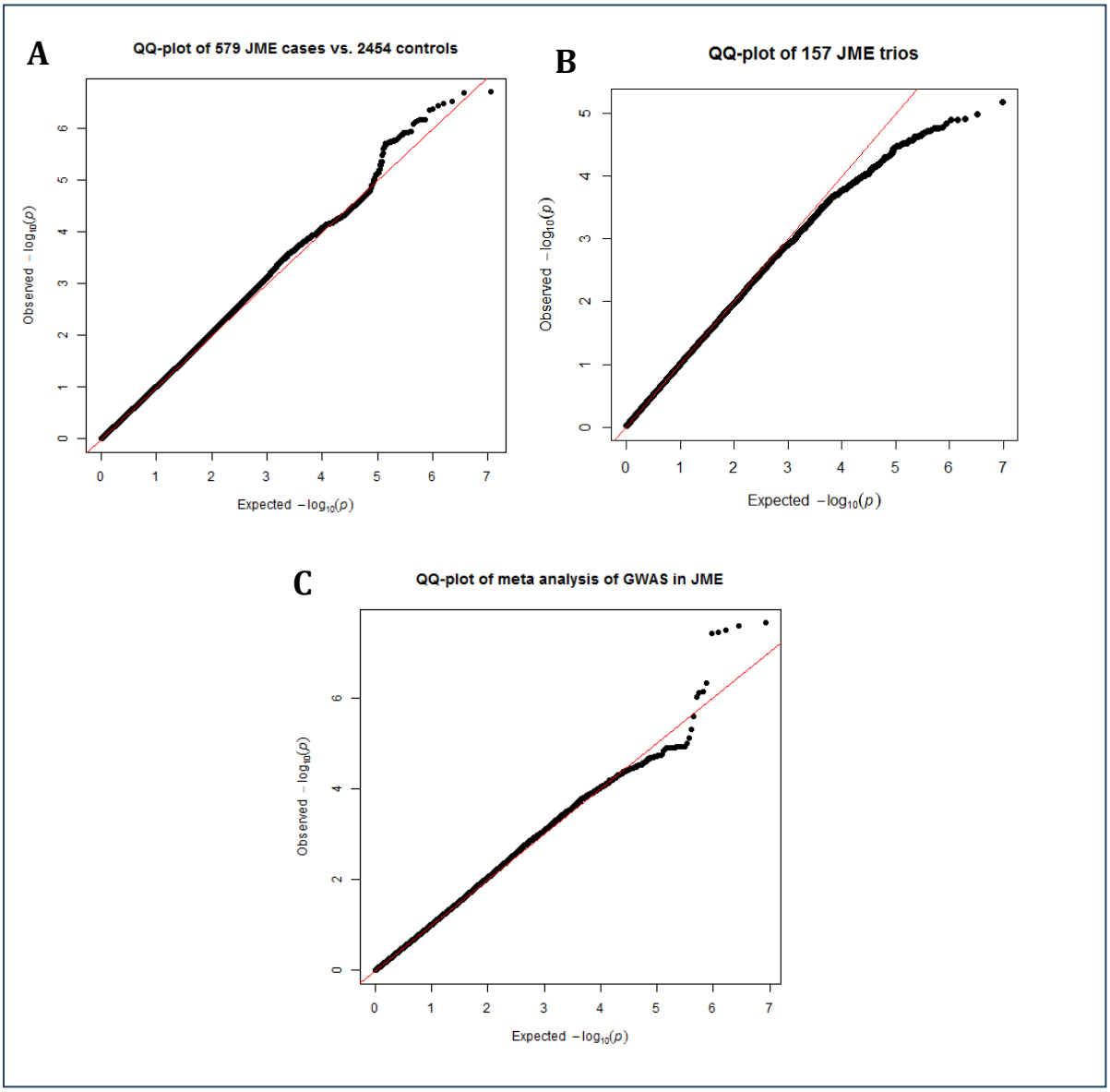


Figure 6-3 | QQ-plots of GWAS in JME: A) LMM, 579 JME cases vs. 2,454 controls; B) TDT, 157 JME trios; C) Genome-wide meta-analysis of GGE.

6.4 Samples with recurrent microdeletions

Table 6-3 | Samples with recurrent microdeletions

Sample_ID	Sex	Chromosomal region	Start	End	Size (kb)	Array	Age-at-onset (years)	Seizure types & comorbidities	IGE syndrome	Inheritance
L1749	F	1q21.1	146,496,383	147,831,171	1,335	SNP6.0	6	abs.	CAE	Unknown
A1207	F	1q21.1	145,930,416	147,789,513	1,859	aCGH	23	abs., GTCS	GAE	Unknown
T17656	M	1q21.1	145,517,561	147,826,789	2,309	Axiom	8	abs.	JAE	*De novo*
D98u3242	F	15q11.2	22,750,301	23,244,372	494	aCGH	4	abs., myocl., GTCS	JME	Maternal
KK0490	F	15q11.2	22,759,242	23,299,047	540	aCGH	15	abs., GTCS	GAE	Maternal
D07u10065	F	15q11.2	22,652,713	23,212,473	560	aCGH	10	abs., GTCS	GAE	Unknown
50Fo1200	F	15q11.2	22,652,713	23,244,372	592	aCGH	Puperty	myocl.	JME	Maternal
119oA	F	15q11.2	22,652,713	23,299,047	646	aCGH	12	GTCS	EGTCS	Unknown
Ao2139	F	15q11.2	22,587,462	23,300,761	713	SNP6.0	7	abs., myocl., legasthenia	JME	Unknown
KK2206A	F	15q11.2	22,652,321	23,419,728	767	aCGH	4	abs.	CAE	Paternal
D09u1781	F	15q11.2	22,317,500	23,237,164	920	SNP6.0	12	myocl., GTCS	JME	Unknown
D03u4524	M	15q11.2	22,652,713	23,077,926	425	aCGH	16	abs., myocl., GTCS	JME	Unknown
46oA	M	15q11.2	22,652,321	23,244,372	592	aCGH	15	abs., myocl., GTCS	JME	Unknown
L1765	M	15q11.2	22,522,798	23,226,468	704	SNP6.0	6	abs.	CAE	Unknown
D09u8942	M	15q11.2	22,673,387	23,487,534	814	SNP6.0	13	myocl.	JME	Unknown
10482	F	15q11.2	22,777,875	23,704,109	926	Axiom	12	abs.	CAE	Paternal
T754	F	15q11.2	22,580,414	23,702,452	1,122	Axiom	16	GTCS	EGTCS	Maternal

SNP6.0: Affymetrix SNP6.0 Genotyping array, Axiom: Affymetrix Axiom® Genotyping array, aCGH: array Comparative Genomic Hybridization, abs.: Absence seizures, myocl.: myoclonic seizures, GTCS: generalized tonic-clonic seizures, EGMA: epilepsy with grand-mals on awakening, FS: Febrile seizures, F: Female, M: Male. Physical positions in NCBI build 37.3, hg19.

Table 6-3 | Samples with recurrent microdeletions (continued)

Sample_ID	Sex	Chromosomal region	Start	End	Size (kb)	Array	Age-at-onset (years)	Seizure types & comorbidities	IGE syndrome	Inheritance
T17044	F	15q11.2	22,580414	23,704,109	1,123	Axiom	10	abs.	CAE	Maternal
T17060	F	15q11.2	22,580414	23,704,109	1,124	Axiom	8	abs.	CAE	Paternal
23P550TR	F	15q11.2	22,763396	23,826,761	1,063	Axiom	13	abs.	JAE	*De novo*
23P785GC	M	15q11.2	22,580414	23,727,655	1,147	Axiom	15	abs., GTCS	JAE	Paternal
157oD	F	15q11.2	22,800539	23,046,902	246	Axiom	3	GTCS	EGTCS	Paternal
EG0766	F	15q11.2	22,580414	23,742,298	1,161	Axiom	17	myocl., GTCS	JME	Maternal
L1371	F	15q13.3	30,950733	32,428,279	1,478	aCGH	4	abs., minor developmental delay	GAE	*De novo*
EG0324	F	15q13.3	30,900866	32,512,208	1,611	SNP6.0	10	abs., GTCS	GAE	Unknown
D07u0771	F	15q13.3	30,807930	32,539,525	1,732	SNP6.0	14	abs.	GAE	Maternal
1674	F	15q13.3	30,730691	32,566,499	1,836	SNP6.0	5	abs.	CAE	Unknown
D04u0213	F	15q13.3	30,920612	32,796,323	1,876	SNP6.0	7	abs., myocl., GTCS	CAE	Paternal
276A	F	15q13.3	30,674083	32,648,839	1,975	aCGH	14	myocl.	JME	Paternal
Ao67	F	15q13.3	30,381105	32,489,233	2,108	aCGH	16	myocl., GTCS	JME	Unknown
40281601	F	15q13.3	28,946780	32,707,226	3,760	aCGH	3	abs., GTCS	GAE	Unknown
D10u0630	M	15q13.3	30,920612	32,539,525	1,619	SNP6.0	15	abs., myocl., GTCS	JME	Unknown
60oA	M	15q13.3	30,381105	32,704,448	2,323	aCGH	12	abs., GTCS	GAE	Unknown
EZ1194	M	15q13.3	30,381,105	32,710,965	2,330	aCGH	6	abs., myocl., GTCS	JME	Unknown

SNP6.0: Affymetrix SNP6.0 Genotyping array, Axiom: Affymetrix Axiom® Genotyping array, aCGH: array Comparative Genomic Hybridization, abs.: Absence seizures, myocl.: myoclonic seizures, GTCS: generalized tonic-clonic seizures, EGMA: epilepsy with grand-mals on awakening, FS: Febrile seizures, F: Female, M: Male. Physical positions in NCBI build 37.3, hg19.

Table 6-3 | Samples with recurrent microdeletions (continued)

Sample_ID	Sex	Chromosomal region	Start	End	Size (kb)	Array	Age-at-onset (years)	Seizure types & comorbidities	IGE syndrome	Inheritance
14P42AE131	F	15q13.3	30,964,572	32,952,193	1,987	Axiom	11	abs.	CAE	*De novo*
14P49AE157	F	15q13.3	30,369,914	32,965,357	2,595	Axiom	10	abs., intellectual disability	CAE	*De novo*
Ao2040	F	16p11.2	29,651,750	30,193,907	542	aCGH	<20	GTCS	EGTCS	Unknown
14P56AE179	F	16p11.2	29,652,,488	30,135,798	483	Axiom	8	abs., EMA	CAE	*De novo*
23P326SS	M	16p11.2	28,944,700	30,432,103	1,487	Axiom	6	GTCS	EGTCS	Maternal
L1808	M	16p11.2	28,944,700	30,466,741	1,522	Axiom	3	GTCS	EGTCS	*De novo*
E1598	F	16p13.11	14,971,403	16,294,706	1,323	SNP6.0	18	GTCS	EGTCS	Unknown
EP204-02	F	16p13.11	14,971,403	16,377,650	1,406	SNP6.0	7	abs.	CAE	Unknown
8P2458	F	16p13.11	14,939,328	16,377,650	1,438	SNP6.0	9	GTCS	EGTCS	Maternal
PYK-043-001	F	16p13.11	14,877,530	16,401,153	1,524	aCGH	5	abs.	CAE	Unknown
KK0987	F	16p13.11	14,791,605	16,401,153	1,610	aCGH	4	abs., FS, learning disability	GAE	*De novo*
Ao1461	F	16p13.11	15,478,837	18,384,481	2,906	aCGH	17	myocl., GTCS, learning disability	JME	Unknown
PYK-017-001	M	16p13.11	15,048,700	16,285,707	1,237	aCGH	7	abs.	CAE	Maternal
KK0204	M	16p13.11	14,877,530	16,859,508	1,982	aCGH	8	abs., GTCS	GAE	*De novo*
14P1F1215	F	16p13.11	14,507,585	16,048,587	1,541	Axiom	2	abs., myocl., GTCS	JME	Paternal

SNP6.0: Affymetrix SNP6.0 Genotyping array, Axiom: Affymetrix Axiom® Genotyping array, aCGH: array Comparative Genomic Hybridization, abs.: Absence seizures, myocl.: myoclonic seizures, GTCS: generalized tonic-clonic seizures, EGMA: epilepsy with grand-mals on awakening, FS: Febrile seizures, F: Female, M: Male. Physical positions in NCBI build 37.3, hg19.

Table 6-3 | Samples with recurrent microdeletions (continued)

Sample_ID	Sex	Chromosomal region	Start	End	Size (kb)	Array	Age-at-onset (years)	Seizure types & comorbidities	IGE syndrome	Inheritance
204AE523	F	16p13.11	14,566,432	16910133	2,343	Axiom	2	abs., EGMA, GTCS	CAE	Maternal
23P049CF	M	16p13.11	14,507,585	16225971	1,718	Axiom	16	myocl., EGMA, FS	JME	Paternal
L1553	F	22q11.21	18,878,339	21456713	2,578	aCGH	20	GTCS	EGTCS	*De novo*
526oS	M	22q11.21	18,878,339	21456713	2,578	aCGH	9	GTCS, minor developmental delay	EGTCS	Unknown
L2148	M	22q11.21	18,876,416	21465835	2,589	SNP6.0	12	GTCS, EGMA	EGMA	Unknown
L1792	M	1q21.1	145,517,561	147826789	2,309	Axiom	–	–	Trio pseudo-control	Unknown
310002235	M	1q21.1	145,932,456	147831171	1,899	SNP6.0	–	–	Control	Unknown
T16070	F	15q11.2	22,580,414	23654207	1,073	Axiom	–	–	Trio pseudo-control	Unknown
SHIP-3663	F	15q11.2	22,673,387	23094223	421	SNP6.0	–	–	Control	Unknown
SHIP-1765	F	15q11.2	22,673,387	23226468	553	SNP6.0	–	–	Control	Unknown
310000142	F	15q11.2	22,673,387	23248421	575	SNP6.0	–	–	Control	Unknown
SHIP-0481	F	15q11.2	22,681,064	23300761	620	SNP6.0	–	–	Control	Unknown
310002175	F	15q11.2	22,673,387	23300761	627	SNP6.0	–	–	Control	Unknown
SHIP-1744	F	15q11.2	22,755,941	23459190	703	SNP6.0	–	–	Control	Unknown
310002138	F	15q11.2	22,673,387	23459190	786	SNP6.0	–	–	Control	Unknown

SNP6.0: Affymetrix SNP6.0 Genotyping array, Axiom: Affymetrix Axiom® Genotyping array, aCGH: array Comparative Genomic Hybridization, abs.: Absence seizures, myocl.: myoclonic seizures, GTCS: generalized tonic-clonic seizures, EGMA: epilepsy with grand-mals on awakening, FS: Febrile seizures, F: Female, M: Male. Physical positions in NCBI build 37.3, hg19.

Table 6-3 | Samples with recurrent microdeletions (continued)

Sample_ID	Sex	Chromosomal region	Start	End	Size (kb)	Array	Age-at-onset (years)	Seizure types & comorbidities	IGE syndrome	Inheritance
310003024	M	15q11.2	22,673,387	23,300,761	627	SNP6.0	–	–	Control	Unknown
SHIP-0777	M	15q11.2	22,681,064	23,389,101	708	SNP6.0	–	–	Control	Unknown
CONSPC2_m_179691	M	15q11.2	22,673,387	23,459,190	786	SNP6.0	–	–	Control	Unknown
SHIP-1348	M	15q11.2	22,586,948	23,417,273	830	SNP6.0	–	–	Control	Unknown
CONBSP_m_213919	M	15q11.2	22,280,465	23,459,190	1,179	SNP6.0	–	–	Control	Unknown
310003270	M	15q11.2	22,129,854	23,459,190	1,329	SNP6.0	–	–	Control	Unknown
SHIP-4057	M	16p11.2	29,517,699	30,191,895	674	SNP6.0	–	–	Control	Unknown
310002282	F	16p13.11	15,053,713	16,377,650	1,324	SNP6.0	–	–	Control	Unknown
CONBSP_m_214168	M	16p13.11	15,491,127	18,361,376	2,870	SNP6.0	–	–	Control	Unknown

SNP6.0: Affymetrix SNP6.0 Genotyping array, Axiom: Affymetrix Axiom® Genotyping array, aCGH: array Comparative Genomic Hybridization, abs.: Absence seizures, myocl.: myoclonic seizures, GTCS: generalized tonic-clonic seizures, EGMA: epilepsy with grand-mals on awakening, FS: Febrile seizures, F: Female, M: Male. Physical positions in NCBI build 37.3, hg19.

6.5 Samples with exon-ablating microdeletions in *NRXN1* and *RBFOX1*

Table 6-4 | Samples with exonic microdeletions in *NRXN1*

Sample_ID	Sex	Chromosomal region	Start	End	Size (kb)	Array	Age-at-onset (years)	Seizure types & comorbidities	IGE syndrome	Inheritance
KK2361	F	2p16.3	50,979,977	51,453,231	333	SNP6.0	5	abs.	CAE	Paternal
8P1844	F	2p16.3	51,080,429	51,682,854	444	SNP6.0	11	abs., GTCS	JAE	Paternal
1696	F	2p16.3	51,152,019	51,209,823	43	SNP6.0	14	myocl., GTCS	JME	De novo
L1748	M	2p16.3	51,185,310	51,379,597	138	SNP6.0	3	abs., learning disability	CAE	Maternal
247oA	M	2p16.3	51,054,002	51,163,990	62	aCGH	16	GTCS	EGTCS	De novo
SHIP-1003	M	2p16.3	50,898,525	51,152,159	254	SNP6.0	-	-	Control	Unknown
SHIP-0946	M	2p16.3	51,022,173	51,163,990	142	SNP6.0	-	-	Control	Unknown

SNP6.0: Affymetrix SNP6.0 Genotyping array, aCGH: array Comparative Genomic Hybridization, abs.: Absence seizures, myocl.: myoclonic seizures, GTCS: generalized tonic-clonic seizures, F: Female, M: Male. Physical positions in NCBI build 37.3, hg19.

Table 6-5 | Samples with exonic microdeletions in *RBFOX1*

Sample_ID	Sex	Chromosomal region	Start	End	Size (kb)	Array	Age-at-onset (years)	Seizure types & comorbidities	IGE syndrome	Inheritance
D07u0680	M	16p13.3	7,035,029	7,137,643	103	SNP6.0	16	myocl., GTCS	JME	Paternal
EG0369	F	16p13.3	5,615,773	6,512,138	896	SNP6.0	3	abs., GTCS	CAE	Maternal
EG0395	M	16p13.3	6,708,812	6,873,788	165	SNP6.0	3	abs., GTCS	CAE	Maternal
EP1613	M	16p13.3	6,796,804	6,865,108	68	SNP6.0	15	abs., FS, EGMA	JAE	Maternal
L2364	M	16p13.3	6,294,809	6,394,343	100	SNP6.0	14	myocl., GTCS	JME	Unknown

SNP6.0: Affymetrix SNP6.0 Genotyping array, abs.: Absence seizures, myocl.: myoclonic seizures, GTCS: generalized tonic-clonic seizures, EGMA: epilepsiy with grand-mals on awakening, FS: Febrile seizures, F: Female, M: Male. Physical positions in NCBI build 37.3, hg19.

Acknowledgements

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study. I would like to extend my sincere thanks to all of them.

I am highly indebted to Prof. Dr. Peter Nürnberg for his guidance and supervision as well as for providing an excellent laboratory environment and also for his support in finalizing this study.

My thanks also go to Prof. Dr. Bernd Wollnik, who kindly took the part as a second referee.

I would like to express my special gratitude and thanks to Dr. habil. Thomas Sander, for providing the subject of this study, as well as the opportunity for collaboration on an international level. His guidance and constant supervision were a major contribution to this study.

My thanks and appreciations go to Dr. Mohammad Reza Toliat, who was a great mentor and support in the laboratory. I also wish to express my gratitude to Dr. Susanne Motameny, Dr. Peter Frommolt, Dr. Wilfried Gunia, Dr. Markus Leber and Dr. Holger Thiele, who all gave me a lot support regarding statistics, bioinformatics, data processing and administration.

My utmost gratitude to Dr. Costin Leu, who told me a lot about statistics and data mining, to Mr. Sebastian Fey, for extending my laboratory expertise, and to Mr. Dennis Lal, who gave me a lot of motivation and assistance. Many thanks as well to Mrs. Ann-Kathrin Ruppert, Mrs. Kerstin Wodecki and Mr. Steffen Pauck for their support and encouragement. Furthermore, I would like to thank the entire team of the CCG for the kind working atmosphere. I apologize to all that have not been mentioned in person or that have been forgotten, I am sure this list is far from complete.

Last but not least I wish to express my gratitude to my parents and to Mrs. Nadine Jansen, who accompanied and encouraged me through all these years.

Eidesstattliche Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Herrn Prof. Dr. P. Nürnberg betreut worden.

_____

Datum, Unterschrift

Curriculum Vitae

Name: Holger Sebastian Trucks

Date of birth: 26 October 1979

Birthplace: Cologne

Citizenship: German

PhD in Biology **2009 – present**

Institute for Genetics, University of Cologne

Subject: "Association mapping of genomic microdeletions and common susceptibility variants predisposing to genetic generalized epilepsies"

Academic studies **2001 – 2008**

Diploma in Biology, University of Cologne

Subject: "Molekulargenetische Grundlagen der Mikrotie"

Internship **2000 - 2001**

One-year internship in graphic design and publishing at rheinsicht GmbH, Cologne

(formerly Hopman Verlag)

Compulsory Community Service **1999 – 2000**

Community service at the Krebelshof e.V. youth center, Cologne

Education **1991 - 1999**

University-entrance diploma at the Leibniz Gymnasium Hackenbroich, Dormagen

Courses & Conferences

Human Genome Analysis: Genetic Analysis of Multifactorial Diseases 2012

Wellcome Trust Advanced Course, Wellcome Trust Genome Campus, Hinxton, UK

BIO.NRW (PhD) Student Convention 2011

BIO.NRW, BayArena Leverkusen

Association Analysis with FAMHAP 2009

Institute for Medical Biometry, Informatics and Epidemiology in Bonn, Germany

List of Publications

EXOSC3 mutations in isolated cerebellar hypoplasia and spinal anterior horn involvement. Biancheri R, Cassandrini D, Pinto F, Trovato R, Di Rocco M, Mirabelli-Badenier M, Pedemonte M, Panicucci C, **Trucks H**, Sander T, Zara F, Rossi A, Striano P, Minetti C, Santorelli FM. J Neurol. 2013 Apr 7. [Epub ahead of print] PMID: 23564332

Rare exonic deletions of the RBFOX1 gene increase risk of idiopathic generalized epilepsy. Lal D, **Trucks H**, Møller RS, Hjalgrim H, Koeleman BP, de Kovel CG, Visscher F, Weber YG, Lerche H, Becker F, Schankin CJ, Neubauer BA, Surges R, Kunz WS, Zimprich F, Franke A, Illig T, Ried JS, Leu C, Nürnberg P, Sander T; EMINet Consortium; EPICURE Consortium.Epilepsia. 2013 Feb;54(2):265-71. doi: 10.1111/epi.12084. Epub 2013 Jan 25.PMID: 23350840

Exon-disrupting deletions of NRXN1 in idiopathic generalized epilepsy. Møller RS, Weber YG, Klitten LL, **Trucks H**, Muhle H, Kunz WS, Mefford HC, Franke A, Kautza M, Wolf P, Dennig D, Schreiber S, Rückert IM, Wichmann HE, Ernst JP, Schurmann C, Grabe HJ, Tommerup N, Stephani U, Lerche H, Hjalgrim H, Helbig I, Sander T; EPICURE Consortium. Epilepsia.2013 Jan 7.doi: 10.1111/epi.12078. [Epub ahead of print] PMID:23294455

Genome-wide association analysis of genetic generalized epilepsies implicates susceptibility loci at 1q43, 2p16.1, 2q22.3 and 17q21.32. EPICURE Consortium; EMINet Consortium, Steffens M, Leu C, Ruppert AK, Zara F, Striano P, Robbiano A,

Capovilla G, Tinuper P, Gambardella A, Bianchi A, La Neve A, Crichiutti G, de Kovel CG, Kasteleijn-NolstTrenité D, de Haan GJ, Lindhout D, Gaus V, Schmitz B, Janz D, Weber YG, Becker F, Lerche H, Steinhoff BJ, Kleefuß-Lie AA, Kunz WS, Surges R, Elger CE, Muhle H, von Spiczak S, Ostertag P, Helbig I, Stephani U, Møller RS, Hjalgrim H, Dibbens LM, Bellows S, Oliver K, Mullen S, Scheffer IE, Berkovic SF, Everett KV, Gardiner MR, Marini C, Guerrini R, Lehesjoki AE, Siren A, Guipponi M, Malafosse A, Thomas P, Nabbout R, Baulac S, Leguern E, Guerrero R, Serratosa JM, Reif PS, Rosenow F, Mörzinger M, Feucht M, Zimprich F, Kapser C, Schankin CJ, Suls A, Smets K, De Jonghe P, Jordanova A, Caglayan H, Yapici Z, Yalcin DA, Baykan B, Bebek N, Ozbek U, Gieger C, Wichmann HE, Balschun T, Ellinghaus D, Franke A, Meesters C, Becker T, Wienker TF, Hempelmann A, Schulz H, Rüschendorf F, Leber M, Pauck SM, **Trucks H**, Toliat MR, Nürnberg P, Avanzini G, Koeleman BP, Sander T. Hum Mol Genet. 2012 Dec 15;21(24):5359-72. doi: 10.1093/hmg/dds373. Epub 2012 Sep 4. PMID: 22949513

Genome-wide linkage meta-analysis identifies susceptibility loci at 2q34 and 13q31.3 for genetic generalized epilepsies. EPICURE Consortium, Leu C, de Kovel CG, Zara F, Striano P, Pezzella M, Robbiano A, Bianchi A, Bisulli F, Coppola A, Giallonardo AT, Beccaria F, Trenité DK, Lindhout D, Gaus V, Schmitz B, Janz D, Weber YG, Becker F, Lerche H, Kleefuss-Lie AA, Hallman K, Kunz WS, Elger CE, Muhle H, Stephani U, Møller RS, Hjalgrim H, Mullen S, Scheffer IE, Berkovic SF, Everett KV, Gardiner MR, Marini C, Guerrini R, Lehesjoki AE, Siren A, Nabbout R, Baulac S, Leguern E, Serratosa JM, Rosenow F, Feucht M, Unterberger I, Covanis A, Suls A, Weckhuysen S, Kaneva R, Caglayan H, Turkdogan D, Baykan B, Bebek N, Ozbek U, Hempelmann A, Schulz H, Rüschendorf F, **Trucks H**, Nürnberg P, Avanzini G, Koeleman BP, Sander T. Epilepsia. 2012 Feb;53(2):308-18. doi: 10.1111/j.1528-1167.2011.03379.x. Epub 2012 Jan 13.PMID: 22242659

Mutations in FKBP10 can cause a severe form of isolated Osteogenesis imperfecta. Steinlein OK, Aichinger E, **Trucks H**, Sander T. BMC Med Genet. 2011 Nov 22;12:152. PMID:22107750

Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. de Kovel CG*, **Trucks H***, Helbig I*, Mefford HC, Baker C, Leu C, Kluck C, Muhle H, von Spiczak S, Ostertag P, Obermeier T, Kleefuss-Lie AA, Hallmann K, Steffens M, Gaus V, Klein KM, Hamer HM, Rosenow F, Brilstra EH, Trenité DK, Swinkels ME, Weber YG, Unterberger I, Zimprich F, Urak L, Feucht M, Fuchs K, Møller RS, Hjalgrim H, De Jonghe P, Suls A, Rückert IM, Wichmann HE, Franke A, Schreiber S, Nürnberg P, Elger CE, Lerche H, Stephani U, Koeleman BP, Lindhout D, Eichler EE, Sander T. Brain. 2010 Jan;133(Pt 1):23-32. Epub 2009 Oct 20. PMID: 19843651 (* sharedfirstauthorship)

Familial and sporadic 15q13.3 microdeletions in idiopathic generalized epilepsy: precedent for disorders with complex inheritance. Dibbens LM, Mullen S, Helbig I, Mefford HC, Bayly MA, Bellows S, Leu C, **Trucks H**, Obermeier T, Wittig M, Franke A, Caglayan H, Yapici Z; EPICURE Consortium, Sander T, Eichler EE, Scheffer IE, Mulley JC, BerkovicSF.HumMol Genet. 2009 Oct 1;18(19):3626-31. Epub 2009 Jul 10. PMID: 19592580