

Molekulargenetische Exploration der idiopathisch generalisierten Epilepsien

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Costin Leu

aus Bukarest

Köln

2012

BERICHTERSTATTER/IN:

PROF. DR. PETER NÜRNBERG

PROF. DR. BRUNHILDE WIRTH

Tag der letzten mündlichen Prüfung:

19.06.2012

INHALTSVERZEICHNIS

Publikation	VI
Zusammenfassung	VII
Abstract.....	IX
Abbildungsverzeichnis	XI
Tabellenverzeichnis	XII
Abkürzungsverzeichnis	XIII
1 EINLEITUNG	1
1.1 Klassifikation epileptischer Syndrome	1
1.2 Genetik von Epilepsiesyndromen.....	4
1.3 Epilepsie als genetisch komplexe Erkrankung	6
1.4 Idiopathische generalisierte Epilepsie	7
1.5 Zielsetzung	9
1.6 Forschungsgrundlage	9
2 METHODEN	11
2.1 Studienteilnehmer	11
2.2 DNA-Aufbereitung und Qualitätskontrolle.....	12
2.2.1 DNA-Konzentrationsbestimmung.....	12
2.2.2 Agarose-Gelelektrophorese	13
2.3 Polymerase-Kettenreaktion.....	14
2.4 SNP-Genotypisierung	14
2.4.1 Affymetrix SNP 6.0	14
2.4.2 GenomeLab SNPstream.....	16
2.4.3 Illumina Infinium.....	18
2.5 Erstellung und Integritätsüberprüfung von Genotyp-Eingabedaten	20
2.6 Hardy-Weinberg-Gleichgewicht.....	21
2.7 Rekombination und Kopplungsungleichgewicht	22
2.8 Populationsstratifizierung	24
2.9 Statistische Tests.....	26
2.10 Fall-Kontroll-Analysen	28
2.10.1 Quotenverhältnis und Relatives Risiko	28
2.10.2 Logistische Regression.....	29
2.10.3 Stratifizierte Assoziationsanalysen	30
2.10.4 Genomic Control.....	30
2.11 Familien-basierte Analysen.....	31

2.11.1	Transmission-Disequilibrium-Test.....	31
2.11.2	Fall-Pseudo-Kontroll-Methode	31
2.11.3	Parametrische Kopplungsanalysen	33
2.11.4	Nichtparametrische Kopplungsanalysen.....	34
2.12	Meta-Analysen.....	35
2.13	Visualisierungen	35
2.14	Bestimmung von Kandidatengenem	35
3	ERGEBNISSE	37
3.1	Genomweite Kopplungsstudie.....	37
3.1.1	Genomweite Kopplungsanalyse in 379 IGE-Multiplexfamilien.....	38
3.1.2	Genomweite Kopplungsanalyse in 235 IAE-Multiplexfamilien.....	40
3.1.3	Genomweite Kopplungsanalyse in 118 JME-Multiplexfamilien.....	40
3.1.4	Zusammenfassung der Kopplungsergebnisse	43
3.2	Genomweite Assoziationsstudie	46
3.2.1	GWAS im IGE-Gesamtkollektiv	48
3.2.2	GWAS in der IAE-Subgruppe	50
3.2.3	GWAS in der JME-Subgruppe	50
3.2.4	Replikation der GWAS-Ergebnisse.....	53
4	DISKUSSION	58
4.1	Genomweite Kopplungsstudie.....	60
4.2	Genomweite Assoziationsstudie	61
4.3	JME-Locus auf 2q32.1-q36.1	62
4.4	IAE-Locus auf 13q31.1-q32.1.....	64
4.5	IGE-Locus auf 17q21.32.....	65
4.6	Ausblick	66
5	ANHANG.....	67
5.1	Kandidatenregionen aus Kopplungsstudien	67
5.2	Qualitätsstatistiken von 21 SNPstream-SNPs	68
6	LITERATURVERZEICHNIS.....	69
	DANKSAGUNG.....	88
	EIDESSTATTLICHE ERKLÄRUNG	90
	CURRICULUM VITAE.....	91

Publikation

Teile dieser Arbeit wurden veröffentlicht:

EPICURE Consortium, Leu, C., de Kovel, C. G., Zara, F., Striano, P., Pezzella, M., Robbiano, A., Bianchi, A., Bisulli, F., Coppola, A., Giallonardo, A. T., Beccaria, F., Trenité, D. K., Lindhout, D., Gaus, V., Schmitz, B., Janz, D., Weber, Y. G., Becker, F., Lerche, H., Kleefuss-Lie, A. A., Hallman, K., Kunz, W. S., Elger, C. E., Muhle, H., Stephani, U., Møller, R. S., Hjalgrim, H., Mullen, S., Scheffer, I. E., Berkovic, S. F., Everett, K. V., Gardiner, M. R., Marini, C., Guerrini, R., Lehesjoki, A. E., Siren, A., Nabbout, R., Baulac, S., Leguern, E., Serratosa, J. M., 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., Rüschen-dorf, F., Trucks, H., Nürnberg, P., Avanzini, G., Koeleman, B. P., Sander, T. (2012). Genome-wide linkage meta-analysis identifies susceptibility loci at 2q34 and 13q31.3 for genetic generalized epilepsies. *Epilepsia*, 53(2), 308-18.

Zusammenfassung

Die idiopathisch generalisierten Epilepsien (IGEs) stellen eine Gruppe von Epilepsiesyndromen dar, die zusammen eine Lebenszeitprävalenz von 0,3% aufweisen und 20-30% aller Epilepsien repräsentieren. Die häufigen, klassischen IGE-Syndrome sind die kindliche Absence-Epilepsie (CAE), die juvenile Absence-Epilepsie (JAE), die juvenile myoklonische Epilepsie (JME) und die Epilepsie mit generalisierten tonisch-klonischen Anfällen (EGTCS). Die Ätiologie dieser vier klinisch überlappenden Subsyndrome ist genetisch determiniert. Molekulargenetische Forschungsansätze sind daher aussichtsreich, einen bedeutenden Beitrag zur Aufklärung der molekularen Mechanismen der Epileptogenese zu leisten. Trotz der hohen Heritabilität von ca. 80%, gelang es bislang nicht, die genetische Architektur der häufigen IGE-Syndrome überzeugend aufzuklären. Lediglich für die seltenen, monogenen Formen der IGEs (ca. 1-2% der Fälle) sind klar prädisponierende Genvarianten identifiziert worden.

In dieser Studie wurden zur Dissektion der genetischen Faktoren bei den häufigen IGE-Syndromen mit genetisch komplexer Disposition zwei verschiedene molekulargenetische Forschungsansätze angewandt. Ziel dieser Strategie war es, die unterschiedlichen genetischen Komponenten des multifaktoriellen polygenen Modells zu erfassen. Mittels genomweiter Kopplungsanalysen in familiären IGE-Formen sollten genetische Faktoren mit relativ hohen Effektstärken identifiziert werden, die zu einer oligonomen Vererbung beitragen. Komplementär sollten mittels genomweiter Assoziationsanalysen genetische Variationen identifiziert werden, die entsprechend der „common disease/common variant“- (CD/CV-) Hypothese meist geringe disponierende Effekte aufweisen, jedoch häufig in der allgemeinen Bevölkerung vorkommen. Darüber hinaus sollte überprüft werden, ob die idiopathischen Absence-Epilepsien (IAE) und JME syndromspezifische Determinanten aufweisen. Die Grundlage für diese Studien war die Rekrutierung der weltweit größten Kollektive von 379 IGE-Multiplexfamilien und 2826 unverwandten IGE-Patienten im Rahmen des europaweiten Forschungsprojektes EPICURE.

In der genomweiten Kopplungsanalyse in 379 IGE-Multiplexfamilien fanden sich suggestive Kopplungshinweise für sechs Regionen (1p36.22, 3p21.1-p13, 5q31.3-q35.1, 13q12.11-q12.13, 13q31.1-q32.1, 19q13.32-q13.42). Für den Chromosomenabschnitt 13q31.1-q32.1, wurde in der Subgruppe mit 235 IAE-Multiplexfamilien eine signifikante Kopplung in der parametrischen Analyse erreicht (rezessives Vererbungsmodell: HLOD = 5,02 bei rs1332470; $\alpha = 0,22$). Die nichtparametrische Kopplungsanalyse in der Subgruppe mit 118 JME-Multiplexfamilien ergab eine signifikante Kopplung in der chromosomalen Region 2q32.1-

q36.1 ($LOD_{NPL} = 3,43$ bei D2S143). Die parametrischen Analysen in der JME-Subgruppe unterstützen den signifikanten Kopplungsbefund ($HLOD > 2,5$ in beiden Vererbungsmodellen bei D2S143). In der chromosomalen Region 5q31.3-q35.1 fanden sich suggestive Kopplungshinweise in beiden Subgruppenanalysen, sowohl in den nichtparametrischen als auch in den parametrischen Kopplungsanalysen mit dem dominanten Vererbungsmodell.

Die Ergebnisse der Kopplungsanalysen unterstützen das oligogene Vererbungsmodell für die familiären IGE-Syndrome. Ein genetischer Risikofaktor in dem Chromosomenabschnitt 5q31.3-q35.1 prädisponiert für ein breites Spektrum von familiären IGE-Syndromen. Interessanterweise befinden sich in dieser Region zwei bekannte Epilepsiegene (*GABRA1*, *GABRG2*), in denen Mutationen bei monogenen IGE-Formen identifiziert wurden. Die zwei Kopplungsloci in den Chromosomenabschnitten 13q31.1-q32.1 und 2q32.1-q36.1 weisen auf subsyndromspezifische Determinanten hin, die differentiell zum Entstehungsrisiko von entweder IAE oder JME beitragen. Die aussichtsreichsten Kandidatengene sind *GPC5* auf 13q31.3 und *ERBB4* auf 2q34.

Die GWAS mit anschließender Replikationsanalyse ergab in der kombinierten Analyse von 2826 IGE-Einzelfällen und 3701 Kontrollen europäischen Ursprungs eine genomweit signifikante Assoziation in der Region 17q21.32 (META-P = $7,9 \times 10^{-9}$ bei rs41336845; OR[T] = 0,76; 95%-KI: 0,70-0,84). In der IAE-Subgruppe mit 1337 IAE-Einzelfällen und 3452 Kontrollen fanden sich suggestive Assoziationen auf Chromosom 2 in den Regionen 2p16.1 (META-P = $1,82 \times 10^{-7}$ bei rs2953442; OR[C] = 1,31; 95%-KI: 1,19-1,45) und 2q22.3 (META-P = $1,53 \times 10^{-7}$ bei rs17741930; OR[C] = 0,69; 95%-KI: 0,60-0,80). In der JME-Subgruppe mit 1067 JME-Einzelfällen und 3287 Kontrollen fand sich eine suggestive Assoziation in der Region 1q43 (META-P = $4,22 \times 10^{-7}$ bei rs1110615; OR[A] = 1,46; 95%-KI: 1,28-1,65). Die genomweit signifikante Assoziation in der chromosomalen Region 17q21.32 mit den häufigen IGE-Syndromen bestätigt erstmals die Existenz eines häufigen Risikofaktors für IGE entsprechend der CD/CV-Hypothese. In der assoziierten Region befindet sich das Gen für die Pyridoxamin 5'-Phosphat Oxidase (*PNPO*), das ein sehr aussichtsreiches Kandidatengen darstellt.

Die vorliegenden genomweiten Kopplungs- und Assoziationstudien identifizieren neue IGE-Loci in den Chromosomenabschnitten 2q32.1-q36.1, 13q31.1-q32.1 und 17q21.32. Die positionelle Identifizierung der zugrundeliegenden Epilepsiegene wird weitere Einblicke in die molekularen Mechanismen der Epileptogenese eröffnen und zur Klärung der genetischen Architektur von Epilepsien beitragen.

Abstract

Idiopathic generalized epilepsies (IGEs) represent a group of epilepsy syndromes affecting about 0.3% of the general population and accounting for 20-30% of all epilepsies. The common classical IGE syndromes include childhood absence epilepsy (CAE), juvenile absence epilepsy (JAE), juvenile myoclonic epilepsy (JME), and epilepsy with generalized tonic-clonic seizures. The etiology of these four clinically overlapping subsyndromes is genetically determined. Therefore, molecular genetic studies present promising perspectives to unravel the molecular mechanism of epileptogenesis. Despite a heritability of 80%, the underlying genetic architecture of the common IGE syndromes remains largely unresolved. Until today, predisposing gene variants could be identified only for the rare, monogenic forms of IGE (about 1-2% of the cases).

In the present study, two molecular genetic research approaches were applied to dissect out genetic factors contributing to the common IGE syndromes with complex genetic disposition. The aim of this strategy was to differentiate various genetic components of the multifactorial polygenic model. Genome-wide linkage analysis in multiplex families with IGE was performed to screen for genetic factors with strong effect sizes, which contribute to an oligogenic inheritance. Complementary genome-wide association studies (GWAS) were performed to identify common genetic variations, which according to the “common disease/common variant” (CD/CV) hypothesis show only small effect sizes. In addition, analyses in two syndrome subgroups, stratified by either idiopathic absence epilepsies (IAE) or JME, were also performed to dissect out syndrome-related variants conferring risk to either absence seizures or JME-related myoclonic seizures. A prerequisite of the present studies was the recruitment of the worldwide largest collective of 379 IGE multiplex families and 2826 unrelated IGE patients as part of the European EPICURE research project.

In the genome-wide linkage analysis of 379 multiplex families, suggestive linkage signals were found in six regions (1p36.22, 3p21.1-p13, 5q31.3-q35.1, 13q12.11-q12.13, 13q31.1-q32.1, 19q13.32-q13.42). In the chromosomal region 13q31.1-q32.1, significant parametric linkage was detected in the subgroup with 235 IAE multiplex families (recessive inheritance model: HLOD = 5.02 at rs1332470; $\alpha = 0.22$). The nonparametric linkage analysis in the subgroup with 118 JME multiplex families showed significant linkage in the chromosomal region 2q32.1-q36.1 ($LOD_{NPL} = 3.43$ at D2S143). The parametric linkage analyses in the JME subgroup support the significant linkage signal (HLOD > 2.5 for both inheritance models at D2S143). The linkage finding at 5q31.3-q35.1 was consistently supported by both nonparametric and parametric linkage results across all family groups.

Altogether, the results of the linkage scan support the oligogenic inheritance model for familial IGE syndromes. A genetic risk factor in the chromosomal region 5q31.3-q35.1 predisposes for a wide spectrum of familial IGE syndromes. Interestingly, two known epilepsy genes encoding the GABA_A-receptor α 1 and γ 2 subunits (*GABRA1*, *GABRG2*) are located in this region. The two linkage loci in the chromosomal regions 13q31.1-q32.1 and 2q32.1-q36.1 imply subsyndrome-specific determinants preferentially predisposing to either IAE or JME. The most promising candidate genes are *GPC5* at 13q31.3 and *ERBB4* at 2q34.

Our GWAS, together with the subsequent replication analysis, investigated 2826 IGE cases and 3701 controls of European origin. Joint analysis revealed a genome-wide significant association at 17q21.32 (META-P = 7.9×10^{-9} at rs41336845; OR[T] = 0.76; 95%-CI: 0.70-0.84). Within the IAE subgroup comprising 1337 IAE cases and 3452 controls, two suggestive association signals on chromosome 2 were found at 2p16.1 (META-P = 1.82×10^{-7} at rs2953442; OR[C] = 1.31; 95%-CI: 1.19-1.45) and 2q22.3 (META-P = 1.53×10^{-7} at rs17741930; OR[C] = 0.69; 95%-CI: 0.60-0.80). In the JME subgroup consisting of 1067 JME cases and 3287 controls, one suggestive association was detected at 1q43 (META-P = 4.22×10^{-7} at rs1110615; OR[A] = 1.46; 95%-CI: 1.28-1.65). Consistent with the CD/CV hypothesis, the genome-wide significant association in the chromosomal region 17q21.32 confirms for the first time the existence of a common risk factor for IGE. A promising candidate gene is located within the associated region; it codes for the pyridoxamine 5'-phosphate oxidase (*PNPO*).

In the present genome-wide linkage and association studies, new IGE loci were detected on the chromosomal regions 2q32.1-q36.1, 13q31.1-q32.1 and 17q21.32. The positional identification of the causal epilepsy gene variants will provide new insights in the molecular mechanisms of epileptogenesis and will contribute to the dissection of the genetic architecture of epilepsy.

Abbildungsverzeichnis

Abb. 1-1	Individuelle genetische Suszeptibilität für Epilepsie	6
Abb. 1-2	Überlappung der komplexen Phänotypen der IGEs	8
Abb. 2-1	Flussdiagramm des <i>Genome-Wide Human SNP Nsp/Sty</i> -Assays	15
Abb. 2-2	Flussdiagramm des SNPstream-Assays.....	17
Abb. 2-3	Flussdiagramm des <i>Illumina Infinium II</i> -Assays	19
Abb. 2-4	Weltweite Populationsstruktur abgebildet durch PCA	25
Abb. 2-5	Trio als Fall / Pseudo-Kontrolle	32
Abb. 3-1	Genomweite Kopplungsanalyse in 379 IGE-Multiplexfamilien.....	39
Abb. 3-2	Genomweite Kopplungsanalyse in 235 IAE-Multiplexfamilien	41
Abb. 3-3	Genomweite Kopplungsanalyse in 118 JME-Multiplexfamilien.....	42
Abb. 3-4	MDS von 1549 IGE-Einzenfällen und 2468 Kontrollen.....	47
Abb. 3-5	Detektierbare ORs mit einer Power von 80% bei $P = 5 \times 10^{-8}$	48
Abb. 3-6	Manhattan Plots der GWAS im IGE-Gesamtkollektiv und in Subgruppen	49
Abb. 3-7	Q-Q-Plots der GWAS im IGE-Gesamtkollektiv und den Subgruppen.....	52
Abb. 3-8	Struktur der signifikant assoziierten Region auf 17q21.32.....	56
Abb. 4-1	Identifizierbarkeit von Variationen nach Frequenz und Effektstärke	58

Tabellenverzeichnis

Tabelle 1-1 Beispiel einer möglichen Klassifikation von Epilepsiesyndromen	3
Tabelle 1-2 Konkordanzraten bei Verwandten und Zwillinge von Epilepsiepatienten	4
Tabelle 1-3 Bekannte Gene die zur Ausprägung eines Epilepsiesyndroms prädisponieren	5
Tabelle 2-1 IGE-Kollektive für Assoziationsanalysen	11
Tabelle 2-2 IGE-Multiplexfamilien für Meta-Kopplungsanalyse	11
Tabelle 2-3 Ethnizitäten der gesammelten IGE-Kollektive	12
Tabelle 2-4 Kopplungsungleichgewicht	23
Tabelle 2-5 Bekannte Regionen mit außergewöhnlichem Langstrecken-LD	26
Tabelle 2-6 Entscheidungen des Signifikanztests	27
Tabelle 2-7 Anzahl an Artefakten bei 1 Million unabhängiger Tests	27
Tabelle 2-8 Vierfeldertafel für die Berechnung des ORs	28
Tabelle 2-9 Interpretation von RR und OR	29
Tabelle 2-10 Anzahl übertragener Allele eines biallelischen Markers für n Trios	31
Tabelle 3-1 Analysegruppen der Kopplungsstudie	37
Tabelle 3-2 Klinische Charakterisierung der betroffenen Familienmitglieder	38
Tabelle 3-3 Power der parametrischen Kopplungsanalyse	38
Tabelle 3-4 Suggestive und signifikante LOD_{NPL} -Scores	43
Tabelle 3-5 Suggestive und signifikante HLOD-Scores	43
Tabelle 3-6 SNPs mit $P_{GC} \leq 10^{-5}$ in der logistischen Regression im IGE-Gesamtkollektiv	50
Tabelle 3-7 SNPs mit $P_{GC} \leq 10^{-5}$ in der logistischen Regression in der IAE-Subgruppe	51
Tabelle 3-8 SNPs mit $P_{GC} \leq 10^{-5}$ in der logistischen Regression in der JME-Subgruppe	51
Tabelle 3-9 Replikation der GWAS und Meta-Analyse von 14 SNPs	54
Tabelle 3-10 Odds Ratios der GWAS, der Replikation und der Meta-Analyse	55
Tabelle 5-1 Kandidatenregionen aus Kopplungsstudien zu Epilepsie	67
Tabelle 5-2 Genotypisierte SNPs aus der Replikationsanalyse	68

Abkürzungsverzeichnis

%	Prozent
°C	Grad Celsius
µl	Mikroliter
A	Adenin
Abb	Abbildung
ADNFLE	Autosomal-dominante nächtliche Frontallappenepilepsie
AEDs	"antiepileptic drugs" - antiepileptische Pharmaka
AFBAC	"Affected Family-Based Controls"- (Fall-Pseudo-Kontroll-) Methode
Anz	Anzahl
BD-P	P-Wert des Breslow-Day-Tests
bp	Basenpaar
bzw	beziehungsweise
C	Cytosin
CAE	"childhood absence epilepsy" - kindliche Absence-Epilepsie
CD/CV	"common disease/common variant"
CD/MV	„common disease multiple rare variant“
Chr	Chromosom
cM	Centimorgan
CMH-P	P-Wert des Cochran-Mantel-Haenszel-Tests
CMH-Test	Cochran-Mantel-Haenszel-Test
CNV	„copy number variations“ - Genkopiezahlvarianten
cQC	„contrast Quality Control“
CR	"Call rate" - Genotypisierungsrate
DNA	Desoxyribonukleinsäure
ddATP	2',3'-Didesoxyadenosin-5'-triphosphat
ddCTP	2',3'-Didesoxycytidin-5'-triphosphat
ddGTP	2',3'-Didesoxyguanosin-5'-triphosphat
ddTTP	2',3'-Didesoxythymidin-5'-triphosphat
dNTP	2'-Desoxyribonukleotid-5'-triphosphat
dsDNA	doppelsträngige Desoxyribonucleinsäure
EEG	Elektroenzephalografie
Effekt R	Effektrichtung
EGTCS	"Epilepsy with Generalized Tonic Clonic Seizures" - Epilepsie mit generalisierten tonisch-klonischen Anfällen

EIEE10	"Early infantile epileptic encephalopathy" - frühinfantile epileptische Enzephalopathie
EM	„Expectation-Maximization“-Algorithmus
EPICURE	"Functional Genomics and Neurobiology of Epilepsy"
F/K	Fälle/Kontrolle
FCD	"focal cortical dysplasia" - fokale kortikale Dysplasie
G	Guanin
GABA	"Gamma-aminobutyric acid" - g-Aminobuttersäure
GC-Korrektur	"genomic control"-Korrektur
GC-Score	„Gencall“-Parameter
gDNA	genomische Desoxyribonukleinsäure
GEFS+	Generalisierte Epilepsie mit Fieberanfällen plus
GM	generelles Betroffenheitsmodell
GMM	„Gaussian Mixture Model“
GRP	Analysegruppe
GSW	generalisierte „spike-wave“-Komplexe im EEG
GTC-Anfall	"generalized tonic-clonic" - generalisierter tonisch-klonischer Anfall
GW	genomweit
GWA	Genomweite Amplifizierung
GWAS	Genomweite Assoziationsstudien
HLOD	Heterogenitäts-LOD-Wert
HWD	"Hardy-Weinberg disequilibrium" - Hardy-Weinberg-Ungleichgewicht
HWE	"Hardy-Weinberg equilibrium" - Hardy-Weinberg-Gleichgewicht
IBD	"identity-by-descent" - Identität durch Abstammung
IBS	„identity by state“ - Identität nach Allelzustand
IGE	„idiopathic generalized epilepsy“ - idiopathische generalisierte Epilepsie
ILAE	"International League Against Epilepsy"
J	Jahre
JAE	"Juvenile absence epilepsy" - juvenile Absence-Epilepsie
JME	"juvenile myoclonic epilepsy" - juvenile myoklonische Epilepsie
kb	Kilobasenpaar
KI (CI)	Konfidenzintervall ("confidence interval")
LD	"Linkage Disequilibrium" - Kopplungsungleichgewicht
LOD Score	"logarithm of the odds ratio"-Wert
log ₁₀	dekadischer Logarithmus
LR-P	P-Wert der logistischen Regression
LR-P _{GC}	P-Wert der logistischen Regression nach der "Genomic Control"-Korrektur

MAF	"Minor allele frequency" - Allelfrequenzen des „Minor“-Allels
Mb	Megabasen
MDS	multidimensionale Skalierung
META-P	P-Wert der Meta-Analyse
Mg ²⁺	Magnesiumionen
MGD	"Mouse Genome Database" - Mausgenom-Datenbank
M-HET-P	Heterogenitäts-P-Werte der Replikationsanalyse bzw. der Meta-Analyse
mm	Millimeter
MP-Familien	Multiplexfamilien
mRNA	messenger-RNA
NCBI	"National Center for Biotechnology Information"
NGFN	Nationales Genomforschungsnetz
NGS	"Next Generation Sequencing" - hochparallele Sequenzierung
nm	Nanometer
NMDA	N-Methyl-D-Aspartat
OH-Enden	Hydroxylgruppe
OMIM	"Online Mendelian Inheritance in Man"
OR	"Odds Ratio" - Quotenverhältnis
p	Bezeichnung für den kurzen Arm eines Chromosoms
P(-Wert)	Signifikanzwert
PCA	"principal components analysis" - Hauptkomponentenanalyse
PCR	"polymerase chain reaction" - Polymerase-Kettenreaktion
POS	Position
PSD	"postsynaptic density" - postsynaptische Dichte
q	Bezeichnung für den langen Arm eines Chromosoms
QQ	Quantile-Quantile
RR	"relative risk" - Relatives Risiko
SM	striktes Betroffenheitsmodell
SMEI	"severe Myoclonic Epilepsy of Infancy" - schwere myoklonische Epilepsie des frühen Kindesalters (Dravet-Syndrom)
SNPs	"single nucleotide polymorphism" - Einzelnukleotidpolymorphismen
T	Thymin
TAMRA	Tetramethylrhodamin
TDT	Transmission-Disequilibrium-Test
UV	Ultraviolett
λ_{GC}	„genomic inflation“ Faktor lambda

1 EINLEITUNG

Epilepsie gehört mit einer Punktprävalenz von 0,5-1% und einer Gesamtlebenszeitprävalenz von 3% (bis 74 Jahren) zu den häufigsten chronischen Erkrankungen des zentralen Nervensystems (Hauser et al., 1993). Epilepsie führt bei den Erkrankten zu einer signifikanten Beeinträchtigung der Lebensqualität, bedingt durch neuropsychologische und psychiatrische Komorbiditäten zusammen mit einer erhöhten Sterblichkeitsrate (Cavalleri et al., 2007a; Jensen, 2011). Klinisch zeichnet sich Epilepsie durch spontan auftretende, unprovizierte Krampfanfälle aus, denen neurophysiologisch eine abnormale paroxysmale Aktivität von Neuronenverbänden zugrunde liegt (Engel, 2006; Moeller et al., 2011). Aus Studien zu der Epidemiologie und Genetik von Epilepsiesyndromen lässt sich entnehmen, dass bis zu 50% aller Epilepsien genetisch determiniert sind (Freitag et al., 2001; Sander und Gaus, 2007). Als multifaktorielle, genetisch komplexe Erkrankung, zeichnet sich Epilepsie jedoch durch eine weitreichende klinische und genetische Heterogenität aus. Bislang konnten nur für die seltenen, monogenen Formen der Epilepsie (ca. 1% der Fälle) prädisponierende Genvarianten identifiziert werden (Pandolfo, 2011). Das Ausmaß der multifaktoriellen Vererbung von Epilepsie zeigt sich in der Pleiotropie der meisten bisher identifizierten Gene (Reid et al., 2009), sowie durch die sehr große Anzahl von potentiellen Kandidatengenen (Meisler et al., 2001; Gurnett und Hedera, 2007). Genetische Studien weisen zudem auf populationspezifische Effektgrößen von prädisponierenden Genvarianten hin (Greenberg et al., 2000; Cavalleri et al., 2007b; Guo et al., 2011). In Mausmodellen von monogenen Epilepsieformen ist nachgewiesen, dass die Ausprägung von bekannten epileptogenen Mutationen vom genetischen Hintergrund abhängt (Bergren et al., 2005; Yu et al., 2006; Tan et al., 2008). Abschließend zeigen virtuelle Simulationsstudien, dass signifikante Änderungen der neuronalen Erregbarkeit durch technisch kaum messbare funktionelle Änderungen von Ionenkanälen verursacht werden können (Thomas et al., 2009).

1.1 Klassifikation epileptischer Syndrome

Ein Leitfaden für eine diagnostische Klassifikation von Epilepsien und epileptischen Anfällen, wurde von der Internationalen Liga Gegen Epilepsie definiert (International League Against Epilepsy, ILAE¹). Die mittlerweile weltweite Akzeptanz hat zu einer wesentlichen Vereinheitlichung der klinischen Diagnose von Epilepsiesyndromen geführt. Neben Vorteilen für die Behandlung von Betroffenen, wurden dadurch genetische Studien präzisiert. Vereinfacht dargestellt, werden Epilepsien nach Anfallsbild in lokalisationsbezogene Epilepsien (auch

¹ ILAE, International League Against Epilepsy, online: <http://www.ilae.org>

fokal; auf eine Großhirnhemisphäre beschränkt) und generalisierte Epilepsien (initial beide Großhirnhemisphären beteiligend) unterteilt. Basierend auf dem Manifestationsalter, dem Erkrankungsverlauf und der vermuteten Ätiologie erfolgt eine Subklassifizierung in idiopathische (genetische Ursache ohne erkennbare strukturelle Läsion), symptomatische (bekannte Ursache) und kryptogene (unbekannte Ätiologie) Formen (Gastaut, 1970; ILAE, 1981; ILAE, 1985; ILAE, 1989).

Die ILAE-Klassifizierung ist ein in Entwicklung befindender Prozess, welcher laufend weiterentwickelt und angepasst wird (Engel, 2001; Engel, 2006; Capovilla et al., 2009). Grundlegendes und bekanntes Problem des Klassifizierungsschemas sind die verwendeten Dichotomien, unter denen nicht alle Epilepsieformen eingeordnet werden können (Engel, 2011). Aus Neuroimaging-Studien wird deutlich, dass lokalisierte und generalisierte Anfälle keine strikt trennbaren Entitäten darstellen (Salek-Haddadi et al., 2009; Moeller et al., 2010). Die Trennung zwischen idiopathischen und symptomatischen Epilepsieformen, die beide genetisch bedingt sein können, ist von dem jeweiligen Kenntnisstand und technischen Untersuchungsmethoden abhängig. Ein Beispiel dafür ist die juvenile myoklonische Epilepsie, eine idiopathische Epilepsie, bei der in vereinzelt Fällen EFHC1-Mutationen und Entwicklungsstörungen des Cortex beschrieben worden sind (de Nijs et al., 2009; Léon et al., 2010; Wong, 2010).

Ein Versuch diese Inkonsistenzen zu bereinigen, stellt der neueste Klassifizierungsvorschlag der ILAE dar (Berg et al., 2010), welcher aber weitgehend mit Kritik aufgenommen worden ist (Überblick in: Panayiotopoulos, 2011). In dieser neuen Klassifizierung wurden Änderungen der bekannten Termini idiopathisch (zu genetisch), symptomatisch (zu strukturell/metabolisch) und kryptogen (zu unbekannt) eingeführt. Problematisch ist dabei, dass mit „genetisch“ ein Begriff eingeführt werden sollte, welcher nach Wortlaut nicht alle genetisch bedingten Epilepsieformen einschließt. Die vorgeschlagene Abschaffung der Klassifizierung in lokalisationsbezogene und generalisierte Epilepsiesyndrome unter Beibehaltung dieser Einteilung für Anfälle ist als „nicht evidenzbasiert“ bezeichnet worden (Ferrie, 2010; Wolf, 2011). Im Rahmen der Diskussion um den Klassifizierungsvorschlag der ILAE von 2010 wurden Gegenvorschläge unterbreitet; ein auf Ätiologie basierender Vorschlag (Shorvon, 2011) und eine Kombination der beiden Vorschläge als zweistufiges System (Wong, 2011). Mit dem Klassifizierungsvorschlag von 2010, wurde jedoch ein flexibles Schema präsentiert, das eine je nach Aufgabenstellung zweckorientierte Klassifizierung erlaubt (Berg und Scheffer, 2011).

In der Tabelle 1-1 ist eine gekürzte Zusammenstellung von epileptischen Syndromen mit den dazugehörigen internationalen Abkürzungen, die in der vorliegenden Studie genannt werden. Die Zusammenstellung kann gleichzeitig als Beispiel für eine zweckorientierte Klassifizierung dienen.

Gruppe von Syndromen	Abkürzung
• Idiopathische fokale Epilepsien von Kleinkindern und Kindern	
Benigne nichtfamiliäre infantile Anfälle	BIS
Benigne kindliche Epilepsie mit zentrotemporalen Spikes (Rolando-Epilepsie)	BCECTS
Benigne okzipitale Epilepsie des Kindesalters Typ Panayiotopoulos	PS
Spät beginnende okzipitale Epilepsie des Kindesalters (Typ Gastaut)	ICOE-G
• Familiäre (autosomal-dominante) fokale Epilepsien	
Benigne familiäre neonatale Anfälle	BFNS
Benigne familiäre infantile Anfälle	BFIS
Benigne familiäre neonatale-infantile Anfälle	BFNIS
Benigne familiäre infantile Anfälle mit familiärer hemiplegischer Migräne	BFIS-FHM
Infantile Konvulsionen mit paroxysmaler Choreoathetose	ICCA
Autosomal-dominante nächtliche Frontallappenepilepsie	ADNFLE
Familiäre Temporallappenepilepsie	FTLE
– Familiäre mesiale Temporallappenepilepsie	MFTLE
– Familiäre laterale Temporallappenepilepsie	ADLTE
Familiäre occipitotemporale Epilepsie mit Migräne und visuelle Aura	EPOLM
familiäre fokale Epilepsie mit variablen Herden	FPEVF
Partielle Epilepsie mit perizentralen Spikes	PEPS
• Idiopathische generalisierte Epilepsien	
Benigne myoklonische Epilepsie bei Kleinkindern	BMEI
Benigne adulte familiäre myoklonische Epilepsie	BAFME
Epilepsie mit myoklonisch-astatischen Anfällen (Doose-Syndrom)	MAE
Epilepsie mit myoklonischen Absencen (Tassinari-Syndrom)	EMA
Augenlid-Myoklonie mit Absencen (Jeavons-Syndrom)	ELMA
Idiopathische generalisierte Epilepsien mit variablen Phänotypen	IGE
– Kindliche Absence-Epilepsie	CAE
– Juvenile Absence-Epilepsie	JAE
– Juvenile myoklonische Epilepsie (Janz-Syndrom)	JME
– Epilepsie mit generalisierten tonisch-klonischen Anfällen	EGTCS
Generalisierte Epilepsie mit Fieberanfällen plus	GEFS+
Früh beginnende Absence-Epilepsie	EOAE
• Reflexepilepsien	
Idiopathische photosensitive Okzipitallappenepilepsie	IPOE
Primäre Leseepilepsie	PRE
• Epileptische Enzephalopathie	
Ohtahara-Syndrom (Frühinfantile epileptische Enzephalopathie)	OS (EIEE)
Frühkindliche myoklonische Enzephalopathie (EIEE3)	EME (EIEE3)
Dravet-Syndrom (Schwere frühkindliche myoklonische Epilepsie, EIEE6)	DS (SMEI)
Borderline SMEI	SMEB
Epilepsie der frühen Kindheit mit migratorischen fokalen Anfällen	MPSI
West-Syndrom (Blitz-Nick-Salaam-Krämpfe)	WS
Lennox-Gastaut-Syndrom (LGS)	LGS
Epilepsie mit kontinuierlichen Spike-Wave-Entladungen im Schlaf	ECSWS
• Epileptische Anfälle, die keine Epilepsiediagnose erfordern	
Benigne neonatale Anfälle	BNA
Fieberkrämpfe	FS

Tabelle 1-1 Beispiel einer möglichen Klassifikation von Epilepsiesyndromen

(verändert nach Engel, 2001)

1.2 Genetik von Epilepsiesyndromen

Das Ausmaß der genetischen Ätiologie von Epilepsie ist in zahlreichen Studien untersucht worden. Aggregationsstudien in Familien zeigen, dass 10% der diagnostizierten Epilepsiepatienten betroffene Verwandte haben (Callenbach et al., 1998; Bianchi et al., 2003). Zwillingsstudien belegen, dass genetische Faktoren für die familiäre Aggregation verantwortlich sind (Berkovic et al., 1998; Johnson et al., 2003; Kjeldsen et al., 2003). So sind monozygote Zwillinge signifikant häufiger konkordant für Epilepsie, Epilepsiesyndrom und Anfallstypen als dizygote Zwillinge. Patienten mit IGE weisen am häufigsten einen familiären Hintergrund (15%) und mit 56% die höchste Konkordanzrate mit ihren Verwandten bezüglich der Epilepsieform (nach Lokalisation und Ätiologie, Tabelle 1-2). Die Subsyndrome von IGE zeigen dagegen keinen Unterschied in der Häufigkeit von betroffenen Verwandten (Janz, 1997).

Lokalisation	Ätiologie des Probanden			Zwillingspaare	
	Idiopathisch ¹	Symptomatisch ¹	Kryptogen ¹	MZ ²	DZ ²
Fokal	51.0 (38.8)	33.9 (18.2)	23.2 (8.3)	0.30 - 0.36	0.05 - 0.10
Generalisiert	65.4 (56.2)	54.6 (39.4)	69.1 (16.7)	0.65 - 0.82	0.12 - 0.26
Undeterminiert	0.0 (0.0)	23.1 (0.0)	35.0 (10.0)		

Tabelle 1-2 Konkordanzraten bei Verwandten und Zwillinge von Epilepsiepatienten

¹ = Konkordanzraten von Probanden zu sämtliche betroffene Verwandte: bezüglich Lokalisation; in Klammern bezüglich Lokalisation und vermutete Ätiologie (nach Bianchi et al., 2003). ² = Konkordanzraten zwischen Zwillingspaaren (aus Berkovic et al., 1998 und Kjeldsen et al., 2003). Legende -> MZ: monozygotisches Zwillingspaar; DZ: dizygotisches Zwillingspaar.

Aus solchen klinischen Familienstudien gibt es sowohl Hinweise für genetische Determinanten, die spezifisch für ein Epilepsiesyndrom oder Subsyndrom prädisponieren (Berkovic et al., 1998; Winawer et al., 2002; Winawer et al., 2005), als auch für genetische Faktoren, die für mehrere Epilepsiesyndrome prädisponieren (Ottman et al., 1989; Scheffer und Berkovic, 1997; Ottman et al., 1998).

In Tabelle 1-3 sind die meisten derzeit bekannten Gene wiedergegeben, bei denen bestätigt ist, dass Mutationen zur Ausprägung von Epilepsie prädisponieren. Die Mehrheit dieser Gene kodieren für neuronale Ionenkanäle, was zu dem Konzept „Epilepsie als Kanalopathie“ geführt hat. Mittlerweile sind allerdings Mutationen in mehreren nicht-Ionenkanal-Genen beschrieben. Neben diesen bekannten Genen gibt es aus Kopplungsstudien Evidenzen für eine größere Anzahl von Kandidatenloci für monogene und genetisch komplexe Formen von Epilepsien. Eine Auswahl dieser Kopplungsbefunde ist in Tabelle 5-1 im Anhang wiedergegeben. Die meisten dieser Regionen sind allerdings nicht repliziert und Mutationsbefunde fehlen (Leu et al., 2012).

Gen	Region	Produkt	Syndrom	Referenz
<i>GABRD</i>	1p36.33	GABA-Rezeptor	GEFS+ JME	Dibbens et al., 2004 Dibbens et al., 2004
<i>SLC2A1</i>	1p34.2	Glucosetransporter	IGE EOAE MAE	Striano et al., 2012 Suls et al., 2009 Mullen et al., 2011
<i>CHRNA2</i>	1q21.3	Acetylcholinrezeptor	ADNFLE	Phillips et al., 2001
<i>KCNJ10</i>	1q23.2	K ⁺ -Kanal	IGE	Lenzen et al., 2005
<i>ATP1A2</i>	1q23.2	Na ⁺ /K ⁺ -ATPase	BFIS mit FHM	Vanmolkot et al., 2003
<i>CACNB4</i>	2q23.3	Ca ²⁺ -Kanal	IGE JME	Escayg et al., 2000a Escayg et al., 2000a
<i>KCNJ3</i>	2q24.1	K ⁺ -Kanal	IGE	Chioza et al., 2002
<i>SCN2A</i>	2q24.3	Na ⁺ -Kanal	BFIS BFNIS GEFS+ EIEE11	Striano et al., 2006 Berkovic et al., 2004a Sugawara et al., 2001 Ogiwara et al., 2009
<i>SCN1A</i>	2q24.3	Na ⁺ -Kanal	FS alleine GEFS+ MPSI SMEI SMEB WS	Mantegazza et al., 2005 Escayg et al., 2000b Carranza Rojo et al., 2011 Striano et al., 2007 Harkin et al., 2007 Wallace et al., 2003
<i>SCN9A</i>	2q24.3	Na ⁺ -Kanal	FS GEFS+ SMEI	Singh et al., 2009 Singh et al., 2009 Singh et al., 2009
<i>GPR98</i>	5q14.3	G-Protein-gekoppelter Rezeptor	FS/GEFS+	Nakayama et al., 2002
<i>GABRA1</i>	5q34	GABA-Rezeptor	CAE JME	Maljevic et al., 2006 Cossette et al., 2002
<i>GABRG2</i>	5q34	GABA-Rezeptor	FS alleine CAE mit FS GEFS+ SMEI	Audenaert et al., 2006 Wallace et al., 2001 Baulac et al., 2001 Jansen et al., 2006
<i>EFHC1</i>	6p12.2	Mikrotubuli assoziiertes Protein	JAE JME	Stogmann et al., 2006 Suzuki et al., 2004
<i>CHRNA2</i>	8p21.2	Acetylcholinrezeptor	ADNFLE	Aridon et al., 2006
<i>CPA6</i>	8q13.2	Carboxypeptidase	FS mit TLE FTLE	Salzmann et al., 2012 Salzmann et al., 2012
<i>KCNQ3</i>	8q24.22	K ⁺ -Kanal	BFNS	Charlier et al., 1998
<i>JRK</i>	8q24.3	Centromer bindendes Protein	IGE	Moore et al., 2001
<i>LGI1</i>	10q23.33	Nogo receptor 1 Ligand	ADLTE	Berkovic et al., 2004b
<i>SCN8A</i>	12q13.13	Na ⁺ -Kanal	SMEI	Martin et al., 2007
<i>GABRB3</i>	15q12	GABA-Rezeptor	CAE	Tanaka et al., 2008
<i>CACNA1H</i>	16p13.3	Ca ²⁺ -Kanal	CAE IGE	Chen et al., 2003 Heron et al., 2007
<i>CACNA1G</i>	17q21.33	Ca ²⁺ -Kanal	IGE	Yalçın, 2012
<i>ME2</i>	18q21.2	Mitochondriales Enzym	IGE	Greenberg et al., 2005
<i>NEDD4L</i>	18q21.31	Ubiquitin-Protein Ligase	IGE mit PPR	Dibbens et al., 2007a
<i>SCN1B</i>	19q13.12	Na ⁺ -Kanal	GEFS+	Wallace et al. 1998
<i>CHRNA4</i>	20q13.33	Acetylcholinrezeptor	ADNFLE	Steinlein et al., 1995
<i>KCNQ2</i>	20q13.33	K ⁺ -Kanal	BFNS	Singh et al., 1998

Tabelle 1-3 Bekannte Gene die zur Ausprägung eines Epilepsiesyndroms prädisponieren

Die Liste ist sortiert nach der genauen chromosomalen Position. Als Legende für die jeweiligen Syndrome dient Tabelle 1-1.

1.3 Epilepsie als genetisch komplexe Erkrankung

Das hohe Ausmaß an genetischer Heterogenität bei den häufigen Epilepsiesyndromen zeichnet sich durch eine rasch abnehmende Vererbungswahrscheinlichkeit jenseits von Verwandten ersten Grades aus, während monozygotische Zwillinge hohe Konkordanzraten aufweisen (Mulley et al., 2005). In einem multifaktoriell-polygenen Modell für Epilepsie ist eine hohe Anzahl von Genen vorstellbar, die jeweils einen additiven oder interaktiven Effekt mit variabler Stärke auf das Erkrankungsrisiko haben. Während das Erkrankungsrisiko in der Allgemeinbevölkerung normalverteilt ist, werden Personen die aufgrund der Anzahl von Risikofaktoren eine gewisse Anfallsschwelle überschreiten, eine Epilepsie entwickeln. Für die Verwandten dieser Personen wird die Verteilung des Erkrankungsrisikos entsprechend in Richtung der Schwelle für Epilepsie verschoben sein (Abb. 1-1).

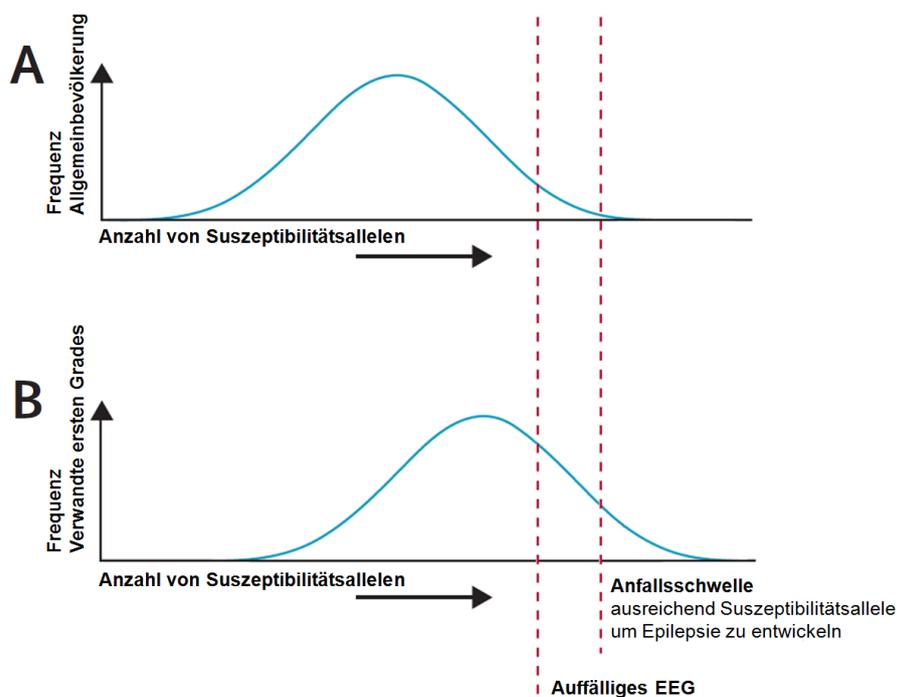


Abb. 1-1 Individuelle genetische Suszeptibilität für Epilepsie

Dargestellt ist die Anzahl von Suszeptibilitätsallelen als Funktion der Frequenz in der Allgemeinbevölkerung (A) und in Verwandte von betroffenen Patienten (B) (verändert nach Helbig et al., 2008)

Bei den monogenen Formen von Epilepsie wird bereits in einem kleineren Rahmen die Komplexität von Epilepsie erkennbar. Jeweils Mutationen in unterschiedlichen Genen können zu der Entstehung von ADNFLE und monogenen Formen von GEFS+ oder IGE führen (siehe Tabelle 1-3). Gleichzeitig zeigen mehrere der bekannten Suszeptibilitätsgene eine ausgeprägte Pleiotropie (unterschiedliche phänotypische Manifestationen eines Gendefekts): *SLC2A1* (Mullen et al., 2010), *GABR*-Gene (Macdonald et al., 2010; Lachance-Touchette et

al., 2011) und *SCN1A* (Harkin et al., 2007). Zusätzlich weisen Studien in Tiermodellen auf Gen-Gen-Interaktionen hin. Zwei einzelne, epileptogene Mutationen in den Genen *Kcna1* und *Cacna1a* ergeben in einer Maus-Doppelmutante einen leichteren Epilepsiephänotyp (Ausbleiben von „spike-wave“-Komplexen im EEG, Verringerung der Anfallshäufigkeit und Stärke; Glasscock et al., 2007). Ebenso zeigen *Scn1a / Scn8a* Doppelmutanten einen Phänotyp mit annähernd normalen Schwellenwert für induzierte Anfälle (Martin et al., 2007). Aus Studien mit Mausmodellen für Epilepsie ist bekannt, dass der genetische Hintergrund die phänotypische Ausprägung von bekannten epileptogenen Mutationen beeinflusst (Jorge et al., 2011).

An dieser Stelle wird die Notwendigkeit für eine zweckorientierte Klassifizierung von Epilepsiesyndromen deutlich, wie sie von Berg et al., 2010 vorgeschlagen wurde. Die genetische Heterogenität ist dermaßen ausgeprägt, dass klinische Phänotypisierung nicht verlässlich zu einer größeren genetischen Homogenität führt. Da gleichzeitig eine bedeutende Anzahl der bekannten Anfälligkeitsgene pleiotrope Effekte haben ist es vorstellbar, dass manche Epilepsieformen trotz unterschiedlicher klinischer Erscheinungsbilder bereits genetisch homogen sind (Mefford und Mulley, 2010). Als Beispiel kann *SCN1A* dienen, das bisher bedeutendste Epilepsiegen. *SCN1A* konnte als Epilepsiegen identifiziert werden, als erkannt wurde, dass mehrere heterogene Epilepsieformen eine gemeinsame genetische Grundlage haben und diese unter GESF+ als neues Syndrom vereinigt wurden (Scheffer und Berkovic, 1997; Johnson, 2011). Das Leitsymptom von GESF+ sind Fieberkrämpfe in der Kindheit, die über das sechste Lebensjahr hinaus bestehen bleiben können (FS+). Ein Teil der Patienten entwickeln eine Fieber-unabhängige generalisierte tonisch-klonische Epilepsie, die mit Absencen, myoklonischen, atonischen oder fokalen Anfällen assoziiert sein kann. In seltenen Fällen werden auch andere Epilepsieformen diagnostiziert (MAE und SMEI). GESF+ wird aufgrund dieser phänotypischen Variabilität in der Regel unter Einbeziehung aller betroffenen Familienmitglieder diagnostiziert.

1.4 Idiopathische generalisierte Epilepsie

Die Gruppe der IGEs umfasst mehrere Subsyndrome, die anhand des Anfallstyps, des Manifestationsalters und des Erkrankungsverlaufs unterschieden werden (ILAE, 1985; ILAE, 1989; Tabelle 1-1). Zusammen repräsentieren die IGEs 20-30% aller Epilepsien (Berg et al., 1999; Jallon et al., 2001). Die kindliche Absence-Epilepsie (CAE), die juvenile Absence-Epilepsie (JAE), die juvenile myoklonische Epilepsie (JME) und die Epilepsie mit generalisierten tonisch-klonischen Anfällen (EGTCS) machen über 90% aller IGE-Fälle aus (Mullins et al., 2007). Diese vier Subsyndrome weisen deutliche Gemeinsamkeiten in der Anfallsentstehung und EEG-Muster auf (Moeller et al., 2011; Zambrelli und Canevini, 2011).

Auch bezüglich des Anfallstyps gibt es Überlappungen: CAE, die in 50% der Erkrankungsverläufe eine vollständige Remission in der Adoleszenz aufweist (Wirrell et al., 1996), entwickelt sich in 18% der Fälle zu einer JME im Jugendalter (Martínez-Juárez et al., 2006); juvenile Absencen finden sich bei ca. 30% der JME-Patienten (Beck-Mannagetta et al., 1991); generalisierte tonisch-klonische Anfälle (GTCS) sind in nur 40-50% isoliert anzutreffen, während 72% aller IGE-Patienten mindestens einen GTC-Anfall erleiden (Crunelli und Leresche, 2002; Mullins et al., 2007). Die Überlappung der klinischen Phänotypen ist in Abb. 1-2 schematisch dargestellt.

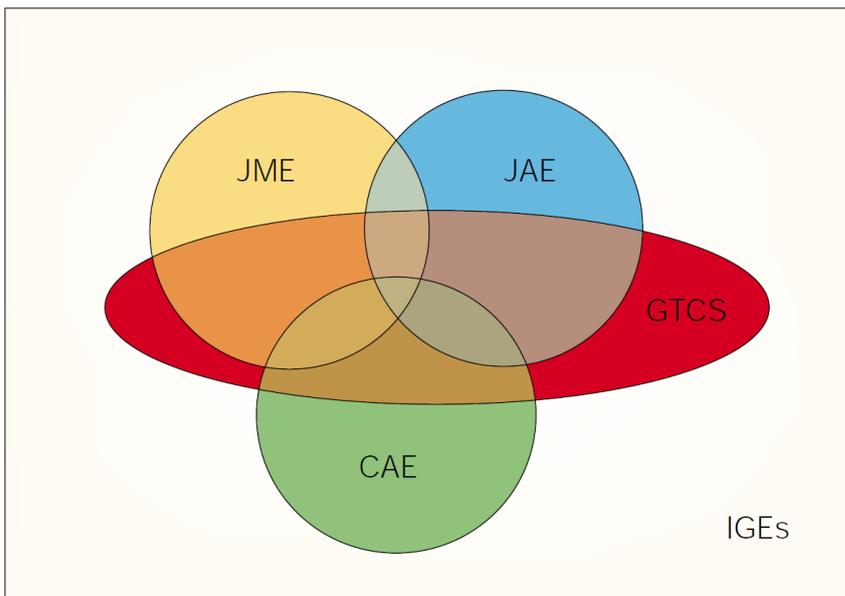


Abb. 1-2 Überlappung der komplexen Phänotypen der IGEs
(nach Crunelli und Leresche, 2002)

Die IGE-Subsyndrome lassen sich, abgesehen von den Anfallstypen, am besten durch den typischen Zeitpunkt des Anfallsbeginns unterscheiden (CAE 6-8J., Ø7,5J.; JAE 10-16J., Ø13,3J.; JME 12-18J., Ø14,6J.; EGTCS 6-28J., Peak 17,1J.; Janz, 1997). Genetische Studien bestätigen dieses Bild einer partiellen Überlappung der IGE-Subgruppen und weisen auf gemeinsame, als auch subsyndromspezifische genetische Faktoren hin (Sander et al., 2000; Durner et al., 2001; Hempelmann et al., 2006).

Die beschriebene genetische Komplexität von Epilepsien, gilt auch für IGE und selbst für die relativ charakteristischen Subsyndrome. Das Ausmaß an genetischer Homogenität, die durch die Subklassifizierung erreicht wird, lässt sich nur schwer einschätzen und ist wahrscheinlich eher geringfügig. So gibt es bereits aus klinischen Studien Evidenzen für Unterschiede zwischen Patientengruppen desselben Subsyndroms. JME-Patienten, die in ihrer Kindheit eine CAE aufgewiesen haben, weisen bei ihren betroffenen Verwandten signifikant höhere Frequenzen von EGTCS, FS oder IGE auf verglichen zu den Familien von klassi-

schen JME-Fällen (Martínez-Juárez et al., 2006). Zusätzlich sprechen Patienten deren CAE sich zu JME entwickelt hat schlechter auf Therapie mit antiepileptischen Pharmaka (AEDs) an und erreichen in nur 7% der Fälle Anfallsfreiheit (Wirrell et al., 1996). Bei CAE-Patienten weisen Aggregationsstudien auf Unterschiede zwischen Fällen von CAE mit und ohne GTCS hin (Kinirons et al., 2008a). Patienten mit einer Absence-Epilepsie und GTCS haben eine schlechtere Prognose und sprechen schlechter auf AEDs an als vergleichbare Patienten ohne GTCS (Crunelli und Leresche, 2002; Curatolo et al., 2009).

1.5 Zielsetzung

Ziel dieser Arbeit war es, populationsübergreifende Erkrankungsloci für häufige IGE Syndrome zu identifizieren. Die positionelle Kartierung von IGE-Genen erfolgte mittels genomweiter Kopplungs- und Assoziationsstudien. Als Forschungsgrundlage wurde IGE ausgewählt, da für dieses Epilepsiesyndrom eine vorwiegend genetische Ätiologie vorliegt. Darüber hinaus sollte überprüft werden, ob idiopathische Absence-Epilepsie (IAE, umfasst CAE und JAE) und JME subsyndromspezifische Determinanten aufweisen, die differentiell zum Entstehungsrisiko von entweder IAE oder JME beitragen.

1.6 Forschungsgrundlage

Innerhalb von mehreren Förderperioden wurde das weltweit größte Kollektiv von 379 IGE-Multiplexfamilien (europäischen Ursprungs) und 1527 IGE-Einzelfällen (Ursprung Deutschland und Anreinerstaaten) gesammelt. Beide Kollektive wurden im Rahmen des Projektes genomweit mit genetischen Sequenzpolymorphismen genotypisiert. Für die Multiplexfamilien wurde der speziell für Kopplungsstudien konzipierte *Illumina HumanLinkage*-SNP-Genotypisierungschip verwendet. Die IGE-Einzelfälle wurden mit dem *Affymetrix Genome Wide Human SNP Array 6.0* genotypisiert. Dieser Chip garantiert eine für genomweite Assoziationsstudien (GWAS) geeignete Abdeckung des gesamten Genoms mit einem durchschnittlichen Markerabstand von ca. 3000 bp. Über NGFN-/NGFNplus-Netzwerke² wurden von der PopGen Biobank (Krawczak et al., 2006) und aus der KORA-Studie (Wichmann et al., 2005) 2461 deutsche Populationskontrollen für die GWAS zur Verfügung gestellt. Für die Replikationsanalyse der GWAS wurden im Rahmen des EPICURE-Konsortiums³ zwei weitere unabhängige IGE-Kollektive gesammelt (breitgefächerter, europäischer Ursprung): 1.) 572 Eltern-Kind-Trios mit IGE erkrankten Kindern und 2.) ein Fall-Kontroll-Kollektiv mit 727 IGE-

² NGFN, Nationales Genomforschungsnetz, online: <http://www.ngfn.de>

³ EPICURE, Functional Genomics and Neurobiology of Epilepsy, online: <http://www.epicureproject.eu>

Betroffenen und 668 Populationskontrollen. Für die Kopplungsanalysen wurden 204 Familien aus zwei früheren Studien (Sander et al., 2000; Hempelmann et al., 2006) und 175 Familien aus dem EPICURE Projekt in eine Meta-Kopplungsanalyse eingeschlossen. Insgesamt standen 2254 IGE-Patienten, 3129 Populationskontrollen, 572 IGE-Trios und 379 IGE-Multiplexfamilien für molekulargenetische Untersuchungen zur Verfügung.

2 METHODEN

2.1 Studienteilnehmer

Für die Studie wurden nur Epilepsie-Patienten berücksichtigt, die eines der häufigen IGE-Subsyndrome aufwiesen (CAE, JAE, JME und EGTCS mit generalisierten „spike-wave“-Komplexen im EEG). Die Probanden wurden gemäß eines standardisierten Studienprotokolls⁴ rekrutiert. Patienten mit bekannten neurologischen Komorbiditäten (Autismus-Spektrum-Störung, Schizophrenie, Psychosen, geistige Behinderung) wurden aus der GWAS ausgeschlossen. In Tabelle 2-1 sind die Besetzungen der drei unabhängigen Kollektive wiedergegeben, die für die GWAS gesammelt worden sind.

GWAS-Kollektive	Patienten	Kontrollen	Trios	CAE	JAE	IAE	JME	EGTCS
GW-Scan	1527	2461	-	480	215	7	586	239
Replikation 1	727	668	-	184	98	30	323	92
Replikation 2	-	-	572	251	66	6	158	91

Tabelle 2-1 IGE-Kollektive für Assoziationsanalysen

Die Phänotypgruppen CAE, JAE und IAE enthalten keine Fälle von JME. Legende -> CAE: Kindliche Absence-Epilepsie; JAE: Juvenile Absence-Epilepsie; IAE: Idiopathische Absence-Epilepsie; JME: Juvenile myoklonische Epilepsie; EGTCS: Epilepsie mit generalisierten tonisch-klonischen Anfällen.

Genomscan	FAM Anz.	IND Anz.	TYP Anz.	IGE Patienten	nicht betroffen	GSW im EEG	FS alleine	andere Epilepsien	nicht bekannt
GS2000	107	551	496	273	251	11	2	9	5
GS2006	97	522	461	261	239	6	7	6	3
EPICURE	175	847	771	448	366	4	9	9	11
Kombiniert	379	1920	1728	982	856	21	18	24	19

Tabelle 2-2 IGE-Multiplexfamilien für Meta-Kopplungsanalyse

Legende -> GS2000 / GS2006: Kopplungsstudien Sander et al., 2000 und Hempelmann et al., 2006; FAM Anz. / IND Anz. / TYP Anz.: Gesamtzahlen von Familien, Individuen und typisierten Personen; GSW: „spike-wave“-Komplexe; FS: Fieberkrämpfe.

Multiplexfamilien wurden ebenfalls über IGE-Patienten ohne neurologische Komorbiditäten ausgewählt. Die für die Familienstruktur essentielle Patienten ohne IGE-Phänotyp oder mit neurologischen Komorbiditäten, wurden mit unbekanntem Phänotyp kodiert. Ein weiteres Einschlusskriterium für Multiplexfamilien war das Vorhandensein von mindestens einem weiteren betroffenen Geschwister mit einem IGE-Subsyndrom (phänotypische Konkordanz war dabei nicht erforderlich). In Tabelle 2-2 sind die Familien wiedergegeben, die in dieser Studie

⁴ Studienprotokoll, online: <http://portal.ccg.uni-koeln.de/ccg/research/epilepsy-genetics/sampling-procedure>

gesammelt worden sind. Zusammen mit den Familien aus zwei früheren Studien gingen diese in eine Meta-Analyse ein.

Tabelle 2-3 listet die Bevölkerungsgruppen auf, aus denen die drei GWAS-Kollektive und die Multiplexfamilien rekrutiert worden sind.

Bevölkerungsgruppe	GWAS		Replikation 1		Replikation 2	Kopplung
	IGE-Fälle	Kontrollen	IGE-Fälle	Kontrollen	IGE-TRIOS	MP-Familien
Österreich	231	-	3	-	1	3
Belgien	41	-	-	-	-	1
Dänemark	93	-	6	-	5	10
Deutschland	885	2461	45	86	5	92
Niederlande	277	-	2	-	5	23
Australien	-	-	244	184	75	33
Neuseeland	-	-	3	-	4	-
Vereinigtes Königreich	-	-	20	-	49	24
Mauritius	-	-	-	-	1	-
Finnland	-	-	7	7	9	7
Aschkenasim	-	-	4	-	6	2
Israel	-	-	-	-	-	4
Schweden	-	-	-	-	1	1
Spanien	-	-	53	59	5	6
Frankreich	-	-	109	89	18	45
ehem. Jugoslawien	-	-	5	-	3	1
Bulgarien	-	-	2	-	4	4
Sinti	-	-	-	-	1	-
Griechenland	-	-	5	-	2	8
Italien	-	-	63	88	260	93
Türkei	-	-	156	155	118	21
Polen	-	-	-	-	-	1
Gesamtzahlen	1527	2461	727	668	572	379

Tabelle 2-3 Ethnizitäten der gesammelten IGE-Kollektive

2.2 DNA-Aufbereitung und Qualitätskontrolle

Alle verwendeten Genotypisierungsmethoden basieren auf Intensitätsmessungen und sind damit abhängig von genauen Gesamtmenge und Qualität der eingesetzten DNA.

2.2.1 DNA-Konzentrationsbestimmung

Die DNA-Konzentration wurde mithilfe der Spektrophotometer Nanodrop ND-1000 und Nanodrop 8000 von NanoDrop⁵ bestimmt. Diese Spektrophotometer ermöglichen die Be-

⁵ Nanodrop, online: <http://www.nanodrop.com>

stimmung von Absorption von DNA-, RNA und Proteinlösungen sowie fluoreszenzmarkierten Microarrays. Für die Messung werden 1-2µl Probe zwischen zwei stumpf endenden und 1 mm voneinander entfernten Lichtwellenleitern platziert. Durch die natürliche Oberflächenspannung wird die Probe in der Messoptik festgehalten. Der erfasste Lichtwellenlängenbereich reicht von 220 bis 750 nm und lässt, neben der Konzentrationsbestimmung, auch Aussagen über die Reinheit der untersuchten Lösung zu. Das Absorptionsmaximum für Nucleinsäuren liegt bei einer Wellenlänge von 260nm, für Proteine liegt es bei 280nm und für organische Verbindungen bei 230nm. Über die Verhältnisse der drei Absorptionsmaxima wird die Reinheit der untersuchten DNA-Lösung quantifiziert. Eine reine DNA-Lösung weist ein A_{260} / A_{280} -Ratio von 1,8-2,0 und ein A_{260} / A_{230} -Ratio von 2,0-2,4 auf. Abweichende Werte weisen auf Kontaminationen mit Proteinen oder organische Verbindungen hin. Durch mehrfache Messung derselben DNA-Probe kann man feststellen, ob die DNA in dem Puffer homogen gelöst ist, was zu einer Verringerung der Fehlerquote bei den nachfolgenden Verdünnungen führt.

2.2.2 Agarose-Gelelektrophorese

Die Elektrophorese im Agarosegel wird zur Visualisierung von genomischer DNA (gDNA) und PCR-Produkten benutzt. Agarose ist ein langkettiges Polysaccharid aus glycosidisch verbundener D-Galactose und 3,6-Anhydrogalactose. Seine Konzentration im Gel bestimmt die Porengröße. Im elektrischen Feld wandern Nucleinsäuren aufgrund der negativ geladenen Phosphatgruppe durch die Gelmatrix stets zur positiv geladenen Anode. Dabei ist die Wanderungsgeschwindigkeit linearer DNA-Moleküle umgekehrt proportional zu dem dekadischen Logarithmus des Molekulargewichtes bzw. der Anzahl an Basenpaaren. Für die Detektion der DNA wird dem Gel Ethidiumbromid zugegeben. Ethidiumbromidmoleküle interkalieren zwischen die Basen der DNA und bilden mit ihnen Charge-Transfer-Komplexe, welche bei Anregung durch UV-Licht fluoreszieren. Zur visuellen Kontrolle der DNA wird eine Geltasche mit einem Größenstandard belegt und alle weiteren Geltaschen mit äquivalenten Mengen der Test-DNA beladen. Die Methode eignet sich, außer für die Auftrennung von PCR-Fragmenten, auch zur visuellen Kontrolle der DNA-Qualität und zur Quantifizierung von Degradation von gDNA, erkennbar als Schmier oder Leiterbanden. Fotografiert wurden die Gele anschließend mit dem UV-Geldokumentationsapparat Alphamager von Biozym⁶.

⁶ Biozym, online: <http://www.biozym.com>

2.3 Polymerase-Kettenreaktion

Mit Hilfe der Polymerase-Kettenreaktion („Polymerase Chain Reaction“, PCR) ist es möglich, DNA-Fragmente, deren Sequenzen zumindest teilweise bekannt sind, in exponentiellem Maße zu amplifizieren. Dabei verwendet man zwei 18-27 bp lange Oligonukleotid-Startermoleküle (Primer), welche die Zielsequenz flankieren und in entgegengesetzter Orientierung zueinander binden. Durch die hitzestabile DNA-Polymerase des thermophilen Bakteriums *Thermus aquaticus* (*Taq*-Polymerase) werden die 3'-OH-Enden der Primer durch Anfügen der in dem Reaktionsansatz bereitgestellten Desoxyribonukleosid-Triphosphate (dNTPs) verlängert, so dass zwei komplementäre DNA-Einzelstränge synthetisiert werden. Nach erneuter Denaturierung und Primeranlagerung dienen neusynthetisierte Einzelstränge ebenfalls als Vorlage. Das PCR-Fragment wird auf diese Weise exponentiell durch einen zyklischen Temperaturverlauf in drei Schritten vermehrt: 1.) Denaturierung bei 95 °C - bei dieser Temperatur werden auch lange dsDNA-Stränge in Einzelstränge getrennt; 2.) Primerhybridisierung - nach der Trennung der DNA-Stränge wird die Temperatur so weit gesenkt, dass die Primer an die jeweilige komplementäre Sequenz binden können. Optimal für die Primerhybridisierung sind Temperaturen von 2-3 °C unterhalb der Schmelztemperatur der Primer-Zielsequenz-Hybride; 3.) Elongation bei 72 °C - die *Taq*-Polymerase hat ihr Reaktionsmaximum bei einer Temperatur von 72 °C. Beginnend von den 3'-OH-Enden der Primer wird ein neuer komplementärer Strang synthetisiert. Diese drei Schritte werden 30-35 mal wiederholt. Die Anzahl der Zyklen ist nicht willkürlich, sondern mäßigt das Auftreten unspezifischer Produkte. Außerdem ist die Reaktion abhängig von der Konzentration der dNTPs und der Effizienz und Lebensdauer der *Taq*-Polymerase; die Amplifikationsrate würde daher ab einem Punkt stagnieren. Abgesehen von der DNA, Primern, dNTPs und Polymerase ist für die Reaktion ein geeignetes Puffersystem nötig. Für die Aktivität der Polymerase sind zusätzlich Mg^{2+} -Ionen notwendig.

2.4 SNP-Genotypisierung

2.4.1 Affymetrix SNP 6.0

Für die genomweite Genotypisierung von 1527 IGE-Einzelfällen wurde der *Affymetrix*⁷ *Genome-Wide Human SNP Array 6.0* verwendet. Dieser Chip ermöglicht pro Person und Chip, 906.600 Einzelnukleotidpolymorphismen (SNPs) zu typisieren.

Der Affymetrix SNP 6.0-Assay basiert auf einer sequenzspezifischen Erfassung von SNPs mittels allelspezifischer Hybridisierung (Kennedy et al., 2003; Matsuzaki et al., 2004). Abb. 2-

⁷ Affymetrix, online: <http://www.affymetrix.com>

1 zeigt das dazugehörige Flussdiagramm der nötigen Experimente. Die eingegebene DNA wird mit zwei Restriktionsenzymen (Nsp I / Sty I) geschnitten und mit Adapter-Molekülen versehen. Mittels PCR und einem generischen Primer, wird die DNA als 200-1000bp große Fragmente amplifiziert. Vor der Hybridisierungsreaktion werden die PCR Fragmente gereinigt, erneut fragmentiert, denaturiert und fluoreszenzmarkiert. Hybridisiert wird an 25bp großen Oligomeren, welche als Cluster von identischen Oligomeren an einer genau festgelegten Position auf der Chipoberfläche vorliegen. Ein solcher Cluster ist spezifisch für ein Allel eines SNPs und wird von Affymetrix „probe“ genannt. Für die zwei Allele eines SNPs werden zwei solcher Cluster gebraucht („probeset“). Pro SNP befinden sich drei solcher „probesets“ physikalisch getrennt auf einem Chip (Chai et al., 2010).

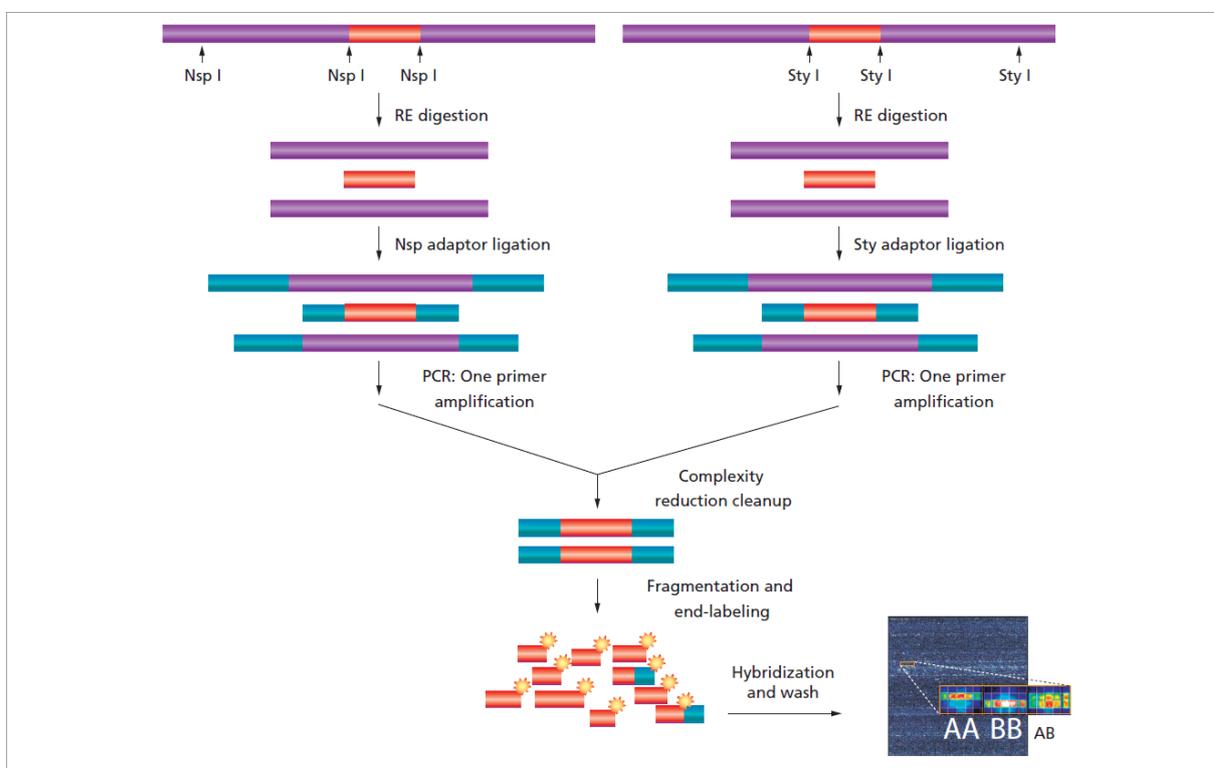


Abb. 2-1 Flussdiagramm des *Genome-Wide Human SNP Nsp/Sty-Assays* (nach Datenblatt 702509 Rev. 2, Affymetrix 2009)

Für die Genotypbestimmung („Genotype-Calling“) wurde die windowsbasierte Affymetrix Genotyping Console 4.1 Software mit dem Birdseed v2⁸ Algorithmus verwendet. Der „Calling“-Algorithmus benutzt ein „Gaussian Mixture Model“ (GMM), welches mit einem „Expectation Maximization“-Algorithmus (EM-Algorithmus) trainiert wird, um zusammenhängende Datenpunkte in einem zweidimensionalen Raum, dargestellt durch die zwei gemessenen Intensitätswerte, pro SNP zu bestimmen. Das erste EM-Modell wird dabei vorgegeben.

⁸ Birdseed, online: <http://www.broadinstitute.org/mpg/birdsuite/birdseed.html>

Für die SNPs der X und Y Chromosomen werden Geschlechtsinformationen bzw. spezielle EM-Modelle vorgegeben. Am besten funktioniert diese Vorgehensweise, wenn alle drei Genotypen eines SNPs durch eine ausreichende Anzahl an Personen vertreten werden. Die Rate mit der Genotypisierungsfehler auftreten, sowie die Genotypisierungsrate (CR), verhalten sich umgekehrt proportional zu der Anzahl der verwendeten Personen (de Andrade et al., 2011). Proben mit einer schlechten Assay-Performanz oder Kontaminationen können zu einer höheren Fehlerrate im „Calling“ führen. Solche Proben können über den Parameter „Contrast Quality Control“ (cQC, Standard min. 0,4) identifiziert werden, ein Messwert für die Diskrimination zwischen Genotyp-Clustern (Affymetrix, 2008). Der cQC-Wert wird berechnet über die Verteilung der Allelintensitäten von 9.340 willkürlich ausgewählten autosomalen SNPs. Zusätzlich werden cQC-Werte für zwei Sets von jeweils 20.000, für eine der Restriktionsenzyme, spezifische SNPs (cQC-Nsp / cQC-Sty) berechnet. Da unter optimalen Assaybedingungen erwartet wird, dass beide enzymatische Reaktionen vergleichbare Ergebnisse erzeugen, wird eine große Differenz zwischen den enzymespezifischen cQC-Werte als Ausfall gewertet (Standard cQC = 0 wenn $cQC-Nsp - cQC-Sty > 2$). Diese standardmäßige Vorgehensweise ist jedoch unzureichend, weil Proben mit einer schlechten Performance in einem der beiden enzymatischen Schritte ($cQC-Nsp / cQC-Sty < 0,4$) diese QC-Bewertung überstehen und das Genotyp-Clustering negativ beeinflussen können (Pluzhnikov et al., 2010). Für die verbleibenden Proben wird über den Parameter „Confidence Threshold“ (Standard 0,1) die maximal zulässige Entfernung vom Clustermittelpunkt vorgegeben. Nach einem ersten „Genotype-Calling“ Durchlauf, können Proben mit schlechten Genotypisierungsraten ($CR < 95\%$) und auffällige Heterozygotieraten ausgeschlossen werden. Das „Genotype Calling“ wird abschließend ein zweites Mal ausschließlich mit Proben, die die Qualitäts Checks bestanden haben, durchgeführt.

In dieser Studie wurden folgende Parameter für das „Genotype Calling“ verwendet: $cQC \geq 0,4$; $cQC-Nsp \geq 0,4$; $cQC-Sty \geq 0,4$; „Confidence Threshold“ = 0,1; $CR \geq 0,95$. Zusätzlich wurde das „Genotype-Calling“ in mehreren Batches ausgeführt, unterteilt nach dem Zeitpunkt der Genotypisierung. Mit dieser Vorgehensweise werden im Durchschnitt höhere CRs erreicht (de Andrade et al., 2011).

2.4.2 GenomeLab SNPstream

Für die Replikationsrunde der GWAS wurde die von Beckman Coulter⁹ entwickelte *GenomeLab SNPstream*-Genotypisierungsmethode verwendet. Mit dieser Multiplexmethode können auf einem Array 48 SNPs im 384er-Plattenformat gleichzeitig typisiert werden.

⁹ Autoprimer, online: <http://www.autoprimer.com>

Der SNPstream-Assay basiert auf Einzelbasenpaar-Primerextension mit fluoreszenzmarkierten Nukleotiden (Nikiforov et al., 1994; Bell et al., 2002). In Abb. 2-2 ist der SNPstream als Flussdiagramm vereinfacht dargestellt. Im ersten Schritt des Assays werden 48 DNA Fragmente, auf denen sich die gewünschten SNPs befinden, in einer Multiplex-PCR amplifiziert. Die Primer für die PCR-Multiplexreaktion und die Einzelbasenpaar-Extension werden über das Webtool Autoprimer¹⁰ erstellt. Die PCR-Fragmente werden gereinigt und der Primerextension-Reaktion unterzogen, bei der die Sequenzprimer um ein Basenpaar verlängert werden, das dem dazugehörigem SNP entspricht. Mit TAMRA markierten Cytosin (C) und Guanin (G) Didesoxyribonukleosid-Triphosphate (ddCTP / ddGTP), zusammen mit Bodipy-Fluorescein markierten Adenin (A) und Thymin (T) Didesoxyribonukleosid-Triphosphate (ddATP / ddTTP), können alle SNPs erfasst werden, außer CG- und AT-SNPs. Für CG- und AT-SNPs werden TAMRA markierte ddGTPs und ddATPs zusammen mit Bodipy-Fluorescein markierten ddCTPs und ddTTPs benutzt. Wird der Schritt der Einzelbasenpaar-Extension getrennt nach diesen zwei SNP-Kategorien durchgeführt, können alle SNP-Varianten in einem Assay typisiert werden (Universal SNPstream-Assay, Abb. 2-2). Jeder Sequenzprimer trägt an seinem 5'-Ende eine einzigartige, 20bp große Erkennungssequenz, die keine Homologie zu gDNA aufweist. Diese 48 Erkennungssequenzen sind zu jeweils einer festgelegten Oligonukleotidsonde auf dem Array komplementär. Nach der Einzelbasenpaar-Primerextension werden die nun fluoreszenzmarkierten Sequenzprimer mittels Hybridisierung an genau definierten Positionen auf dem Array rekrutiert. Abschließend werden die Arrays in einem Laserscanner abgelesen.

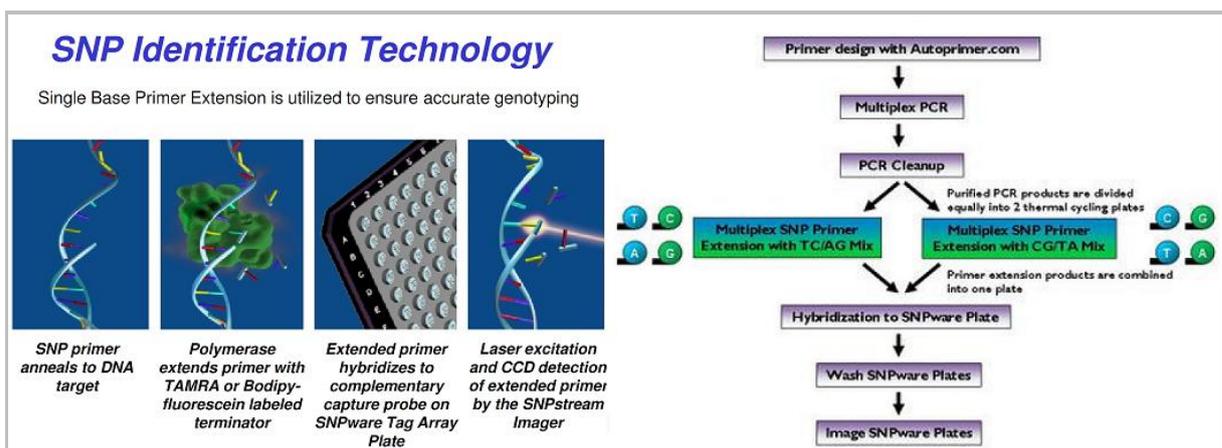


Abb. 2-2 Flussdiagramm des SNPstream-Assays
(verändert nach Beckman Coulter)

¹⁰ Beckman Coulter, online: <http://www.beckmancoulter.com>

Für die Genotypbestimmung verwendet die Beckman Coulter Software keinen speziellen „Calling“-Algorithmus. Die SNP-Genotypen werden anhand der Intensitätswerte der zwei Fluoreszenzgruppen direkt ermittelt. Dabei bekommen Proben, die eine festgesetzte Intensitätsschwelle oder die Intensitätswerte der Nullkontrollen des Arrays unterschreiten, keinen Genotyp zugewiesen. Diese „Calling“-Vorgehensweise erfordert eine manuelle Nachbearbeitung der meisten genotypisierten SNPs.

2.4.3 Illumina Infinium

Für die Genotypisierung von 175 Multiplexfamilien wurden *Illumina*¹¹ *HumanLinkage*-Genotypisierungschips mit 6090 SNPs verwendet. Auf diesen Chips befinden sich, speziell für Kopplungsstudien ausgewählte, Tag-SNPs in einem durchschnittlichen Abstand von 500 kb, die eine genomweite Abdeckung garantieren, ohne zueinander im Kopplungsungleichgewicht (linkage disequilibrium, LD) zu stehen. Im Zuge von Produktaktualisierungen seitens Illumina, wurden in dieser Studie Versionen dieses Chips verwendet auf denen 12 beziehungsweise 24 Personen gleichzeitig genotypisiert werden können. Für die Genotypisierungsarbeiten wurden die „Illumina Infinium II“ beziehungsweise „Infinium HD Ultra“ Assays verwendet.

Der Infinium Assay basiert auf Hybridisierung und direkte, enzymatische Erfassung von SNPs (Steemers et al., 2006). In Abb. 2-3 ist der Assay als Flussdiagramm vereinfacht dargestellt. In den ersten Schritten wird die eingegebene DNA über eine genomweite Amplifizierung (GWA) und anschließende Fragmentierung für die Hybridisierung auf dem Chip vorbereitet. Hybridisiert wird an 50bp große Oligomere deren 3'OH-Enden direkt an dem dazugehörigen SNP enden. Jeweils 100.000 baugleiche Oligomere haften auf einer 3µm großen Kugel („Bead“) auf der Chipoberfläche. Pro SNP sind im Durchschnitt 30 dieser „Beads“ auf einem Chips vertreten, womit eine Redundanz der Signale gewährleistet ist. Unter stringenten Hybridisierungsbedingungen wird gewährleistet, dass keine unspezifischen Hybridisierungen vorhanden sind. Anschließend fügt eine DNA-Polymerase ein einzelnes markiertes Basenpaar (Cytosin (C), Guanin (G) mit Biotin; Adenin (A), Thymin (T) mit einem Dinitrophenyl) entsprechend dem SNP-Genotyp an. Nach einer abschließenden Fluoreszenzfärbung werden die Chips mit einem hochauflösendem Scanner abgelesen. Die Positionen der „Beads“ auf einem Chip müssen dabei eingegeben werden, da diese für jeden einzelnen Chip einzigartig sind. Für die Identifizierung der Positionen tragen die SNP-Oligomere synthetische Erkennungssequenzen, die von Illumina über mehrere Hybridisierungsreaktionen mit fluoreszenzmarkierte Oligomere dekodiert werden (Gunderson et al., 2004).

¹¹ Illumina, online: <http://www.illumina.com>

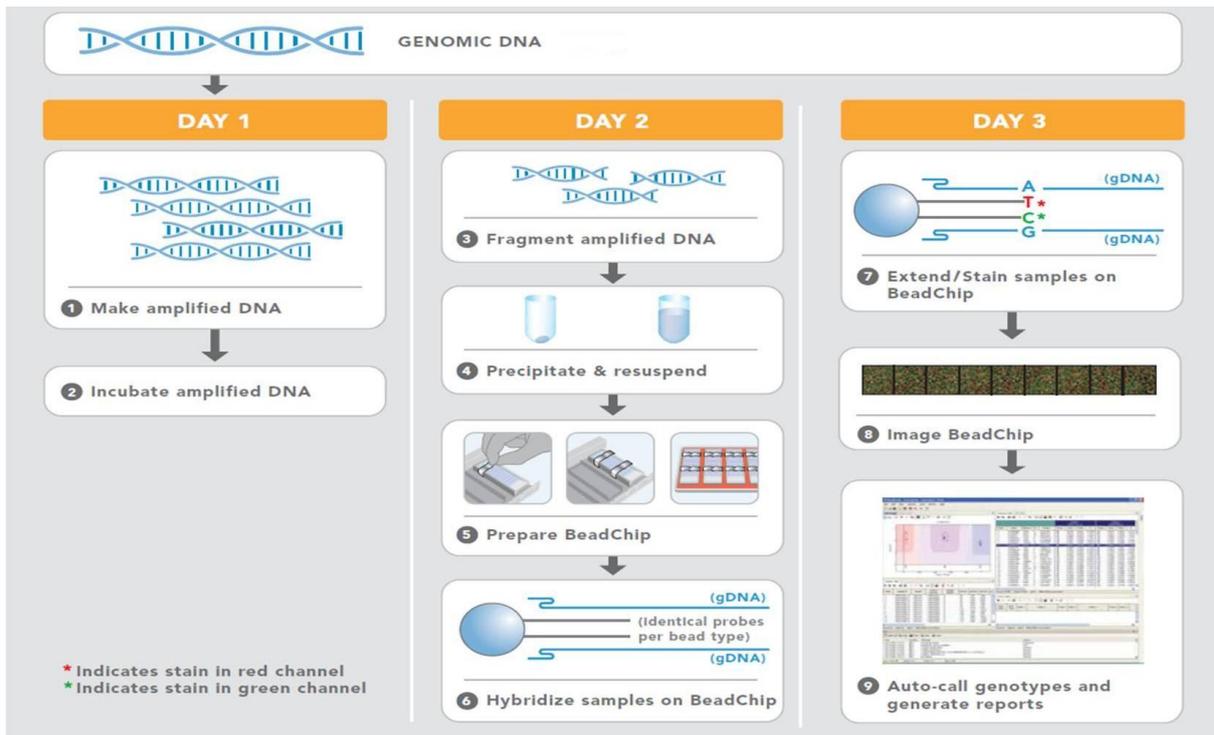


Abb. 2-3 Flussdiagramm des *Illumina Infinium II*-Assays
(nach *Infinium II Assay Workflow*; Illumina, 2006)

Für die Genotypbestimmung wurde die GenomeStudio Software v2011.1 verwendet. Der automatische „Calling“-Algorithmus der GenomeStudio Software benutzt vordefinierte Intensitätsbereiche für die Allelbestimmung und ist unabhängig von der analysierten Probenanzahl. Dabei wird über den Parameter „Gencall Score“ (GC-Score, Standard 0,15) die maximal zulässige Entfernung eines Signals vom Zentrum des am nächsten liegenden Intensitätsbereiches bestimmt. Der „Calling“-Algorithmus von Illumina versagt, wenn die genotypisierten Proben, aufgrund von Batcheffekten starke Intensitätsverschiebungen aufweisen. Aufgrund vorgegebener Intensitätsbereiche werden den Personen außerhalb ihres zugehörigen Intensitätsbereiches automatisch keine oder fehlerhafte Allele zugeordnet. Zusätzlich sind SNPs auf dem X-Chromosom fehleranfällig, weil keine Geschlechtsinformation für die Allelzuordnung benutzt wird.

Die GenomeStudio Software bietet die Möglichkeit fehlerhaftes „Genotype-Calling“ zu erkennen und manuell oder automatisch zu berichtigen. Die absolute Mehrheit der SNPs mit fehlerhaften Clusterzuordnungen können anhand folgender Parameter identifiziert werden: niedrige Genotypisierungsraten („call rate“, CR), schlechte Clustertrennung, niedrige durchschnittliche Intensitäten, niedrige Allelfrequenzen des „Minor“-Allels (MAF) und Heterozygotieüberschuß (Technischer Hinweis; Illumina, 2010). Bei Angabe von Familienstrukturen kann auch nach Mendelfehlern gescreent werden.

Genotypisierungschips oder Personen mit einer schlechten Assay-Performanz werden auf Chipsebene über interne Kontrollen identifiziert. Auf SNP-Ebene wird eine schlechte Assay-Performanz über die CR oder das Verhältnis von CR zu GC-Score identifiziert. Im Fall von Kontaminationen ist man nicht auf die Heterozygotierate angewiesen, die erst bei signifikanten Kontaminationen messbar vom Durchschnitt abweicht. Kontaminationen lassen sich bei einer Person über einen Plot aller SNP-Intensitäten darstellen. Anhand dieses Plots lässt sich auch die DNA mit der kontaminiert wurde identifizieren, sofern diese mit demselben Chip genotypisiert worden ist.

In dieser Studie wurden sämtliche auffälligen SNPs in allen „Calling“-Runden überprüft, berichtigt oder ausgeschlossen. Diese lückenlose Vorgehensweise führt zu einer sehr hohen Reliabilität der SNP-Genotypen. Für Multipoint Kopplungsstudien ist eine hohe Genotyp Reliabilität essentiell. Abgesehen von Artefakten (Typ-I-Fehler) führen fehlerhaft genotypisierte SNPs vor allem zu einem hohen Verlust an Informativität (Power, Typ-II-Fehler). Bereits eine Fehlerrate von nur 0,5% führt zu einer Reduzierung der maximalen erreichbaren LOD Scores um 10-30% (Douglas et al., 2000).

2.5 Erstellung und Integritätsüberprüfung von Genotyp-Eingabedaten

Erstellung

Die Erstellung von Genotyp-Eingabedaten für die Auswerteprogramme ist ein entscheidender Schritt. Bei genomweiten Daten kann dieser Schritt einen höheren Zeitaufwand verursachen, als die anschließenden statistischen Berechnungen.

Affymetrix- und SNPstream-Daten wurden mit Linux Standardprogrammen auf Kommandozeilenebene konvertiert. Vor allem im Fall von sehr großen Datenmengen bietet diese Vorgehensweise einen bedeutenden Zeitvorteil im Vergleich zu anderen geläufigen Programmiersprachen. Häufig gebrauchte Befehlsketten wurden dabei als Bash-Skripte automatisiert. Illumina-Daten wurden mittels ALOHOMORA¹² (Rüschendorf und Nürnberg, 2005) und Perl-Skripten für Kopplungsanalysen konvertiert. Mikrosatelliten- und SNP-Daten wurden, geordnet nach der geschlechtsneutralen genetischen Position auf der Rutgers map v.2 (Matisse et al., 2007), vereinigt. Die genetischen Positionen wurden mithilfe des Webtools Rutgers Map Interpolator¹³ ermittelt.

¹² ALOHOMORA, online: <http://gmc.mdc-berlin.de/alohomora>

¹³ Rutgers Map Interpolator, online: <http://compgen.rutgers.edu/mapinterpolator>

Integritätsüberprüfung

Nach der Erstellung von Eingabedaten kann eine Überprüfung der Geschlechtszuordnung durch Analyse der Heterozygotie für X-chromosomale SNPs etwaige Zuordnungsfehler identifizieren. Die Integrität von Familien- oder Triodaten wird über die Überprüfung der Stammbaumstruktur bestätigt. Zur Überprüfung der Stammbaumstruktur wurde PEDIGREE-EXPLORER¹⁴ (Steffens, Institut für Medizinische Biometrie, Informatik und Epidemiologie, IMBIE) verwendet. Um das Vorhandensein von unbekanntem Verwandtschaftsverhältnissen oder duplizierten Proben auszuschließen, wurden die genetischen Ähnlichkeiten zwischen allen Personenpaaren, auf Grundlage der Bewertung der Markerähnlichkeit nach Allelzustand („identity by state“, IBS) mit PLINK¹⁵ (Purcell et al., 2007) und GRR¹⁶ (Abecasis et al., 2001) überprüft.

Bereinigung von Genotypisierungsfehlern

Abschließend wurden vor jeder Analyse die bereinigten Eingabedaten nach potentiellen Genotypisierungsfehlern überprüft. Im Fall von Familiendaten wurden Vererbungsfehler mit PedCheck¹⁷ (O'Connell und Weeks, 1998) oder PLINK entfernt. Verbleibende potentielle Genotypisierungsfehler können mithilfe von MERLIN¹⁸ (Abecasis et al., 2002) über Vergleiche der Haplotypen verwandter Personen identifiziert werden. Im Fall von unverwandten Fall-Kontroll-Daten können Genotypisierungsfehler, außer durch eine Überprüfung des Genotyp-Clusterings, nicht direkt erkannt werden. SNPs mit Hardy-Weinberg-Gleichgewicht-Verletzungen (Kap. 2.7), niedrige Allelfrequenzen oder niedrige Genotypisierungsraten sind Hinweise für potentielle Genotypisierungsfehler.

2.6 Hardy-Weinberg-Gleichgewicht

In einer idealen diploiden Population sind die Frequenzen von SNP-Genotypen einfache Funktionen ihrer Allelfrequenzen. Dieser Zusammenhang ist in der Populationsgenetik als Hardy-Weinberg-Gleichgewicht (HWE) bekannt (Hardy, 1908; Weinberg, 1908). Eine ideale Population wird gekennzeichnet durch: identische Allelfrequenzen der beiden Geschlechter, unendliche Populationsgröße (kein genetischer Drift in kleinen Populationen), Zufallspaarung (Panmixie), keine Selektion, keine Mutationen, keine Migration (Wanderung von Individuen zwischen verschiedenen Populationen).

¹⁴ PEDIGREE-EXPLORER, online: <http://pedigreeexplorer.meb.uni-bonn.de>

¹⁵ PLINK, online: <http://pngu.mgh.harvard.edu/~purcell/plink/index.shtml>

¹⁶ GRR, Graphical Representation of Relationships, online: <http://www.sph.umich.edu/csg/abecasis/GRR>

¹⁷ PedCheck, online: <http://watson.hgen.pitt.edu/register/docs/pedcheck.html>

¹⁸ MERLIN, online: <http://www.sph.umich.edu/csg/abecasis/Merlin/index.html>

Bei einem diallelischen Locus mit den Allelen A, B und die dazugehörigen Allelhäufigkeiten p_A und p_B , ergeben sich folgende Gleichungen für die Genotyphäufigkeiten:

$$P(AA) = p_A^2 \qquad P(AB) = 2 \times p_A \times p_B \qquad P(BB) = p_B^2$$

Dabei gilt:

$$p_A^2 + 2p_A p_B + p_B^2 = 1$$

Da in der Realität keine idealen Populationen existieren, gilt das HWE für relativ große Populationen, in denen alle Effekte die Allelfrequenzen verändern können, weniger ins Gewicht fallen. Desweiteren konnte gezeigt werden, dass nicht alle Bedingungen einer idealen Population notwendig für das Erreichen eines HWE sind (Stark, 2006).

Genotypfrequenzen, die sich im Hardy-Weinberg-Ungleichgewicht (HWD) befinden, können auf eine versteckte Populationsstruktur (Kap. 2.9) und Genotypisierungsfehler hinweisen. In Kohorten mit ausschließlich betroffenen Individuen kann ein HWD auch auf einer Assoziation eines Genotyps mit dem Phänotyp hinweisen (Wang und Shete, 2012). Allgemein wird deswegen der Test auf HWE in Fall-Kontroll-Studien ausschließlich in der Gruppe der Kontrollen angewandt. Da dieser Tests nur im Fall von hohen Fehlerraten Genotypisierungsfehler anzeigt, ist die Eignung des HWE-Tests in Frage gestellt worden (Salanti et al., 2005; Zou und Donner, 2006). Zusätzlich tendiert der Test zu einer überhöhten Typ-I-Fehlerrate, die bei einer hohen Anzahl von Tests mit sehr niedrigen Signifikanzschwellen kontrolliert werden muss ($P < 10^{-7}$; McCarthy et al., 2008). Für die Ermittlung von HWD wurde in dieser Arbeit der exakte HWE-Test verwendet (Wigginton et al., 2005), welcher in PLINK implementiert ist und eine verringerte Typ-I-Fehlerrate aufweist.

2.7 Rekombination und Kopplungsungleichgewicht

Allele eines Markers, die sich auf demselben Chromosom befinden, können gemeinsam weitervererbt werden. Die Wahrscheinlichkeit dafür ist abhängig von der physikalischen und der genetischen Distanz. Die genetische Distanz beschreibt die durchschnittliche Rekombinationsrate zwischen zwei Loci (Strachan und Read, 2005: 463). Die physikalischen und genetischen Abstände sind zueinander nicht konstant, weil die Rekombinationswahrscheinlichkeiten vom Geschlecht (Kong et al., 2002) und von dem chromosomalen Bereich abhängen (Jeffreys et al., 2001). Die Rekombinationshäufigkeit (Theta, θ) ist nie größer als 50% ($\theta = 0,5$), auch wenn die Distanz zwischen zwei Loci mehr als nur ein einziges Crossing-over-Ereignis zulässt. Bei der maximalen genetischen Distanz ($\theta = 0,5$) sind die Allelverteilungen zweier Loci voneinander völlig unabhängig. Dieser Zustand wird als Kopplungsgleichgewicht bezeichnet. In allen anderen Fällen ($\theta < 0,5$) sind die Loci gekoppelt und ihre jeweiligen

Allele treten häufiger gemeinsam auf, als bei zufälliger Verteilung zu erwarten wäre. Als Maß für diese Kopplung dient das Kopplungsungleichgewicht („linkage disequilibrium“, LD).

Für zwei diallelische Loci M_1, M_2 mit den Allelen A_i, B_i ($i = 1,2$) und den Allelhäufigkeiten p_{A_i}, p_{B_i} , ergeben sich die Haplotyphäufigkeiten aus dem Produkt der Allelhäufigkeiten. Bei Kopplungsungleichgewicht ergibt sich:

$$P(A_i B_j) = p_{A_i} \times p_{B_j} \quad i, j \in \{1,2\}$$

Die Abweichung von diesem Produkt ergibt das Kopplungsungleichgewicht:

$$D_{ij} = P(A_i B_j) - p_{A_i} \times p_{B_j} \quad i, j \in \{1,2\}$$

Da das LD stark von den Allelfrequenzen abhängt, wird das Disequilibriumskoeffizient D_{ij} als Maß für das LD normiert (Lewontin et al., 1988). Bei einem diallelischen Locus sind durch $D = D_{11}$, im Bezug auf dem ersten Haplotyp $A_1 B_1$ ($D_{11} = P(A_1 B_1) - p_{A_1} \times p_{B_1}$) die Abweichungen der anderen drei Haplotyphäufigkeiten festgelegt (Tabelle 2-4).

Allele	A_1	A_2
B_1	$p_{A_1} \times p_{B_1} + D$	$p_{A_2} \times p_{B_1} - D$
B_2	$p_{A_1} \times p_{B_2} - D$	$p_{A_2} \times p_{B_2} + D$

Tabelle 2-4 Kopplungsungleichgewicht

Die Normalisierung erfolgt über Division mit dem maximal möglichen Absolutbetrag:

$$D_{Max} = \min(p_{A_1} \times p_{B_2}, p_{A_2} \times p_{B_1}), \text{ falls } D_{11} > 0$$

oder

$$D_{Max} = \min(p_{A_1} \times p_{B_1}, p_{A_2} \times p_{B_2}), \text{ falls } D_{11} < 0$$

Damit gilt $D' = D_{11} / D_{Max}$. Dabei bedeutet $D' = 0$ absolutes Kopplungsgleichgewicht, während maximales LD bei $D' = 1$ vorliegt. Aus D leitet sich ein alternatives Maß für LD, das Korrelationsmaß r^2 (Slatkin, 2008):

$$r^2 = [P(A_1 B_1) \times P(A_2 B_2) - P(A_1 B_2) \times P(A_2 B_1)]^2 / p_{A_1} \times p_{B_1} \times p_{A_2} \times p_{B_2}$$

Das LD ist nicht stabil, es sinkt im Verlauf jeder neuen Generation t in Abhängigkeit von den Rekombinationsfrequenz θ :

$$D_{11}(t+1) = (1 - \theta) \times D_{11}(t)$$

Auf LD zwischen Markern basieren zahlreiche genetische Kartierungsmethoden. Bei Kopplungs- und Assoziationsstudien wird LD-Information verwendet, um unbekannte krankheitsrelevante Variationen zu identifizieren, die im starken LD zu den bekannten, genotypisierten Markern stehen. Die Abhängigkeit der Allelverteilungen zwischen Loci im LD wird bei SNPs dazu benutzt, um fehlende SNP-Genotypen anhand von benachbarten,

genotypisierten SNPs herzuleiten (Marchini und Howie, 2010). Für dieses Verfahren hat sich die Bezeichnung „Imputing“ durchgesetzt. Das Ziel dieser Verfahrensweise, ist in der Regel, die Anzahl der SNPs zu maximieren, die in einer GWAS getestet werden. Solche „in silico“ hergeleitete Genotypen führen bei einer GWAS zu einer Erhöhung der Wahrscheinlichkeit (Power) echte Assoziationen festzustellen (Becker et al., 2009). Zusätzlich werden Meta-Analysen von Studien ermöglicht, die auf unterschiedliche SNP-Chip-Daten basieren.

2.8 Populationsstratifizierung

Separationen durch geographische, politische oder kulturelle Barrieren gefolgt von genetischer Drift, Selektion und Mutationen, führen zur Bildung von Unterpopulationen, bzw. Ausbildung von Strata innerhalb einer Population. Solche, meist als ethnische Gruppen erkennbare, Populationsgruppen können unterschiedliche Frequenzen von SNP-Allelen oder krankheitsrelevanten Variationen aufweisen. Durch assortative Paarungen können sich weitere Untergruppierungen innerhalb von ethnischen Gruppen ausbilden (versteckte Populationsstruktur). Da genetische Assoziation auf der Differenz von Allelfrequenzen zwischen Gruppen mit unterschiedlichen Phänotypen basieren, kann Populationsstratifizierung zu Artefakten, bzw. künstlichen Assoziationen führen, die nicht Phänotyp-/Krankheitsrelevant sind (Thomas und Witte, 2002).

Auf der Basis von genomweiten SNP-Genotypdaten lässt sich durch Hauptkomponentenanalysen („principal components analysis“, PCA) oder multidimensionaler Skalierung („multidimensional scaling“, MDS), Populationsstratifizierung erkennen und abbilden (Abb. 2-4). In europäische Populationen sind die genetischen und geographischen Distanzen stark miteinander korreliert (Lao et al., 2008).

Das Prinzip beider Methoden ist die Verdichtung von Daten (in diesem Fall die genetische Information in Form von X Individuen beschrieben durch Y SNPs) in unabhängige Hauptkomponenten, die nach ihrem Anteil zu der totalen Varianz der Daten sortiert werden.

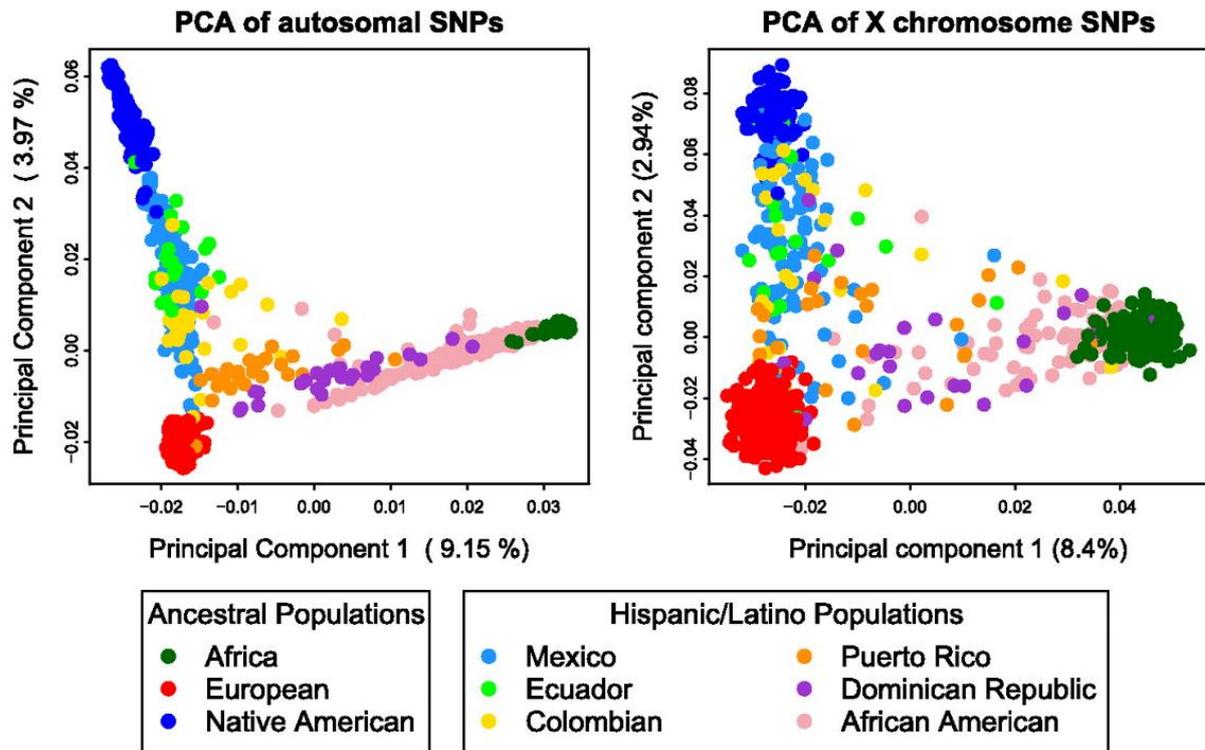


Abb. 2-4 Weltweite Populationsstruktur abgebildet durch PCA

Dargestellt sind die ersten zwei Hauptkomponenten („principal components“) einer PCA, mit autosomale oder geschlechtschromosomale SNPs. In Klammern ist der Anteil zu der totalen Varianz der jeweiligen Hauptkomponente wiedergegeben (verändert nach Bryc et al., 2010).

Die ersten Hauptkomponenten enthalten damit die meiste Information, während der Informationsgehalt, mit steigender Hauptkomponentennummer gegen Null tendiert. PCA lässt sich mit der EIGENSOFT¹⁹ Software (Patterson et al., 2006; Price et al., 2006) und MDS mit PLINK durchführen. Da beide Methoden zu ähnlichen Ergebnissen führen (Wang et al., 2009), wurde in dieser Arbeit aufgrund der direkteren Auswertbarkeit von Ergebnissen PLINK-MDS verwendet. Um sicherzustellen, dass die MDS die reine Populationsstruktur abbildet, wurden potentiell fehlgenotypisierte SNPs (Qualitätskriterien: CR > 0,99; MAF > 0,15, HWE $P > 10^{-3}$), SNPs aus Regionen mit weitreichenden LD-Muster (Langstrecken-LD), SNP-Paare mit $r^2 > 0,1$, sowie geschlechtschromosomale SNPs ausgeschlossen. Starkes LD-Muster kann durch Redundanz einen signifikanten Anteil des Informationsgehaltes einzelner Hauptkomponenten ausmachen, so dass diese nicht mehr die echte Populationsstruktur darstellen (Fellay et al., 2007; Price et al., 2008).

¹⁹ EIGENSOFT, online: http://genetics.med.harvard.edu/reich/Reich_Lab/Software.html

Chromosom	Position Anfang (NCBI 36.3)	Position Ende (NCBI 36.3)
1p33-p32.3	48000000	52100000
2p11.2-q11.2	85941853	100500000
2q21.2-q22.1	134382738	138000000
2q32.1-q32.2	182882739	190000000
3p21.31	47500000	50000000
3p12.2-p12.1	83500000	87000000
3p11.2-q11.2	89000000	97500000
3q29	196882966	198870687
5p12-q11.2	44000000	51500000
5q21.1	98000000	100500000
5q23.3-q31.1	129000000	132000000
5q31.2	135500000	138500000
6p22.2-p21.32	25000000	33500000
6p12.1-q12	57000000	64000000
6q24.1	140000000	142500000
7p11.2-q11.22	55000000	66193285
8p23.1	7225962	12487029
8p11.21-q11.21	43000000	50000000
8q23.2-q23.3	112000000	115000000
10p11.21-q11.21	37000000	43000000
11p11.2-q12.1	45000000	57000000
11q14.2-q14.3	87500000	90500000
12p11.21-q12	33000000	40000000
12q24.11-q24.13	109500000	112021663
15q13.2-q13.3	28524207	30602466
15q24.1-q24.2	72151413	73356183
17q12	31888441	33393152
17q21.31-q21.32	40899921	41989253
20q11.22-q11.23	32000000	34500000

Tabelle 2-5 Bekannte Regionen mit außergewöhnlichem Langstrecken-LD

(zusammengetragen aus: Fellay et al., 2007; Price et al., 2008; Deng et al., 2008; Antonacci et al., 2009; Kasperaviciute et al., 2010)

In Tabelle 2-5 sind die Regionen gelistet, die aufgrund von Langstrecken-LD von der MDS ausgeschlossen worden sind. LD zwischen weit entfernten Markern kann durch Inversionspolymorphismen bedingt sein, die hohe Frequenzen in der untersuchten Bevölkerungsgruppe aufweisen (Bansal et al., 2007).

2.9 Statistische Tests

Statistische Testverfahren werden verwendet, um eine Aussage über die Gültigkeit oder Ungültigkeit einer aufgestellten Hypothese zu erhalten. Dabei werden zwei Hypothesen gegeneinander getestet. In der Genetik wird die Hypothese getestet, dass eine Variation, oder eine genetische Region im Zusammenhang mit dem untersuchten Phänotyp steht. Ausgehend

von der Nullhypothese (H_0), dass kein Zusammenhang mit dem untersuchten Phänotyp besteht, wird diese Alternativhypothese (H_1) getestet (Tabelle 2-6).

Entscheidung	Wirklichkeit	
	Nicht-Ablehnung von H_0	Richtige Nicht-Ablehnung ($1 - \alpha$)
Ablehnung von H_0	Falsche Ablehnung (α -, Typ-I-Fehler)	Richtige Ablehnung ($1 - \beta$, Power)

Tabelle 2-6 Entscheidungen des Signifikanztests

Ein solcher Test liefert keine absolute Aussagen sondern Wahrscheinlichkeiten und gilt deswegen als Signifikanztest. Die Nullhypothese wird abgelehnt, wenn das Ergebnis des Tests, der Signifikanzwert (P-Wert), unter einem im vorraus vorgegebenen Signifikanzniveau liegt. Das Signifikanzniveau α ist dabei die maximal zugelassene Wahrscheinlichkeit eines Typ-I-Fehlers. Auch wenn das Signifikanzniveau frei wählbar ist, hat sich in der wissenschaftlichen Literatur $\alpha = 0,05$ (5%) etabliert. Vereinfacht dargestellt bedeutet ein $\alpha = 0,05$ dass maximal einer von 20 Tests die Nullhypothese rein zufällig ablehnen wird. Die Typ-I-Fehlerrate beträgt damit 0% bis 5%, je nachdem ob H_0 falsch oder richtig ist (Goodman et al., 2008). Die Typ-I-Fehlerrate steigt also direkt mit einer größer werdenden Anzahl von Tests an. Das Problem des multiplen Testens wird über die Bonferroni-Korrektur beseitigt (Bland und Altman, 1995). Dabei ergibt die Division der Typ-I-Fehlerwahrscheinlichkeit durch die Zahl der durchgeführten Tests (α / n -Tests) ein neues Signifikanzniveau. In Tabelle 2-7 ist das Problem des multiplen Testens am Beispiel einer Studie mit 1.000.000 unabhängigen Tests wiedergegeben. Da bei einem Signifikanzniveau von 0,05, 50.000 Artefakten zu erwarten wären, beträgt ein angemessenes Signifikanzniveau $0,05 / 1.000.000 = 5 \times 10^{-8}$. In dieser Arbeit wird der Ausdruck „Artefakte“ für Typ-I-Fehler verwendet, weil der allgemein gebräuchliche Ausdruck „falsch-positiv“ im Fall von Assoziationsstatistiken irreführend ist.

Signifikanzniveau α	Artefakte (Typ-I-Fehler)
0,05	50000
0,01	10000
0,001	1000
1,00E-04	100
1,00E-05	10
1,00E-06	1
5,00E-07	0,5
1,00E-07	0,1
5,00E-08	0,05
1,00E-08	0,01

Tabelle 2-7 Anzahl an Artefakten bei 1 Million unabhängiger Tests

Die falsche Nicht-Ablehnung von H_0 führt zu Typ-II-Fehler. Die Wahrscheinlichkeit für Typ-II-Fehler (β) kann nur als Funktion des Stichprobenumfanges bei einem festgelegten Signifikanzniveau berechnet werden. Je größer der Stichprobenumfang und das gewählte Signifikanzniveau α , desto niedriger fällt β aus. Da bei genomweiten Studien die durchgeführten Tests nicht völlig voneinander unabhängig sind, wird β mit steigender Anzahl an Tests kleiner. Direkt von β leitet sich die Wahrscheinlichkeit ab, dass alle falschen Nullhypothesen einer Studie abgelehnt werden ($\text{Power} = 1 - \beta$). Für die Einschätzung der Power der GWAS-Studiekollektive wurde in dieser Studie der CaTS Power Calculator²⁰ (Skol et al., 2006) verwendet. Die Power einer GWAS ist, neben Stichprobenumfang und Signifikanzniveau, von weiteren Faktoren abhängig: Häufigkeit des Phänotyps, genetisches Modell, relatives Risiko (Kap. 2.11.1) und Allelfrequenz der krankheitsrelevanten Variation, LD zwischen getesteten Marker und krankheitsrelevante Variation.

2.10 Fall-Kontroll-Analysen

2.10.1 Quotenverhältnis und Relatives Risiko

In der Epidemiologie sind das Quotenverhältnis („Odds Ratio“, OR) und das Relative Risiko (RR) statistische Maßzahlen für die Stärke des genetischen Effektes eines Risikofaktors für einen bestimmten Phänotyp. Es werden dabei Personen mit und ohne einen Risikofaktor bezüglich des Auftretens einer Erkrankung verglichen (Tabelle 2-8).

	mit Risikofaktor	ohne Risikofaktor
Fälle	a	b
Kontrollen	c	d

Tabelle 2-8 Vierfeldertafel für die Berechnung des ORs

Dabei gilt:

$$RR = [a / (a + c)] / [(b / (b + d))]$$

$$OR = (a / c) / (b / d) = (a \times d) / (c \times b)$$

Das RR definiert die Wahrscheinlichkeit, dass eine Krankheit bei den Trägern eines Risikofaktor auftritt ($a / (a + c)$) relativ zu der Wahrscheinlichkeit, dass dieselbe Erkrankung bei Personen auftritt, die den Risikofaktor nicht aufweisen ($b / (b + d)$). Das OR definiert das Chancenverhältnis zu erkranken in einer Gruppe mit dem Risikofaktor (a / c), verglichen zu der Gruppe ohne Risikofaktor (b / d). Die zwei Kenngrößen liefern vergleichbare Ergebnisse je niedriger die Inzidenzrate der untersuchten Erkrankung ist:

²⁰ CaTS Power Calculator, online: <http://www.sph.umich.edu/csg/abecasis/CaTS>

wenn: $a + c \sim c$ und $b + d \sim d$; dann gilt $OR \sim RR = (a / c) / (b / d)$

Beide Kenngrößen nehmen Werte zwischen 0 und ∞ an und werden folgendermaßen interpretiert:

RR bzw. OR	Interpretation bezüglich der Entstehung einer Erkrankung
< 1	reduziertes Risiko
= 1	kein Einfluss auf die Erkrankung
> 1	erhöhtes Risiko

Tabelle 2-9 Interpretation von RR und OR

Eine vollständige Interpretation der Kenngrößen ist erst mit dem dazugehörigen Konfidenzintervall (KI) möglich. Der Konfidenzintervall beschreibt den Wertebereich, in dem sich mit einer vorher vorgegebenen Wahrscheinlichkeit der wahre Wert des untersuchten Merkmals befindet. In der medizinischen Statistik wird in der Regel eine 95%-ige Wahrscheinlichkeit verwendet, die zu dem 95%-Konfidenzintervall (95%-KI) führt. Aus diesem Konfidenzintervall lässt sich entnehmen, ob die berechnete Effektgröße eines Risikofaktors statistisch signifikant ist. Das OR oder RR eines Risikofaktors ist dann statistisch signifikant, wenn der Wert 1 nicht zu dem 95%-KI gehört.

2.10.2 Logistische Regression

Wie in Kap. 2.9 beschrieben kann Populationsstratifizierung bei Fall-Kontroll-Analysen zu Assoziationen führen, die nicht Phänotyp-/Krankheitsrelevant sind, sondern auf eine Schichtung der Stichprobe zurückzuführen sind (Artefakte). Eine Schichtung der Stichprobe aufgrund von unterschiedlichen Genotypisierungsbedingungen führt ebenfalls zu Artefakten. In einem logistischen Regressionsmodell mit der Schichtzugehörigkeit aller Personen als Kovarianten, kann die Wahrscheinlichkeit für das Vorkommen von stratifizierungsbedingten Artefakten reduziert werden (Zhu et al., 2002; Price et al., 2006). Im Fall von genomweiten Daten werden die Populationschichten mithilfe einer MDS oder PCA-Analyse bestimmt. Die Hauptkomponenten, die eine Schätzung der Zugehörigkeit der Personen zu den Schichten darstellen, dienen dann als Kovarianten in der logistischen Regression.

GWAS-P-Werte wurden in dieser Studie mit PLINK über das additive Modell der logistischen Regression mit dem Geschlechtsstatus als zusätzliche Kovariante bestimmt. Das additive Modell bedeutet, dass das Risiko von Personen mit homozygotem Genotyp für das Risikoallel doppelt so groß ist wie das Risiko von heterozygoten Trägern des Risikoallels.

2.10.3 Stratifizierte Assoziationsanalysen

Eine direktere Methode eine Populationsstratifizierung zu kontrollieren bietet sich an, wenn die Personen einer Stichprobe einer abgrenzbaren Schicht bzw. Subgruppe zugeordnet werden können. Im Fall von genomweiten Daten wird eine solche Zuordnung durch Clustering der Personen anhand ihrer genetischen Distanzen realisiert (Pritchard et al., 2000). Wenn genomweite Daten fehlen, führt die Zuordnung der Personen in Cluster nach der ethnischen Herkunft zu vergleichbaren Ergebnissen (Wang et al., 2010; Chokkalingam et al., 2011).

Für stratifizierte Analysen mit Fall-Kontroll-Daten wurde in dieser Studie der bei PLINK implementierte Cochran-Mantel-Haenszel- (CMH-) 2x2xK Test verwendet. Bei diesem Test wird jede Schicht oder Cluster K der Kontingenztabelle als partielle Vierfeldertafel (2x2) behandelt. Die Teststatistik basiert damit auf das durchschnittliche Quotenverhältnis (OR) über alle Schichten. Die Validität eines solchen Tests ist dabei abhängig von einer homogenen Verteilung der ORs in den verschiedenen Cluster. Die homogene Verteilung der ORs kann mit dem Breslow-Day-Test (H_0 = homogene ORs, gleiche Effektrichtung in allen Strata) überprüft werden.

2.10.4 Genomic Control

Die „Genomic Control“- (GC-) Methode wird verwendet, um Assoziationstests anhand der Streuung ihrer Teststatistiken zu korrigieren. Durch diese Korrektur kann bei vorhandener Populationsstratifizierung die Wahrscheinlichkeit für Typ-I-Fehler zu reduziert werden (Devlin und Roeder, 1999; Marchini et al., 2004; Devlin et al., 2004). Bei vorhandener Populationsstratifizierung kann der Median der Teststatistiken größer ausfallen als der Median der erwarteten Chi-Quadrat (X^2) Verteilung. Aus den berechneten X^2 Werte einer GWAS wird der „genomic inflation“ Faktor λ_{GC} errechnet:

$$\lambda_{GC} = \text{median}(X_1^2, X_2^2, X_3^2, \dots, X_n^2) / 0,456$$

GC korrigierte P-Werte erhält man durch Multiplikation ihrer Teststatistiken mit dem λ Faktor. Da der λ_{GC} Faktor größer wird mit steigender Stichprobengröße, kann für einen Vergleich von unterschiedlichen Studien ein zu einer Stichprobengröße von 1.000 Fällen und 1.000 Kontrollen äquivalenter λ_{GC} Faktor berechnet werden (Freedman et al., 2004; de Bakker et al., 2008):

$$\lambda_{1000} = 1 + (\lambda - 1) \times [(1 / N_{\text{Fälle}} + 1/N_{\text{Kontrollen}}) / (1 / 1000 + 1 / 1000)]$$

„Genomic Control“ wurde in dieser Studie mittels PLINK angewandt.

2.11 Familien-basierte Analysen

2.11.1 Transmission-Disequilibrium-Test

Der Transmission-Disequilibrium-Test (TDT) wird bei Trio-Daten angewandt und ist ein Test auf Assoziation und Kopplung (Spielman et al., 1993). Die getesteten Hypothesen lauten: H_0 = keine Kopplung oder keine Assoziation liegt vor; H_1 = sowohl Kopplung als auch Assoziation liegen vor. Mit diesem Test können Varianten erfasst werden, die signifikant häufiger an betroffene Kinder transmittiert werden, als zufällig zu erwarten wäre. Pro Trio werden die transmittierten und nichttransmittierten Allele der Eltern mit heterozygotem Genotyp gezählt (Tabelle 2-10).

		transmittiert		Gesamt
		A ₁	A ₂	
nicht transmittiert	A ₁	n ₁₁	n ₂₁	n ₁₁ + n ₂₁
	A ₂	n ₁₂	n ₂₂	n ₁₂ + n ₂₂
Gesamt		n ₁₁ + n ₁₂	n ₂₁ + n ₂₂	2n

Tabelle 2-10 Anzahl übertragener Allele eines biallelischen Markers für n Trios

H_0 ist $n_{21} = n_{12}$, also wird die Teststatistik von McNemar angewandt (McNemar, 1947):

$$T_{\text{TDT}} = (n_{21} - n_{12})^2 / (n_{21} + n_{12})$$

Der TDT bleibt auch bei vorhandener Populationsstruktur bzw. Unterschiede in den Allelverteilungen zwischen verschiedenen Strata einer Stichprobe, als Assoziationstest valide ohne Erhöhung der Wahrscheinlichkeit für Typ-I-Fehler (Ewens und Spielman, 1995). Allerdings bedeutet diese Eigenschaft nicht, dass der TDT ohne Rücksicht auf die Populationsstruktur angewandt werden darf. Bei vorhandener Populationsstratifizierung, ist die Wahrscheinlichkeit für Typ-II-Fehler des TDTs erhöht (Sebro und Rogus, 2010). Dieser Verlust an Power ist durch die reduzierte Heterozygotie einer Population, die in Unterpopulationen stratifiziert ist, zu erklären (Wahlund-Effekt). Die reduzierte Heterozygotie resultiert in einer niedrigeren Häufigkeit von Eltern mit heterozygoten Genotypen und damit zu einem Informationsverlust beim TDT.

2.11.2 Fall-Pseudo-Kontroll-Methode

Die Fall-Pseudo-Kontroll- („affected family-based controls“-, AFBAC-) Methode basiert, analog zum TDT, auf Alleltransmissionen zu einem betroffenen Kind (Terwilliger und Ott, 1992; Thomson, 1995). Im Gegensatz zu dem TDT wird bei dieser Methode die vollständige Infor-

mation der Eltern verwendet. Die nichttransmittierten Allele der Eltern werden dabei zu einer virtuellen Kontrolle zusammengefügt (Abb. 2-5). Die Methode entspricht dann einer Fall-Kontroll-Analyse mit perfekt zugeordneten Fall-Kontroll-Kollektiven.

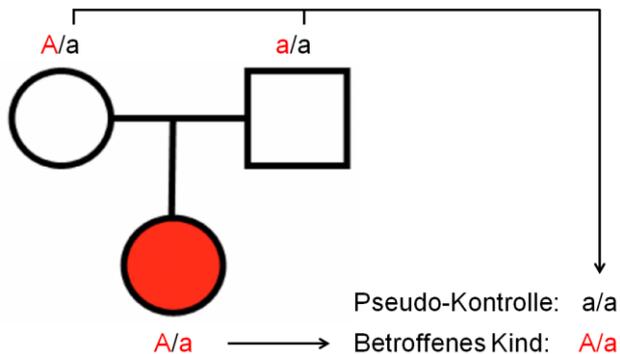


Abb. 2-5 Trio als Fall / Pseudo-Kontrolle

Dargestellt ist ein Trio mit einem biallelischen Marker A/a . Die jeweils transmittierten Allele sind rot gekennzeichnet.

Der Vorteil dieser Vorgehensweise ist die Möglichkeit Trio-Daten einer stratifizierten Fall-Kontroll-Analyse unterziehen zu können. Aufgrund der Wahrscheinlichkeit des TDTs für Typ-II-Fehler (Kap 2.11.1) ergibt sich bei vorhandener Populationsstratifizierung, ein Powergewinn gegenüber dem TDT, ohne eine Erhöhung der Wahrscheinlichkeit für Typ-I-Fehler. Ein weiterer Vorteil dieser Vorgehensweise ist die Möglichkeit, Trio-Daten mit unverwandten Fällen und Kontrollen in einer gemeinsamen Analyse testen zu können. Mit einer gemeinsamen Analyse von Trios sowie unverwandten Fällen und Kontrollen kann eine höhere Power erreicht werden (Nagelkerke et al., 2004; Guo et al., 2009). Dabei bleibt der Test valide (bezüglich Typ-I-Fehlerrate), solange alle Personen aus derselben Population stammen (Guo et al., 2009). Stratifizierte Teststatistiken oder logistische Regressionsmodelle können angewandt werden, um die Wahrscheinlichkeit für Typ-I-Fehler zu kontrollieren (Crossett et al., 2010).

In dieser Studie wurde die AFBAC-Methode in der Replikationsrunde der GWAS angewandt. Die zwei vereinigten Trio- und Fall-Kontroll-Kollektive wurden eingeteilt in Cluster nach der ethnischen Herkunft und einer stratifizierten Analyse mit dem CMH-Test (Kap 2.10.3) unterzogen. Für die Konvertierung der Trio-Daten zu Fall-Pseudo-Kontrollen und anschließende Analyse wurde PLINK verwendet. Diese Vorgehensweise ist bereits erfolgreich für die Replikation von GWAS-Ergebnissen in einem ähnlich strukturierten Datensatz europäischen Ursprungs angewandt worden (Jin et al., 2010).

2.11.3 Parametrische Kopplungsanalysen

Mit parametrischen Kopplungsanalysen wird nach Markern getestet, die innerhalb einer Familie im LD mit einer genetischen Variation stehen, welche für die Ausprägung eines bestimmten Phänotyps verantwortlich ist (Bickeböllner und Fischer, 2007: 157-227). Das zugrundeliegende Vererbungsmodell muss dabei vorgegeben werden.

Das Vererbungsmodell wird über die Penetranzraten des Risikoallels f_n (n = Anzahl der Risikoallele) und das Vorkommen von erkrankten Personen, die kein Risikoallel aufweisen (Phänokopierate f_0) bestimmt (dominant: $f_2 = f_1 \gg f_0$; rezessiv: $f_2 \gg f_1 = f_0$).

Ausgehend von H_0 ($\theta = 0,5$; freie Rekombination zwischen den Loci) und H_1 ($\theta < 0,5$; Loci befinden sich im LD) wird das Ausmaß der Kopplung mittels dem dekadischen Logarithmus der Wahrscheinlichkeit für Kopplung ($L(H_1) / L(H_0)$) bemessen. Für Kernfamilien gilt (Morton, 1955):

$$\text{LOD}(\theta) = \log_{10} [L(\theta) / L(0,5)], \text{ mit } 0 \leq \theta < 0,5$$

Das globale Maximum der LOD-Funktion ergibt den LOD-Wert als Maß für Kopplung. Im Unterschied zu normalen statistischen Tests (Kap. 2.10) kann bei der LOD-Methode H_0 als richtig angenommen und nicht nur abgelehnt werden. Kopplung wird ausgeschlossen bei einem LOD-Wert ≤ -2 . Bei komplexen Krankheiten kann, unter Annahme von genetischer Heterogenität, die LOD-Methode auf eine beliebige Anzahl von Familien ausgeweitet werden. Die genetische Heterogenität wird geschätzt durch den Anteil der Familien die Evidenz für Kopplung zu demselben Locus erbringen (α). Der Heterogenitäts-LOD-Wert (HLOD) ergibt sich aus:

$$\text{HLOD}(\theta) = \log_{10} [L(\alpha, \theta) / L(\alpha = 0, \theta = 0,5)]$$

Wobei $L(\alpha, \theta)$ die Summe der einzelnen Wahrscheinlichkeiten $L_i(\alpha, \theta)$ der i -ten Familie aus einer Stichprobe mit n Familien darstellt (Abreu et al., 2002):

$$L_i(\alpha, \theta) = \sum_{i=1}^n \alpha \times L_i(\theta) + (1 - \alpha) \times L_i(0,5)$$

Die HLOD Funktion wird dann für beide Variablen maximiert. Wenn genetische Heterogenität gegeben ist, erreicht die HLOD-Methode eine höhere Power als nichtparametrische IBD-basierte Methoden (Kap. 2.11.4) (Hodge et al., 2002; Holliday et al., 2005).

Eine Kopplung gilt dann als wahrscheinlich, sobald $\text{HLOD} \geq 3$ ist. Ein HLOD-Wert von 3 bedeutet, dass die Wahrscheinlichkeit für eine Kopplung genau 1000mal größer ist als die Wahrscheinlichkeit, dass beide Loci frei rekombinieren können. Die Schwelle für suggestive Kopplung wird dabei definiert als die Schwelle die erreicht werden muss, damit pro Analyse eines ganzen Genoms H_0 maximal ein Mal falsch abgelehnt wird. Die Schwelle bei der H_0 maximal 0,05 mal pro Analyse eines ganzen Genoms falsch abgelehnt wird, definiert die

Schwelle für signifikante Kopplung (Strachan und Read, 2005: 517). Die zu erreichenden Signifikanzschwellen hängen von der Anzahl und Informativität der getesteten Markern ab (Lander und Kruglyak, 1995; Wiltshire et al., 2002). Mit steigender Marker Anzahl und Informativität steigt auch die Wahrscheinlichkeit für Typ-I-Fehler. Die Wahl des Vererbungsmodells beeinflusst nicht die Wahrscheinlichkeit für Typ-I-Fehler. Allerdings wird die Wahrscheinlichkeit für Typ-II-Fehler größer, je stärker das gewählte Vererbungsmodell von dem wirklichen Vererbungsmodell abweicht (Whittemore und Halpern, 2001; Hodge et al., 2002).

Für die Bestimmung der Schwellenwerte für suggestive und signifikante Kopplung in dem IGE-Datensatz, wurden empirische genomweite Typ-I-Fehlerraten mit MERLIN v1.1.2 basierend auf 5.000 Simulationen je Analysegruppe und für zwei Vererbungsmodelle (dominant, rezessiv) berechnet. Um der Analyse von zwei Vererbungsmodellen Rechnung zu tragen, wurde zu den berechneten Schwellenwerten 0.3 addiert (Abreu et al., 2002).

Für die Berechnung von genomweiten HLOD-Werten für die 379 IGE-Familien wurde ALLEGRO v2²¹ (Gudbjartsson et al., 2005) verwendet. Es wurden dabei zwei Vererbungsmodelle getestet: 1.) dominant: Allelfrequenz des Risikoallels = 0,01; $f_0 = 0,05\%$; $f_1 = 70\%$; $f_2 = 70\%$. 2.) rezessiv: Allelfrequenz des Risikoallels = 0,1; $f_0 = 0,05\%$; $f_1 = 0,05\%$; $f_1 = 70\%$. In der Analyse der Subgruppen wurden die Penetranzraten auf 50% reduziert. Die gewählten Vererbungsmodelle gewährleisteten eine breite Abdeckung der möglichen oligogenen Vererbungsmodelle, mit dem Ziel einer möglichst geringen Wahrscheinlichkeit für Typ-II-Fehler.

Die Power der Kopplungstudie und den jeweiligen Vererbungsmodellen wurde mit SLINK²² (Schäffer et al., 2011) geschätzt.

2.11.4 Nichtparametrische Kopplungsanalysen

Nichtparametrische Kopplungsanalysen sind modellfreie Kopplungsverfahren, die nicht auf der Schätzung der Rekombinationsraten basieren (Bickeböller und Fischer, 2007: 157-227). Basierend auf der Identität ererbter Anlagen („identity-by-descent“, IBD) zwischen betroffenen Mitgliedern einer Familie werden mittels „allele-sharing“-Methoden (Whittemore und Halpern, 1994; Kong und Cox, 1997) Markerallele identifiziert, die innerhalb einer Familie mit der Erkrankung kosegregieren.

Nichtparametrische „allele-sharing“ LOD-Werte (LOD_{NPL}) wurden mit ALLEGRO v2 nach dem linearen S_{all} Modell berechnet. Das lineare S_{all} Modell ist für komplexe Erkrankungen konzipiert und erkennt kleine Veränderungen des IBD-Status zwischen allen Betroffenen einer Familie kombiniert über eine große Anzahl von Familien (Kruglyak et al., 1996; Kong und Cox, 1997).

²¹ Allegro v2, online: <http://www.decode.com/software>

²² SLINK, online: <http://watson.hgen.pitt.edu/docs/SLink.html>

Die Schwellenwerte für genomweite suggestive und signifikante Kopplung wurden, analog zu Kap. 2.11.3, mit MERLIN berechnet.

2.12 Meta-Analysen

Meta-Analysen von getrennt ausgewerteten Assoziationsstatistiken, entweder im Rahmen von mehrstufigen GWAS oder vollkommen separaten Studien, erhöhen die Wahrscheinlichkeit für die Identifizierung von genetischen Varianten, die niedrige Effektgrößen aufweisen (Skol et al., 2006; de Bakker et al., 2008). In einer Meta-Analyse werden die Ergebnisse der jeweiligen Assoziationsstatistiken als P-Werte miteinander kombiniert. Die Kombination von P-Werten hat dieselbe Power genetische Varianten zu identifizieren, wie eine gebündelte Assoziationsanalyse aller Personen auf Genotypebene (Lin und Zeng, 2010).

Meta-Analysen wurden in dieser Studie mit METAL²³ (Willer et al., 2010) durchgeführt. METAL kombiniert P-Werte, gewichtet nach der Quadratwurzel der Stichprobengröße (Stouffer et al., 1949), über eine Umwandlung in Z-Scores anhand der jeweiligen Effektrichtung. Da GWAS oft einen starken Überschuss an Kontrollen haben, wird bei ungleichen Fall-Kontroll-Anzahlen die Verwendung einer effektiven Stichprobengröße empfohlen:

$$N_{\text{eff}} = 4 / (1 / N_{\text{Fälle}} + 1 / N_{\text{Kontrollen}})$$

N_{eff} ist dabei eine lineare Funktion wenn $N_{\text{Fälle}} = N_{\text{Kontrollen}}$. Mit steigender Differenz zwischen $N_{\text{Fälle}}$ und $N_{\text{Kontrollen}}$, tendiert N_{eff} gegen die vierfache Anzahl der kleinsten Gruppe.

2.13 Visualisierungen

Für die Visualisierung von Ergebnissen wurden folgende Anwendungen verwendet:

- Gnuplot²⁴: MDS-, QQ- und Chromosomen-Plots der GWAS; sämtliche Abbildungen der Kopplungsstudie
- Haploview²⁵ (Barrett et al., 2005): regionale LD-Muster
- LocusZoom²⁶ (Pruim et al., 2010): regionale Assoziations-Plots

2.14 Bestimmung von Kandidatengen

Über menschliche Krankheiten stehen zahlreiche veröffentlichte Forschungsarbeiten zur Verfügung. Literaturdatenbanken wie PubMed und OMIM (Online Mendelian Inheritance in Man)

²³ METAL - Meta Analysis Helper, online: <http://www.sph.umich.edu/csg/abecasis/metal/index.html>

²⁴ Gnuplot, online: <http://www.gnuplot.info>

²⁵ Haploview, online: <http://www.broad.mit.edu/mpg/haploview>

²⁶ LocusZoom, online: <http://www.sph.umich.edu/csg/abecasis/metal/index.html>

(Hamosh et al., 2005) von NCBI²⁷ (Wheeler et al., 2007) bieten eine Informationsmöglichkeit über bisher erforschte Aspekte einer Krankheit oder einzelner Gene. Für die direkte Suche nach Kandidatengenen wurde in dieser Studie das Webtool GeneDistiller 2 (Seelow et al., 2008) verwendet. Das Webtool verbindet alle gängigen Datenbanken, erlaubt Volltextsuche bezogen auf die eingegebenen genetischen Regionen und gibt bekannte Gen-Phänotyp Verbindungen an. Dafür werden auch Informationen von Mausmodellen aus der MGD-Datenbank²⁸ (Blake et al., 1997; Eppig et al., 2012) verwendet.

²⁷ NCBI, National Center for Biotechnology Information, online: <http://www.ncbi.nlm.nih.gov>

²⁸ MGD, Mouse Genome Database, online: <http://www.informatics.jax.org>

3 ERGEBNISSE

3.1 Genomweite Kopplungsstudie

Eine genomweite Kopplungsstudie in 379 IGE-Multiplexfamilien wurde mit dem Ziel durchgeführt Loci zu identifizieren die zur Ausprägung von familiären IGE Syndromen prädisponieren. Um das vollständige Spektrum der familiären IGE-Syndrome berücksichtigen zu können, wurde ein generelles Betroffenheitsmodell gewählt. In diesem Modell wurden Familienmitglieder mit jeglichem IGE-Syndrom als „betroffen“ klassifiziert. Familienmitglieder mit anderen Epilepsie- oder Anfallsformen, auffälligem EEG ohne Anfälle oder fehlenden klinischen Informationen wurden als „unbekannt“ kodiert. Alle verbliebenen Familienmitglieder wurden als „unbetroffen“ klassifiziert (Tab. 3-1, Tab. 3-2). Um nach erkrankungsassoziierten Loci differenzieren zu können, die zum Entstehungsrisiko von Absencen oder Myoklonien beitragen, wurden weitere Kopplungsanalysen in zwei Subgruppen durchgeführt. Die Zuordnung zu beiden Subgruppen erfolgte durch das Vorkommen von mindestens einem Familienmitglied mit jeweils IAE bzw. JME. Die beiden Familien-Subgruppen schlossen ein: 1.) 235 IAE-Multiplexfamilien mit 567 betroffenen Familienmitgliedern, davon 87,5% mit einer idiopathischen Absence-Epilepsie (IAE, CAE oder JAE); 2.) 118 JME-Multiplexfamilien mit 289 betroffenen Familienmitgliedern, davon 74,4% mit JME. Für die Kopplungsanalysen in den Subgruppen wurde ein striktes Betroffenheitsmodell ausgewählt, bei dem ausschließlich Familienmitglieder mit IAE oder JME als „betroffen“ klassifiziert wurden. Alle anderen IGEs (unklassifizierte IGE, EGTCS) wurden als „unbekannt“ kodiert (Tab. 3-1, Tab. 3-2).

Die Power der genomweiten Kopplungsanalysen ist in Tabelle 3-3 wiedergegeben. Dabei hatte die Kopplungsanalyse mit allen 379 IGE-Multiplexfamilien eine Power von ~100%, um einen genomweit signifikanten HLOD-Wert = 3.5 zu erreichen, sofern 30% der Familien Kopplung zu demselben Locus aufwiesen ($\alpha = 0,3$) wenn ein rezessives Vererbungsmodell angenommen wurde.

Familien Gruppe	Betroffenheitsmodell	Anzahlen			Betroffenheitsstatus		
		Familien	Individuen	typisiert	betroffen	unbetroffen	unbekannt
IGE	GM	379	1920	1728	982	856	82
IAE	SM	235	1213	1095	567	539	107
JME	SM	118	624	560	289	264	71

Tabelle 3-1 Analysegruppen der Kopplungsstudie

Legende -> GM / SM: generelles bzw. striktes Betroffenheitsmodell.

Familien Gruppe	IAE o. JME	IAE u. JME	JME	EGTCS	IGE unklass.	gsw-EEG o. E.	FS alleine	andere Epilepsien
IGE	504	60	198	205	15	21	18	24
IAE	442	54	71	60	7	18	9	3
JME	74	54	161	37	5	9	7	3

Tabelle 3-2 Klinische Charakterisierung der betroffenen Familienmitglieder

Legende -> IAE o. JME: IAE-Patienten ohne JME; IAE u. JME: Patienten mit IAE und JME; IGE unklass.: unklassifizierte IGE; gsw-EEG o. E.: generalisierte „spike-wave“-Komplexe im EEG ohne Epilepsiediagnose. Die klinischen Phänotypen, die als „betroffen“ kodiert wurden, sind dunkel hervorgehoben.

FAM Gruppe	FAM Anz.	Betroffenheitsmodell	Vererbungsmodell	Phenokopierate	Allelfreq. Risikoallel	Heterog. α	HLOD Power [%]		
							2,0	3,0	3,6
IGE	379	GM	AD70	0,05%	0,01	30%	96	86	78
IGE	379	GM	AR70	0,05%	0,10	30%	100	100	99
IAE	235	SM	AD50	0,05%	0,01	30%	17	4	0
IAE	235	SM	AR50	0,05%	0,10	30%	58	33	21
JME	118	SM	AD50	0,05%	0,01	30%	29	11	6
JME	118	SM	AR50	0,05%	0,10	30%	63	40	28

Tabelle 3-3 Power der parametrischen Kopplungsanalyse

Dargestellt ist die Power der Kopplungsanalyse um HLOD-Werte von 2,0; 3,0; 3,6 zu erreichen. Legende -> GM / SM: generelles bzw. striktes Betroffenheitsmodell; AD70: autosomal dominantes Vererbungsmodell mit 70% Penetranz; AR70: autosomal rezessives Vererbungsmodell mit 70% Penetranz; AD50: autosomal dominantes Vererbungsmodell mit 50% Penetranz; AR50: autosomal rezessives Vererbungsmodell mit 50% Penetranz.

Basierend auf den berechneten empirischen genomweiten Typ-I-Fehlerraten des IGE-Datensatzes wurden folgende Schwellen für genomweite Signifikanz bestimmt: 1.) Nichtparametrische Kopplungsanalysen: suggestive Kopplung bei $LOD_{NPL} > 1,80$; signifikante Kopplung bei $LOD_{NPL} > 3,15$. 2.) Parametrische Kopplungsanalysen: suggestive Kopplung bei $HLOD > 2,13$; signifikante Kopplung bei $HLOD > 3,49$.

3.1.1 Genomweite Kopplungsanalyse in 379 IGE-Multiplexfamilien

Das Ergebnis der genomweiten, nichtparametrischen Kopplungsanalyse in allen 379 IGE-Multiplexfamilien ist in Abb. 3-1A aufgeführt. Es wurden keine signifikanten Kopplungen ermittelt. Suggestive Kopplungshinweise fanden sich für vier Regionen: 3p14.2 ($LOD_{NPL} = 2,96$ bei rs624755; chromosomale Position chr3:61733962, NCBI 37.3), 5q34 ($LOD_{NPL} = 1,95$ bei rs1432881; chr5:166932520), 13q12.12 ($LOD_{NPL} = 2,42$ bei rs1008812; chr13:23966145) und 19q13.42 ($LOD_{NPL} = 2,86$ bei rs9788; chr19:53719250) (Tabelle 3-4).

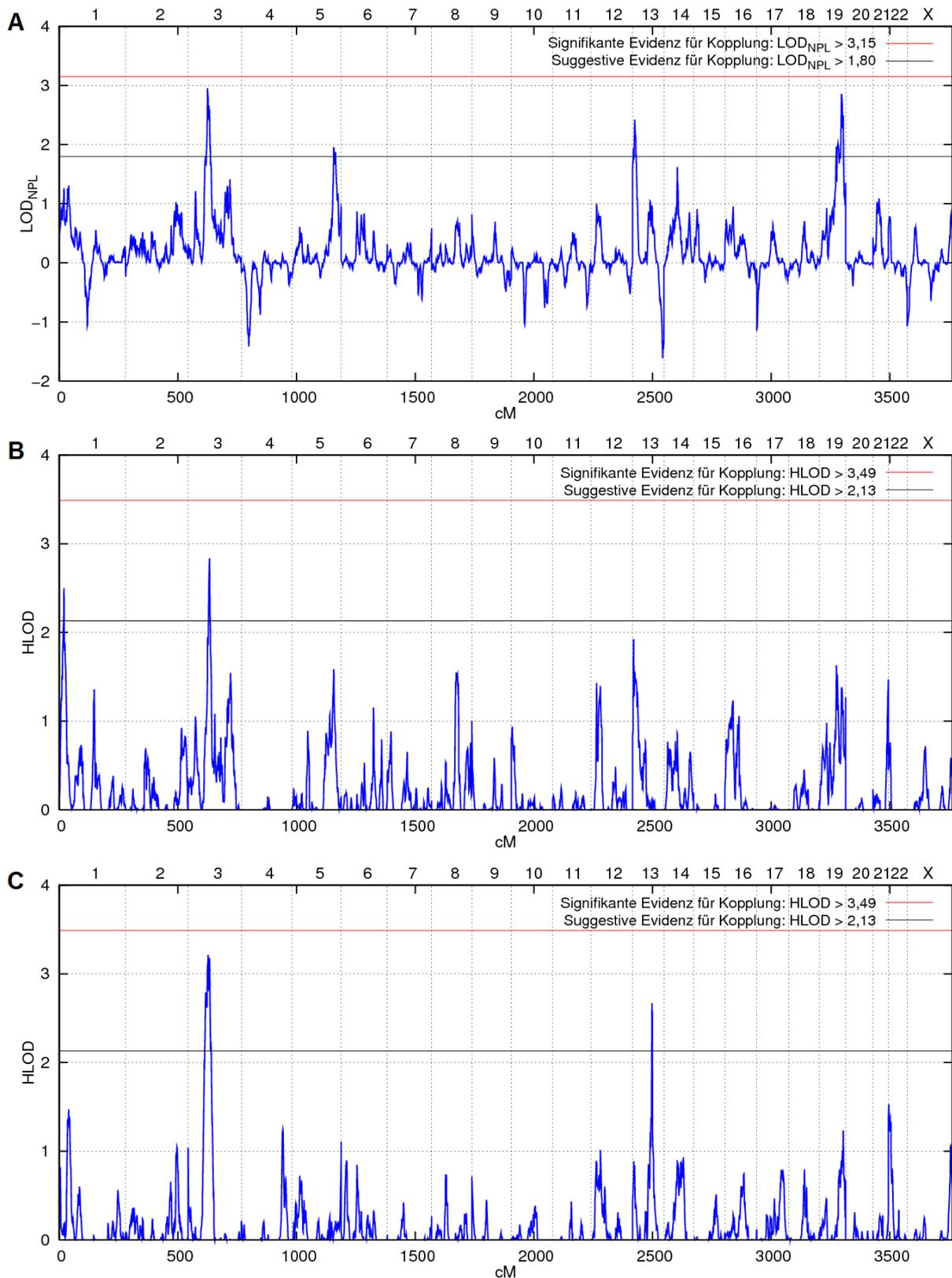


Abb. 3-1 Genomweite Kopplungsanalyse in 379 IGE-Multiplexfamilien

Dargestellt sind die genomweiten LOD-Werte als Funktion der genetischen Karte: (A) nichtparametrische Kopplungsanalyse; (B) parametrische Kopplungsanalyse unter der Annahme eines dominanten Vererbungsmodells mit 70% Penetranz; (C) parametrische Kopplungsanalyse unter der Annahme eines rezessiven Vererbungsmodells mit 70% Penetranz. Die Schwellen für suggestive und signifikante Kopplung sind als schwarze bzw. rote horizontale Linien gezeichnet.

Die Ergebnisse der genomweiten parametrischen Kopplungsanalysen sind in den Abb. 3-1B (dominantes Vererbungsmodell) und Abb. 3-1C (rezessives Vererbungsmodell) aufgeführt. Es wurden keine signifikanten Kopplungen ermittelt. In Übereinstimmung mit der nichtparametrischen Kopplungsanalyse ergaben die HLOD-Analysen suggestive Kopplungshinweise mit beiden Vererbungsmodellen bei 3p14 (dominantes Vererbungsmodell: HLOD = 2,84 bei rs782728; chr3:66326302; $\alpha = 0,20$ // rezessives Vererbungsmodell: HLOD = 3,21 bei rs1374679; chr3:63075267; $\alpha = 0,13$). Suggestive Kopplungshinweise ergaben sich für zwei weitere Regionen: 1p36.22 (dominantes Vererbungsmodell: HLOD = 2,50 bei rs12136213; chr1:10046460; $\alpha = 0,17$) und 13q31.3 (rezessives Vererbungsmodell: HLOD = 2,67 bei D13S1230; chr13:90036642; $\alpha = 0,11$) (Tabelle 3-5).

3.1.2 Genomweite Kopplungsanalyse in 235 IAE-Multiplexfamilien

Das Ergebnis der genomweiten, nichtparametrischen Kopplungsanalyse in der Subgruppe mit 235 IAE-Multiplexfamilien ist in Abb. 3-2A aufgeführt. Ein suggestiver Kopplungshinweis ergab sich für die Region auf 5q34 ($LOD_{NPL} = 2,31$ bei rs244903; chr5:167913510) (Tabelle 3-4).

Die Ergebnisse der genomweiten, parametrischen Kopplungsanalysen sind in den Abb. 3-2B (dominantes Vererbungsmodell) und Abb. 3-2C (rezessives Vererbungsmodell) aufgeführt. In Übereinstimmung mit der nichtparametrischen Kopplungsanalyse ergab die HLOD-Analyse mit dem dominanten Vererbungsmodell ein suggestives Kopplungshinweis bei 5q34 (HLOD = 3,23 bei rs357608; chr5:150840380; $\alpha = 0,31$). In der HLOD-Analyse mit dem rezessiven Vererbungsmodell fand sich ein signifikanter Kopplungshinweis bei 13q31.3 (HLOD = 5,02 bei rs1332470; chr13:91417190; $\alpha = 0,22$) (Tabelle 3-5).

3.1.3 Genomweite Kopplungsanalyse in 118 JME-Multiplexfamilien

Das Ergebnis der genomweiten, nichtparametrischen Kopplungsanalyse in der Subgruppe mit 118 JME-Multiplexfamilien ist in Abb. 3-3A aufgeführt. Ein signifikanter Kopplungshinweis ergab sich für eine Region bei 2q34 ($LOD_{NPL} = 3,43$ bei D2S143; chr2:214973108). Ein suggestiver Kopplungshinweis fand sich in der 5q34 Region ($LOD_{NPL} = 2,62$ bei rs1025482; chr5:166893257) (Tabelle 3-4).

Die Ergebnisse der genomweiten parametrischen Kopplungsanalysen sind in den Abb. 3-3B (dominantes Vererbungsmodell) und Abb. 3-3C (rezessives Vererbungsmodell) aufgeführt.

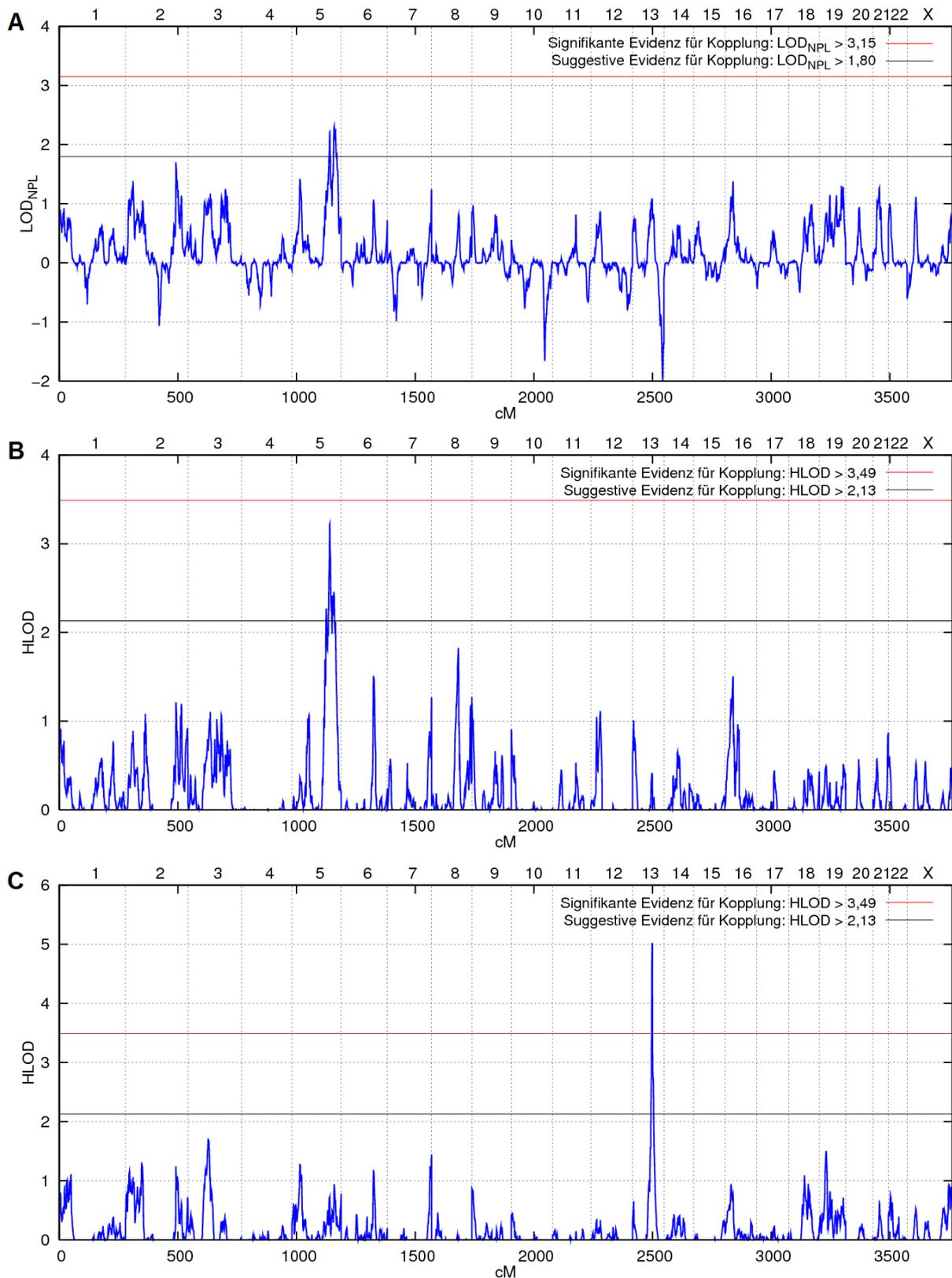


Abb. 3-2 Genomweite Kopplungsanalyse in 235 IAE-Multiplexfamilien

Dargestellt sind die genomweiten LOD-Werte als Funktion der genetischen Karte: (A) nichtparametrische Kopplungsanalyse; (B) parametrische Kopplungsanalyse unter der Annahme eines dominanten Vererbungsmodells mit 70% Penetranz; (C) parametrische Kopplungsanalyse unter der Annahme eines rezessiven Vererbungsmodells mit 70% Penetranz. Die Schwellen für suggestive und signifikante Kopplung sind als schwarze bzw. rote horizontale Linien gezeichnet.

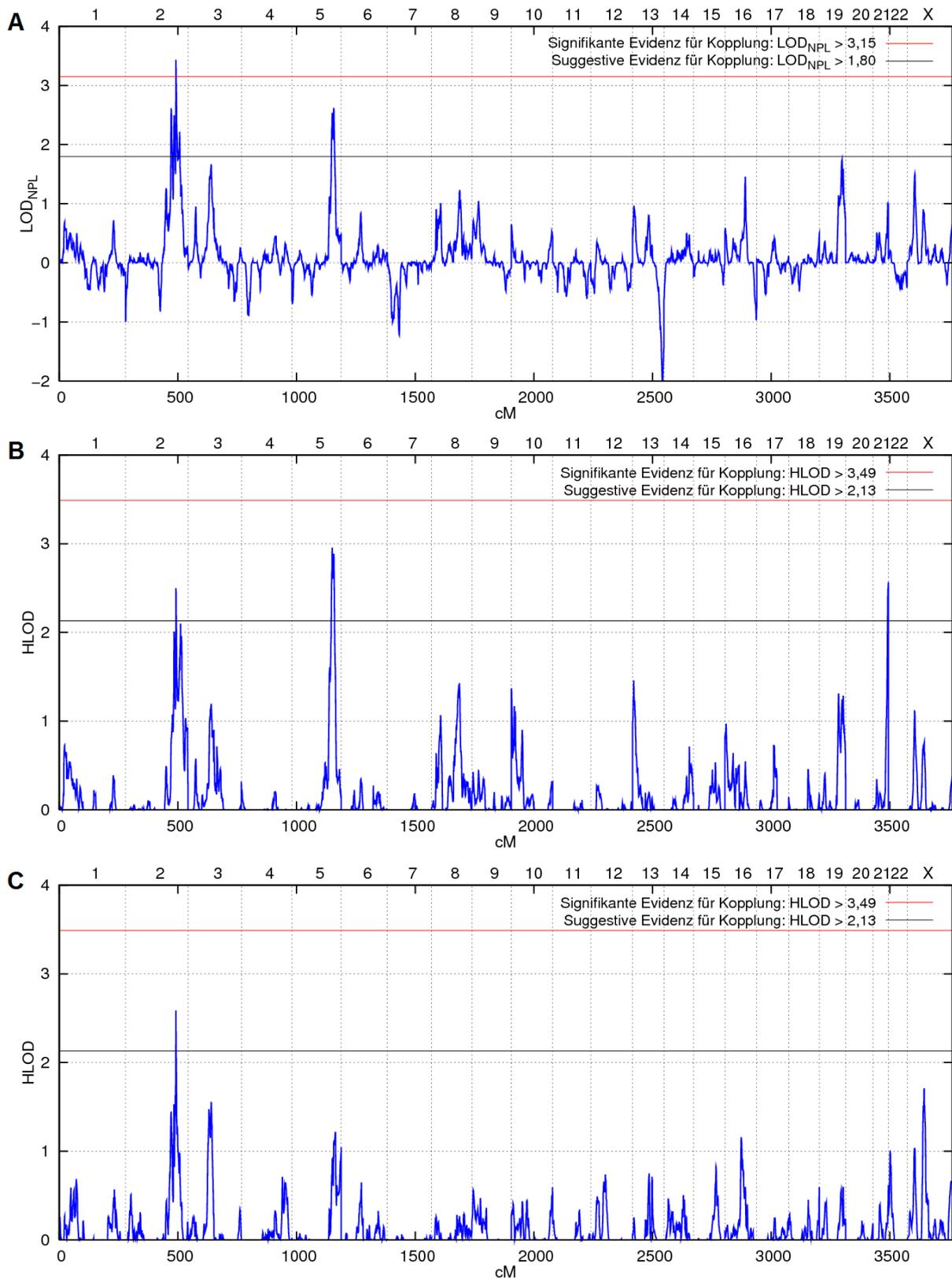


Abb. 3-3 Genomweite Kopplungsanalyse in 118 JME-Multiplexfamilien

Dargestellt sind die genomweiten LOD-Werte als Funktion der genetischen Karte: (A) nichtparametrische Kopplungsanalyse; (B) parametrische Kopplungsanalyse unter der Annahme eines dominanten Vererbungsmodells mit 70% Penetranz; (C) parametrische Kopplungsanalyse unter der Annahme eines rezessiven Vererbungsmodells mit 70% Penetranz. Die Schwellen für suggestive und signifikante Kopplung sind als schwarze bzw. rote horizontale Linien gezeichnet.

Beide Regionen mit Kopplungshinweisen aus der nichtparametrischen Analyse (2q34; 5q34) wurden mit suggestiven Kopplungshinweisen durch die HLOD-Analyse unterstützt. Die HLOD-Analyse ergab suggestive Kopplungshinweise mit beiden Vererbungsmodellen bei 2p34 (dominantes Vererbungsmodell: HLOD = 2,50 bei D2S143; chr2:214973108; $\alpha = 0,39$ // rezessives Vererbungsmodell: HLOD = 2,59 bei D2S143; chr2:214973108; $\alpha = 0,25$). In der HLOD-Analyse mit dem dominanten Vererbungsmodell fanden sich zwei weitere suggestive Kopplungshinweise: 5q34 (HLOD = 2,96 bei rs2069347; chr5:162867195; $\alpha = 0,40$) und 21q22.3 (HLOD = 2,57 bei rs2839377; chr21:48077812; $\alpha = 0,39$) (Tabelle 3-5).

3.1.4 Zusammenfassung der Kopplungsergebnisse

Insgesamt konnten acht Kopplungsregionen festgestellt und anhand der suggestiven Kopplungssignale eingegrenzt werden (Tabellen 3-4 und 3-5).

FAM Gruppe	Kopplungsregion, $LOD_{NPL} > 1,80$			max. LOD_{NPL}	bei Marker	Position NCBI 37.3
	Chromosom	Start	Ende			
IGE	3p14.3-p13	54595678	70445977	2,96	rs624755	61733962
IGE	5q34-q35.1	166153356	168828290	1,95	rs1432881	166932520
IGE	13q12.11-q12.13	21295194	27652291	2,42	rs1008812	23966145
IGE	19q13.32-q13.42	47529740	55428459	2,86	rs9788	53719250
IAE	5q32-q35.1	149609393	172580866	2,31	rs244903	167913510
JME	2q32.1-q36.1	188115398	223823132	3,43	D2S143	214973108
JME	5q33.3-q35.1	158556738	169032520	2,62	rs1025482	166893257

Tabelle 3-4 Suggestive und signifikante LOD_{NPL} -Scores

Dargestellt sind die Regionen in denen suggestive Kopplungssignale erreicht werden. Der signifikante Kopplungshinweis bei 2p34 für JME ist fett hervorgehoben. Alle physikalischen Positionen sind nach der NCBI Version 37.3 des menschlichen Genoms.

FAM Gruppe	Vererb. Modell	Kopplungsregion, $HLOD > 2,13$			max. HLOD	α	bei Marker	Position NCBI 37.3
		Chromosom	Start	Ende				
IGE	AD70	1p36.22	9370911	10767902	2,50	0,17	rs12136213	10046460
IGE	AD70	3p14.2-p13	63563882	70299390	2,84	0,20	rs782728	66326302
IGE	AR70	3p21.1-p13	52833805	72679023	3,21	0,13	rs1374679	63075267
IGE	AR70	13q31.2-q31.3	88297101	92549631	2,67	0,11	D13S1230	90036642
IAE	AD50	5q31.3-q35.1	139646776	168828290	3,23	0,31	rs357608	150840380
IAE	AR50	13q31.1-q32.1	82259163	95654167	5,02	0,22	rs1332470	91417190
JME	AD50	2q34-q35	213496624	216195171	2,50	0,39	D2S143	214973108
JME	AD50	5q33.3-q35.1	158556738	168828290	2,96	0,40	rs2069347	162867195
JME	AD50	21q22.3	46148874	48077812	2,57	0,39	rs2839377	48077812
JME	AR50	2q34-q35	213496624	216195171	2,59	0,25	D2S143	214973108

Tabelle 3-5 Suggestive und signifikante HLOD-Scores

Dargestellt sind die Regionen in denen suggestive Kopplungssignale erreicht werden. Der signifikante Kopplungshinweis bei 13q31.3 für IGE ist fett hervorgehoben. Alle physikalischen Positionen sind nach der NCBI Version 37.3 des menschlichen Genoms.

Aus den genomweiten Kopplungsanalysen mit allen 379 IGE-Multiplexfamilien ergaben sich suggestive Kopplungshinweise für sechs Regionen (1p36.22, 3p21.1-p13, 5q31.3-q35.1, 13q12.11-q12.13, 13q31.1-q32.1, 19q13.32-q13.42).

Für eine dieser Regionen im Bereich der Zytobande 13q31.1-q32.1 (chr13:82259163-95654167) wurde in der Subgruppe mit 235 IAE-Multiplexfamilien signifikante Evidenz für Kopplung in der parametrischen Analyse erreicht (max. HLOD = 5,02 bei rs1332470; chr13:91417190; $\alpha = 0,22$). Der signifikante Kopplungshinweis wird von der nichtparametrischen Analyse nicht unterstützt (max. $LOD_{NLP} = 1,07$ bei rs1332470). Eine mögliche Erklärung dafür ist, dass unbetroffene Personen einen signifikanten Beitrag zu dem Kopplungsergebnis haben. Post-Hoc-Analysen ergaben keinen Hinweis, dass der signifikante parametrische Kopplungsbefund ein Artefakt der Familienstruktur oder des Markerdatensatzes ist: 1.) Analyse mit Kernfamilien (Eltern und Kinder): HLOD = 4,27; 2.) Entfernung von Markern mit $r^2 > 0,1$: keine Veränderung, HLOD = 5,02. Die Kopplungsregion weist 22 Gene auf (ohne Pseudogene). Das aussichtsreichste Kandidatengen ist *GPC5* (glypican 5), dessen Genprodukt als Modulator von Zell-Zell Wechselwirkungen in der Neurogenese involviert ist (Luxardi et al., 2007). Ein weiterer Kandidatengen ist *SOX21* (SRY-related HMG-box gene 21) dessen Genprodukt, als neuronal exprimierter Transkriptionsfaktor, essentiell für die neuronale Differenzierung ist (Sandberg et al., 2005).

Die nichtparametrische Kopplungsanalyse in der Subgruppe mit 118 JME-Multiplexfamilien ergab eine signifikante Kopplung in der chromosomalen Region 2q32.1-q36.1 (chr2:188115398-223823132; max. $LOD_{NPL} = 3,43$ bei D2S143; chr2:214973108). Die parametrischen Analysen unterstützen den signifikanten Kopplungshinweis (HLOD > 2,5 in beiden Vererbungsmodellen bei D2S143). Die Kopplungsregion weist 256 Gene auf (ohne Pseudogene). Davon sind 204 proteinkodierend und 30 weisen einen neurologischen Phänotyp in der MGD-Datenbank auf. Aus diesen ist *ERBB4* (v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)) ein hochrangiges Kandidatengen für IGE. *ERBB4* ist an der Synaptogenese der GABA-Neuronen beteiligt (Yau et al., 2003; Ting et al., 2011) und führt als „Knockout“ in der Maus zu spontanen Krampfanfällen (Tan et al., 2011). Weitere interessante Kandidatengene sind *CREB1* (cAMP responsive element binding protein 1), *SLC4A3* (solute carrier family 4, anion exchanger, member 3) und *CDK5R2* (cyclin-dependent kinase 5, regulatory subunit 2 (p39)).

Sowohl in den nichtparametrischen, als auch parametrischen Kopplungsanalysen, ergaben sich suggestive Kopplungshinweise in der Region auf 5q31.3-q35.1 besonders akzentuiert für ein dominantes Vererbungsmodell (überspannter Bereich: chr5:139646776-172580866). In diesem Bereich befindet sich ein Gen-Cluster von GABA_A-Rezeptor-Untereinheiten

(GABRB2, GABRA6, GABRA1, GABRG2). Mutationen in den Genen *GABRA1* und *GABRG2* sind bei mehreren Formen von Epilepsie beschrieben worden (CAE, JME, GESF+, SMEI; Tabelle 1-3).

3.2 Genomweite Assoziationsstudie

Zur Identifikation von genetischen Varianten, die häufig in der Allgemeinbevölkerung vorkommen und zur Ausprägung der IGEs prädisponieren, wurde eine GWAS mit 1527 IGE-Einzelfällen und 2461 Populationskontrollen durchgeführt. Als Teil einer mehrstufigen Assoziationsstudie wurden die „Top-Hits“ der genomweiten Suche in zwei unabhängige IGE-Kollektiven gezielt überprüft: 1.) 572 IGE-Trios und 2.) 727 IGE-Einzelfällen und 668 Populationskontrollen. In einer Meta-Analyse (Kap. 2.12) wurden abschließend die Ergebnisse der drei Kollektive kombiniert.

Zur Identifikation von genetischen Varianten, die differentiell zum Entstehungsrisiko von entweder IAE oder JME beitragen, wurden GWAS in diesen zwei Subgruppen durchgeführt. Die IAE-Subgruppe setzte sich zusammen aus: 1.) GW Scan: 702 IAE-Einzenfällen und 2461 Populationskontrollen; 2.) Replikation: 323 IAE-Trios, 312 IAE-Einzenfällen und 668 Populationskontrollen. Die JME-Subgruppe bestand aus: 1.) GW Scan: 586 JME-Einzenfällen und 2461 Populationskontrollen; 2.) Replikation: 158 JME-Trios, 323 JME-Einzenfällen und 668 Populationskontrollen. Ein Überblick der Phänotypverteilung und Ethnizitäten der gesammelten IGE-Kollektive ist in Kap. 2.1 gegeben.

GWAS-Qualitätskontrollen

Die aufgeführten Studienteilnehmer der GWAS sind das Resultat einer vorangehenden Qualitätskontrolle. Von ursprünglich 1595 IGE-Einzelfällen und 2518 Populationskontrollen, wurden 125 Personen aufgrund folgender Kriterien ausgeschlossen: 1.) nicht übereinstimmende Geschlechtsinformation ($n = 1$); 2.) auffällige \emptyset -Heterozygotieraten von autosomalen SNPs, \emptyset -HET $> 29,5\%$ ($n = 1$); 3.) duplizierte oder verwandte Personen mit einem IBS $> 1,55$ ($n = 12$); 4.) genetische Ausreißer identifiziert durch einen direkten Vergleich der paarweisen IBS-Werten (Personen mit $> 25\%$ ihrer IBS-Abstände mehr als vier Standardabweichungen entfernt vom Mittelwert aller IBS-Abstände der engsten Nachbarn; $n = 82$) gefolgt von einer MDS der genomweiten IBS-Matrix (verbleibende Populationsstratifizierung; $n = 29$, Abb. 3-4). Personen mit einer Array CR $< 95\%$ wurden im Vorfeld während des „Genotype Calling“ ausgeschlossen (Kap. 2.5.1).

Der genomweite Datensatz von 906.600 SNPs wurde der Berechnung der Assoziationsanalyse von potentiellen Genotypisierungsfehlern bereinigt. SNPs wurden nach folgenden Qualitätskriterien ausgeschlossen: 1.) CR $< 95\%$ in den Fällen sowie in den Kontrollen; 2.) CR-Differenz zwischen den Fällen und den Kontrollen $> 1\%$; 3.) MAF $< 2\%$ in den Fällen sowie in den Kontrollen; 4.) Abweichung vom HWE mit $P < 10^{-4}$ in den Kontrollen;

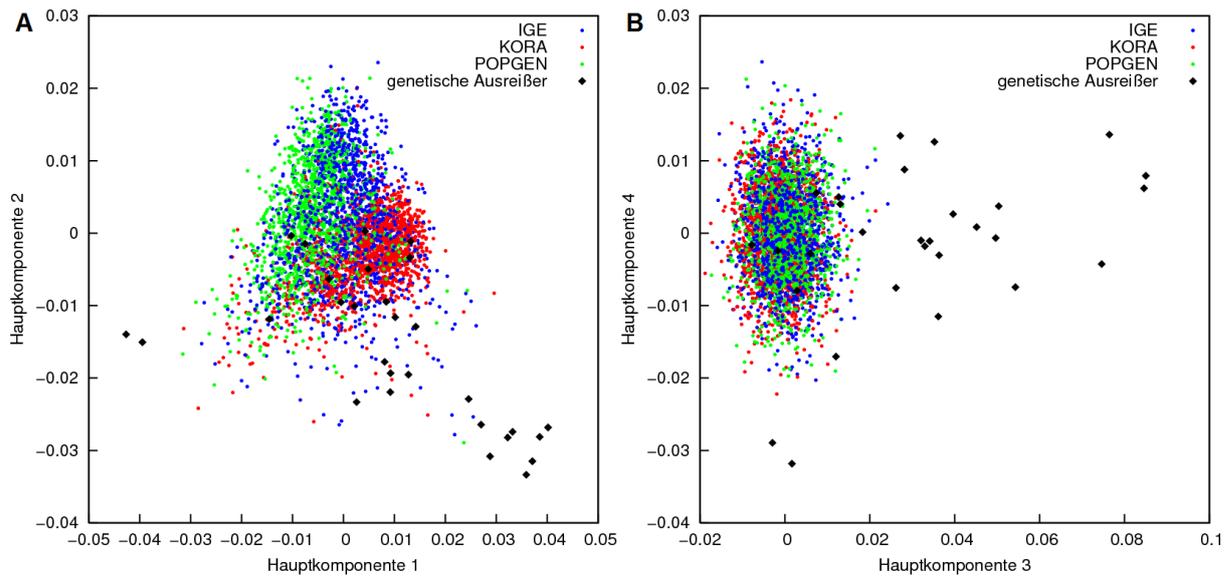


Abb. 3-4 MDS von 1549 IGE-Einzenfällen und 2468 Kontrollen

Dargestellt ist die Populationsstruktur abgebildet durch (A) die ersten zwei Hauptkomponenten und (B) Hauptkomponenten 3 und 4 einer MDS der IBS-Matrix von 1549 IGE-Einzenfällen und 2468 Kontrollen. Die MDS basiert auf 23.973 unkorrelierte ($r^2 < 0,1$) und qualitätsgefilterte SNPs (CR > 99%, MAF > 15%, HWE $P > 10^{-3}$). Legende -> blaue Punkte: 1527 IGE-Fälle; rote Punkte: 1381 KORA Kontrollen; grüne Punkte: 1080 POPGEN Kontrollen; schwarze Rauten: 29 genetische Ausreißer.

5.) Abweichungen der Allelfrequenzen ($P < 10^{-4}$) zwischen drei Teilgruppen der IGE-Erkrankten und zwischen zwei Teilgruppen der Kontrollen. Zusätzlich wurde die Zuverlässigkeit der „Genotyp-Diskriminierung“ von allen GWAS-„Top-Hits“ mit $P < 10^{-5}$ überprüft und SNPs mit fehlerhaften Genotypisierungen ausgeschlossen.

Nach der Anwendung aller Qualitätskontrollen wurden in der GWAS im IGE-Gesamtkollektiv 629.549, in der IAE-Subgruppe 613.727 und in der JME-Subgruppe 616.386 SNPs eingeschlossen.

Signifikanzschwellen

Als Schwelle für eine genomweit signifikante Assoziation wurde $P < 5 \times 10^{-8}$ gewählt (nach GC-Korrektur, Kap. 2.11.4). $P = 5 \times 10^{-8}$ ist als Signifikanzniveau für GWAS in der wissenschaftlichen Literatur etabliert und entspricht nach der Bonferroni-Methode einem Signifikanzniveau von 0,05 bei 1.000.000 unabhängigen Tests. Von 629.549 SNPs, die in dieser Studie getestet worden sind, befinden sich 298.883 im starken LD zueinander ($r^2 > 0.8$). Aus 330.666 SNPs mit $r^2 < 0.8$ und drei getesteten IGE-Phänotypen (IGE, IAE und JME) resultieren 991.998 relativ unabhängige Tests. Als Schwelle für eine genomweit suggestive Assoziation wurde $P < 10^{-6}$ gewählt (1 / 1.000.000; Simino et al., 2011; Wright et al., 2011). Für die Replikationsanalyse wurden SNPs mit einem Assoziationshinweis von $P_{GC} \leq 10^{-5}$ ausgewählt.

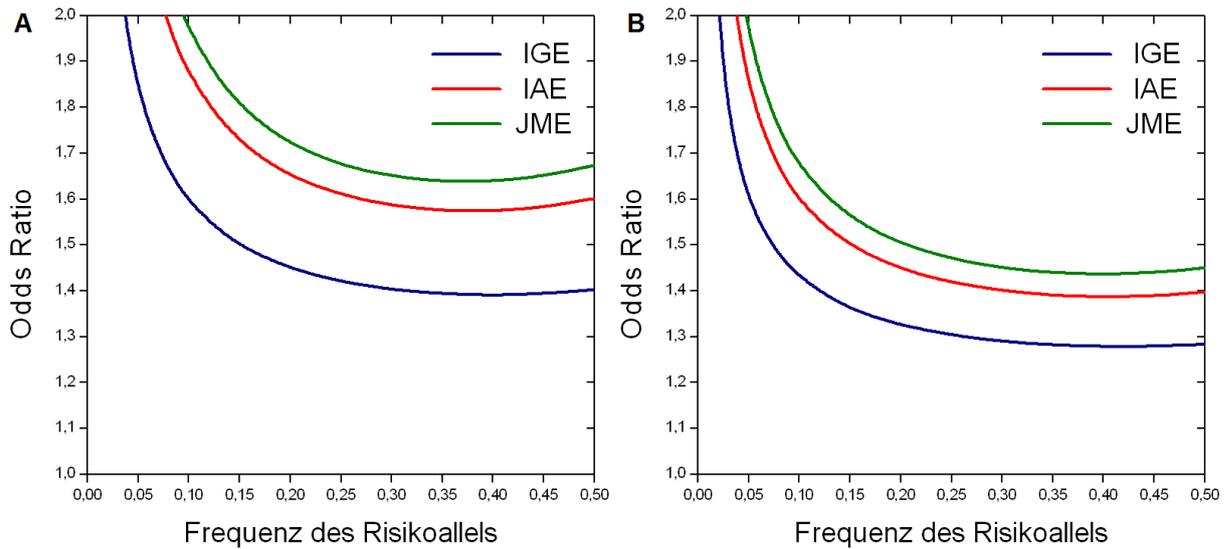


Abb. 3-5 Detektierbare ORs mit einer Power von 80% bei $P = 5 \times 10^{-8}$

Dargestellt sind die detektierbaren ORs als Funktion der Allelfrequenz des Risikoallels für: (A) GWAS für IGE, IAE und JME; (B) Meta-Analyse der GWAS und der Replikationsanalyse für IGE, IAE und JME. Die Power-Analysen wurde unter folgenden Annahmen durchgeführt: Prävalenz von IGE = 0,3%; IAE = JME = 0,1%; Signifikanzschwelle $P = 5 \times 10^{-8}$; additives genetisches Modell; max. LD zwischen Risikoallel und genotypisierten Marker.

Power-Analysen

Die Meta-Analyse der drei IGE-Kollektive (GWAS und Replikation) erreicht eine Power von 80%, um Risikoallele mit einem $OR \geq 1,33$ bei einer Signifikanzschwelle von $P = 5 \times 10^{-8}$ zu identifizieren. Für Power-Analysen wurden folgende Annahmen getroffen: Prävalenz der Erkrankung = 0,3%, additives genetisches Modell, maximales LD zwischen Risikoallel und genotypisierten Markerallel und eine Frequenz des Risikoallels von 20%. Unter den gleichen Bedingungen, aber mit einer reduzierten Prävalenz von 0,1%, wurde in der IAE-Subgruppe eine Power von 80% für $OR \geq 1,45$ und in der JME-Subgruppe für $OR \geq 1,51$ erreicht. Ohne Meta-Analyse erreichte die GWAS eine Power von 80% bei einem $OR \geq 1,46$ für IGE, $OR \geq 1,66$ für IAE und $OR \geq 1,73$ für JME. Die detektierbaren ORs bei einer Power von 80% sind in den Abb. 3-5A (GWAS) und Abb. 3-5B (Meta-Analyse: GWAS und Replikationsanalyse) wiedergegeben.

3.2.1 GWAS im IGE-Gesamtkollektiv

Das Ergebnis der GWAS von 1527 IGE-Einzelfällen und 2461 Populationskontrollen ist in Abb. 3-6A dargestellt. Als Statistik wurde das additive Modell der logistischen Regression verwendet, mit dem Geschlechtsstatus und den ersten vier Hauptkomponenten der MDS als Kovarianten (Kap. 2.11.2). Anschließend wurden alle P-Werte aus der logistischen Regression (LR-P) nach der „Genomic Control“-Methode für potentielle Stratifikationseffekte korrigiert (LR- P_{GC}).

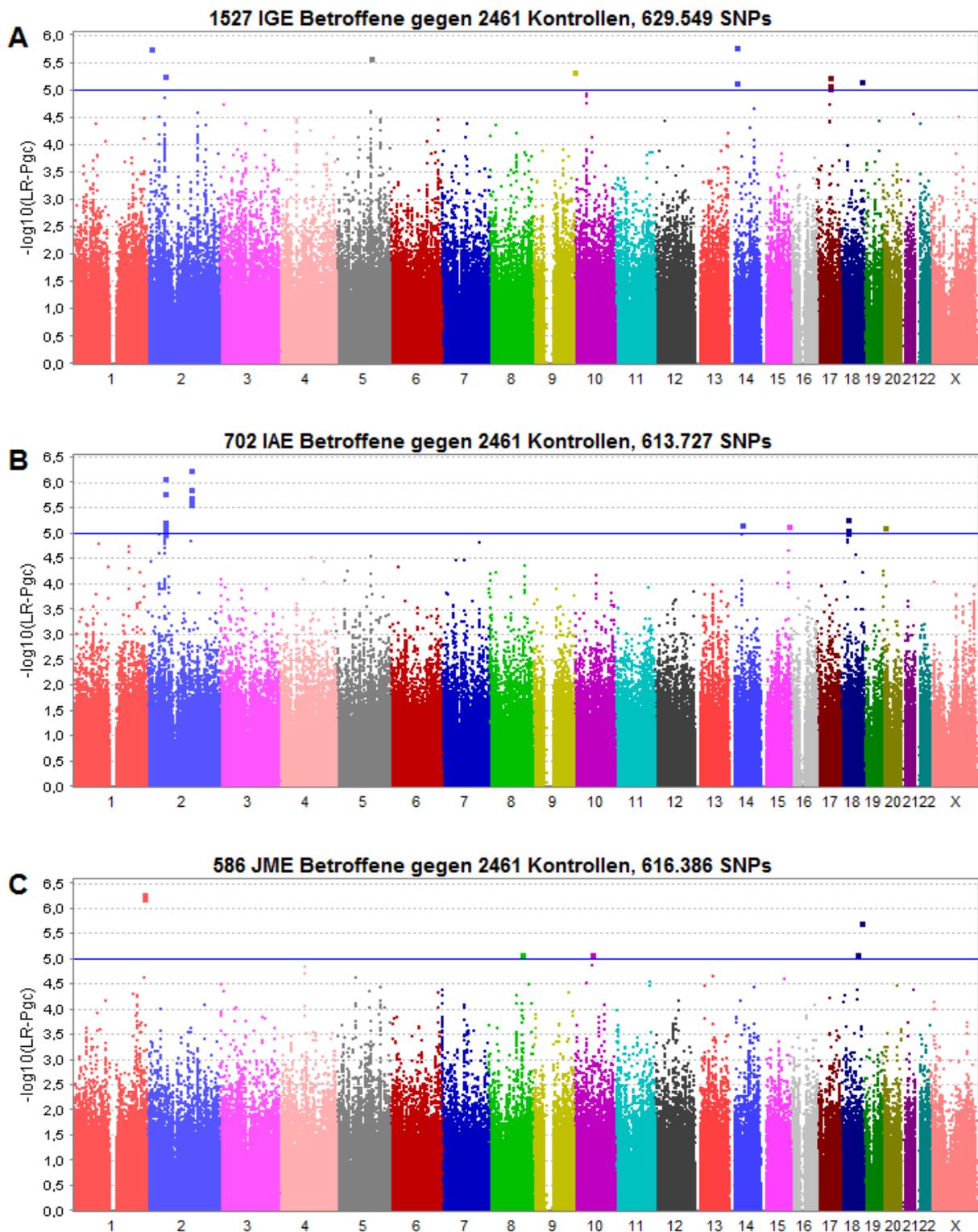


Abb. 3-6 Manhattan Plots der GWAS im IGE-Gesamtkollektiv und in Subgruppen

Dargestellt sind die Manhattan Plots der GWAS im IGE-Gesamtkollektiv (A), in der IAE-Subgruppe (B) und in der JME-Subgruppe (C). Gezeichnet sind $-\log_{10}(\text{LR-P}_{\text{GC}})$ -Werte der SNPs sortiert nach der physikalischen Position (IGE: $n = 629.549$; IAE $n = 613.727$; JME $n = 616.386$). Die blaue horizontale Linie markiert die Schwelle bei $\text{LR-P}_{\text{GC}} < 10^{-5}$ für eine Selektion in die Replikationsanalyse.

Aus der Abweichung der unkorrigierten LR-P-Werten von der erwarteten Verteilung resultierte ein λ_{GC} von 1,12 ($\lambda_{GC1000} = 1,06$). Dementsprechend wies der Quantile-Quantile- (QQ-) Plot auf einem inflationären Verhalten der Teststatistiken vor der GC-Korrektur (Abb. 3-7A).

Von den 629.549 analysierten SNPs wies keiner genomweit signifikante LR- P_{GC} -Werte auf. Insgesamt 10 SNPs aus sieben unabhängigen Regionen erreichten die Schwelle von LR- $P_{GC} \leq 10^{-5}$ für die Replikationsanalyse (Tabelle 3-6).

SNP	CHR	POS NCBI 37.3	Allele A1/A2	CR F/K	MAF F/K	HWE-P F/K	OR (95%-KI)	LR- P_{GC}
rs2268330	14	23795976	G/C	1,00/1,00	0,26/0,21	0,35/0,07	1,33 (1,19-1,49)	1,64E-06
rs388556	2	13392540	G/C	1,00/1,00	0,25/0,31	0,22/0,78	0,76 (0,69-0,85)	1,75E-06
rs289034	5	114122165	T/C	1,00/1,00	0,43/0,38	0,18/0,61	1,27 (1,16-1,40)	2,56E-06
rs1110146	9	140777370	C/A	1,00/1,00	0,27/0,33	1,00/0,96	0,78 (0,70-0,86)	4,47E-06
rs11885093	2	57941185	T/C	1,00/1,00	0,39/0,44	0,79/0,24	0,79 (0,72-0,87)	5,35E-06
rs11650615	17	46123698	G/C	1,00/1,00	0,25/0,30	0,95/0,67	0,77 (0,70-0,86)	5,77E-06
rs4892034	18	70399988	T/A	0,99/0,98	0,24/0,20	0,36/0,80	1,32 (1,18-1,47)	6,72E-06
rs4982753	14	23814569	T/C	0,99/1,00	0,28/0,24	0,61/0,15	1,29 (1,16-1,44)	7,41E-06
rs41336845	17	46215110	T/C	1,00/1,00	0,17/0,21	0,78/0,76	0,75 (0,66-0,84)	8,10E-06
rs16953235	17	46244948	G/A	1,00/1,00	0,17/0,21	0,85/0,71	0,75 (0,67-0,85)	8,96E-06

Tabelle 3-6 SNPs mit $P_{GC} \leq 10^{-5}$ in der logistischen Regression im IGE-Gesamtkollektiv

Legende -> CHR: Chromosom; POS NCBI 37.3; physikalische Position nach der NCBI Version 37.3; CR, MAF, HWE-P F/K: CR, MAF und HWE-P-Werte jeweils getrennt nach Fälle und Kontrollen; OR (95%-KI): Odds Ratio mit 95%-Konfidenzintervall bezogen auf Allel A1.

3.2.2 GWAS in der IAE-Subgruppe

Das Ergebnis der GWAS in der Subgruppe mit 702 IAE-Einzelfällen und 2461 Populationskontrollen ist in Abb. 3-6B dargestellt. Aus der Abweichung der unkorrigierten LR-P-Werten von der erwarteten Verteilung resultierte ein λ_{GC} von 1,07 ($\lambda_{GC1000} = 1,06$) (Abb. 3-7B). Von den 613.727 analysierten SNPs wies keiner genomweit signifikante LR- P_{GC} -Werte auf. Insgesamt 26 SNPs aus sechs unabhängigen Regionen erreichten die Schwelle von LR- $P_{GC} \leq 10^{-5}$ für die Replikationsanalyse (Tabelle 3-7).

3.2.3 GWAS in der JME-Subgruppe

Das Ergebnis der GWAS in der Subgruppe mit 586 JME-Einzelfällen und 2461 Populationskontrollen ist in Abb. 3-6C dargestellt. Aus der Abweichung der unkorrigierten LR-P-Werten von der erwarteten Verteilung resultierte ein λ_{GC} von 1,06 ($\lambda_{GC1000} = 1,06$) (Abb. 3-7C). Von den 616.386 analysierten SNPs wies keiner genomweit signifikante LR- P_{GC} -Werte auf. Insgesamt sieben SNPs aus fünf unabhängigen Regionen erreichten die Schwelle von LR- $P_{GC} \leq 10^{-5}$ für die Replikationsanalyse (Tabelle 3-8).

SNP	CHR	POS NCBI 37.3	Allele A1/A2	CR F/K	MAF F/K	HWE-P F/K	OR (95%-KI)	LR-P _{GC}
rs1035837	2	145379270	A/C	0,99/0,99	0,11/0,16	0,69/0,37	0,61 (0,50-0,73)	5,60E-07
rs2953442	2	58083420	C/T	1,00/1,00	0,37/0,30	0,26/0,57	1,39 (1,23-1,58)	8,20E-07
rs1427500	2	145388519	T/C	1,00/1,00	0,11/0,16	0,56/0,33	0,62 (0,52-0,75)	1,32E-06
rs1520965	2	58063748	C/A	1,00/1,00	0,37/0,30	0,29/0,56	1,38 (1,22-1,57)	1,55E-06
rs17741930	2	145362155	C/T	1,00/1,00	0,11/0,16	0,84/0,36	0,62 (0,52-0,75)	1,97E-06
rs17690892	2	145360235	G/A	1,00/1,00	0,11/0,16	0,84/0,44	0,62 (0,52-0,75)	2,04E-06
rs1346343	2	145364356	A/G	1,00/1,00	0,11/0,16	1,00/0,13	0,63 (0,52-0,76)	2,37E-06
rs10496964	2	145359909	T/C	1,00/1,00	0,11/0,16	0,84/0,64	0,63 (0,52-0,76)	2,55E-06
rs10853686	18	25870524	T/C	1,00/1,00	0,36/0,43	0,12/0,35	0,74 (0,66-0,84)	5,05E-06
rs2717024	2	58161397	C/G	1,00/1,00	0,47/0,41	0,05/0,13	1,34 (1,19-1,52)	5,98E-06
rs1447579	14	42125681	T/C	0,99/0,99	0,22/0,27	0,50/0,03	0,70 (0,61-0,82)	6,56E-06
rs12904369	15	97335674	C/A	1,00/1,00	0,30/0,23	0,72/0,57	1,38 (1,20-1,57)	7,21E-06
rs2717023	2	58161355	C/T	1,00/1,00	0,47/0,41	0,07/0,13	1,34 (1,18-1,52)	7,47E-06
rs1568253	2	58170004	T/C	1,00/1,00	0,47/0,40	0,06/0,16	1,34 (1,18-1,52)	7,52E-06
rs16993982	20	7744461	C/T	0,99/1,00	0,20/0,16	0,01/0,09	1,46 (1,24-1,71)	7,54E-06
rs11684106	2	58075701	C/A	1,00/1,00	0,35/0,29	0,41/0,22	1,36 (1,19-1,55)	7,96E-06
rs2717040	2	58169494	T/C	1,00/1,00	0,47/0,41	0,07/0,11	1,34 (1,18-1,51)	8,28E-06
rs2678897	2	58169418	G/A	1,00/1,00	0,47/0,40	0,05/0,13	1,34 (1,18-1,52)	8,35E-06
rs2717031	2	58166468	T/C	1,00/1,00	0,47/0,40	0,04/0,09	1,34 (1,18-1,52)	8,64E-06
rs2717002	2	58142647	T/G	1,00/1,00	0,47/0,41	0,17/0,05	1,34 (1,18-1,51)	8,67E-06
rs7236787	18	25800169	G/A	1,00/1,00	0,27/0,34	0,18/0,59	0,73 (0,64-0,84)	8,68E-06
rs2290867	2	57952870	A/C	1,00/1,00	0,44/0,38	0,36/0,11	1,34 (1,18-1,51)	8,78E-06
rs2717055	2	58044220	G/A	1,00/1,00	0,44/0,38	0,19/0,10	1,34 (1,18-1,51)	9,31E-06
rs9959163	18	25811156	G/A	1,00/1,00	0,27/0,34	0,18/0,65	0,73 (0,64-0,84)	9,47E-06
rs1518395	2	58208074	A/G	1,00/1,00	0,47/0,40	0,23/0,17	1,33 (1,18-1,51)	9,93E-06
rs1474215	2	58032029	C/G	1,00/1,00	0,44/0,38	0,19/0,14	1,34 (1,18-1,51)	9,99E-06

Tabelle 3-7 SNPs mit $P_{GC} \leq 10^{-5}$ in der logistischen Regression in der IAE-Subgruppe

Legende -> CHR: Chromosom; POS NCBI 37.3; physikalische Position nach der NCBI Version 37.3; CR, MAF, HWE-P F/K: CR, MAF und HWE-P-Werte jeweils getrennt nach Fälle und Kontrollen; OR (95%-KI): Odds Ratio mit 95%-Konfidenzintervall bezogen auf Allel A1.

SNP	CHR	POS NCBI 37.3	Allele A1/A2	CR F/K	MAF F/K	HWE-P F/K	OR (95%-KI)	LR-P _{GC}
rs12059546	1	239970097	G/A	1,00/1,00	0,24/0,17	0,82/0,62	1,52 (1,29-1,77)	5,15E-07
rs1110615	1	239974170	A/G	1,00/1,00	0,24/0,17	0,82/0,43	1,51 (1,29-1,77)	6,30E-07
rs4892034	18	70399988	T/A	0,99/0,98	0,26/0,20	0,67/0,80	1,47 (1,26-1,72)	1,92E-06
rs12963674	18	55978595	T/G	1,00/1,00	0,08/0,05	0,16/0,10	1,83 (1,42-2,37)	7,93E-06
rs1497633	8	109751881	C/G	1,00/1,00	0,32/0,39	0,26/0,73	0,72 (0,62-0,83)	8,10E-06
rs16906789	10	56751811	G/A	0,99/0,99	0,25/0,19	0,37/0,84	1,45 (1,24-1,69)	8,18E-06
rs1391196	8	109745945	A/T	1,00/1,00	0,32/0,39	0,26/0,73	0,72 (0,62-0,83)	8,40E-06

Tabelle 3-8 SNPs mit $P_{GC} \leq 10^{-5}$ in der logistischen Regression in der JME-Subgruppe

Legende -> CHR: Chromosom; POS NCBI 37.3; physikalische Position nach der NCBI Version 37.3; CR, MAF, HWE-P F/K: CR, MAF und HWE-P-Werte jeweils getrennt nach Fälle und Kontrollen; OR (95%-KI): Odds Ratio mit 95%-Konfidenzintervall bezogen auf Allel A1.

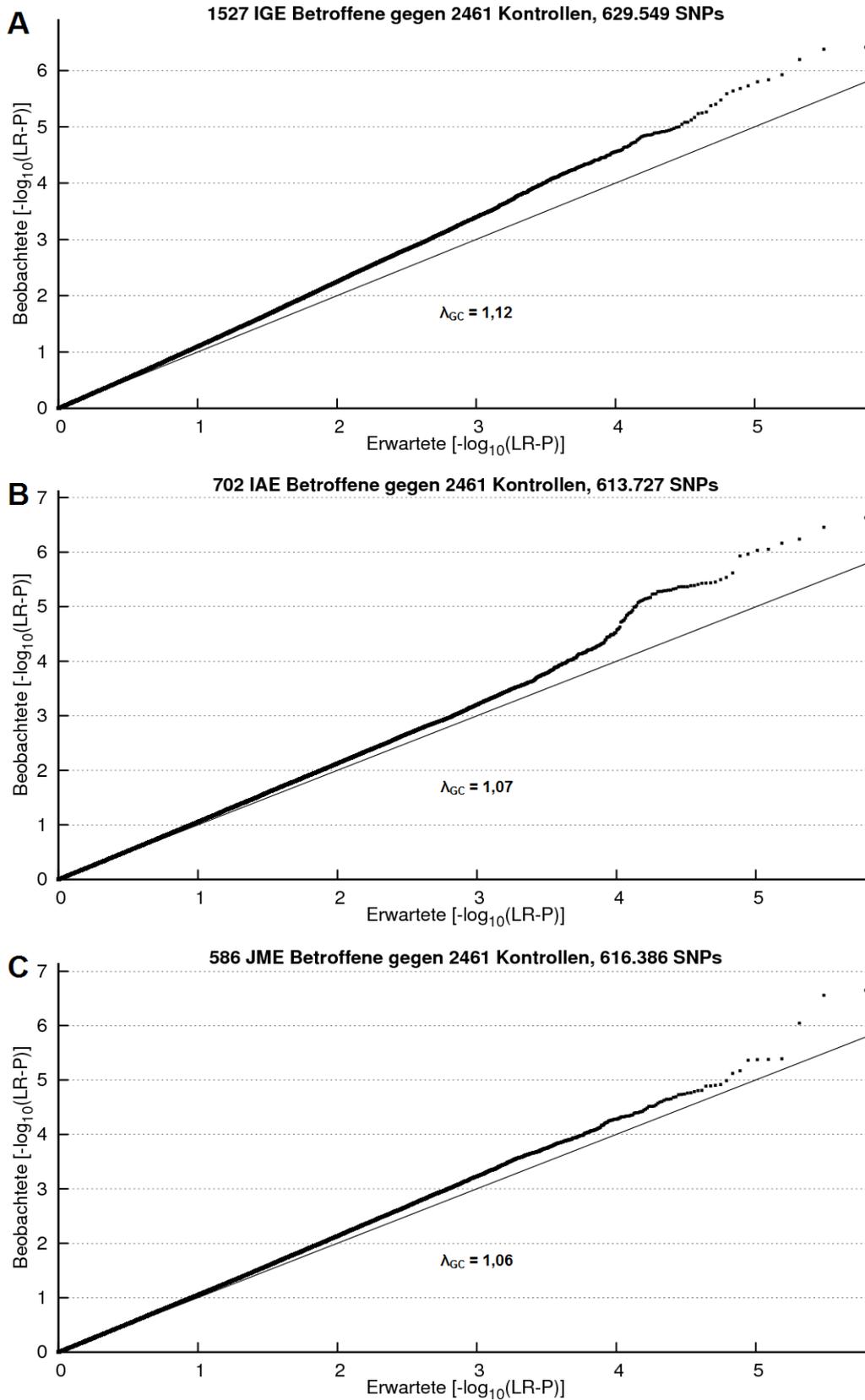


Abb. 3-7 Q-Q-Plots der GWAS im IGE-Gesamtkollektiv und den Subgruppen

Gezeichnet Verteilung der $-\log_{10}(\text{LR-P})$ -Werte (vor der GC-Korrektur) gegen die erwartete Verteilung in (A) der IGE-Gesamtgruppe bei 629.549 SNPs, (B) der IAE-Subgruppe bei 613.727 SNPs und (C) der JME-Subgruppe bei 616.386 SNPs.

3.2.4 Replikation der GWAS-Ergebnisse

Von den 42 SNPs mit $LR-P_{GC} \leq 10^{-5}$ aus der GWAS im IGE-Gesamtkollektiv oder in einer der Subgruppen, wurden 21 SNPs für die Replikationsanalyse ausgewählt (14 unabhängige chromosomale Regionen). Sieben dieser SNPs wurden aus den nachfolgenden Analysen ausgeschlossen (nicht erfolgreiche SNPstream-Genotypisierung: $n = 3$; Qualitätskriterien nicht bestanden: $n = 4$). Neben den üblichen Qualitätskriterien wurden die genotypisierten SNPs auf Konkordanz und Replizierbarkeit hin überprüft. Für diesen Zweck wurden zusätzlich 383 IGE-Einzelfälle aus dem genomweiten Scan und 79 Duplikate aus dem Replikationskollektiv genotypisiert. Die Genotypen der verbleibenden 14 SNPs wiesen eine Konkordanz von $>99\%$ zu Affymetrix SNP 6.0-Genotypen und eine Replizierbarkeit von 100% in den duplizierten Proben auf (Anhang, Tabelle 5-1). Die verbleibenden 14 SNPs beschreiben 10 unabhängige chromosomale Regionen. Die Replikations-P-Werte wurden nach der AFBAC-Methode mit dem stratifizierten CMH-Test berechnet (Kap. 2.11.3 und 2.12.2). Die Stichproben wurden dafür in 19 Cluster nach der ethnischen Herkunft eingeteilt (Tabelle 2-3).

In der Replikationsanalyse im IGE-Gesamtkollektiv erreichten zwei SNPs im LD in der chromosomalen Region 17q21.32 die Schwelle für eine signifikante Assoziation ($P < 0,05 / 42$) nach der Bonferroni Korrektur für multiples Testen (rs41336845; chr17:46215110; CMH-P = $2,54 \times 10^{-4}$; OR[T] = 0,76; 95%-KI: 0,65-0,88 // rs16953235; chr17:46244948; CMH-P = $7,92 \times 10^{-4}$; OR[G] = 0,78; 95%-KI: 0,67-0,90). Es ergaben sich keine Hinweise auf Heterogenität in den Strata des Fall-Pseudo-Kontroll-Kollektivs der Replikationsanalyse (rs41336845: BD-P = 0,98 // rs16953235: BD-P = 0,99). In der Meta-Analyse der GWAS mit der Replikationsanalyse erreichten beide SNPs die Schwelle für eine genomweit signifikante Assoziation (rs41336845; META-P = $7,90 \times 10^{-9}$; OR[T] = 0,76; 95%-KI: 0,70-0,84 // rs16953235; META-P = $2,67 \times 10^{-8}$; OR[G] = 0,77; 95%-KI: 0,70-0,85; Abb. 3-8).

In den Subgruppenanalysen wurden weder in der Replikationsanalyse, noch in der Meta-Analyse signifikante Assoziationsergebnisse erreicht. In der IAE-Subgruppe erreichen zwei SNPs suggestive META-P-Werte: rs2953442 bei 2p16.1 (chr2:58083420; META-P = $1,82 \times 10^{-7}$; OR[C] = 1,31; 95%-KI: 1,19-1,45) und rs17741930 bei 2q22.3 (chr2:145362155; META-P = $1,53 \times 10^{-7}$; OR[C] = 0,69; 95%-KI: 0,60-0,80). In der JME-Subgruppe erreichten zwei SNPs im LD bei 1q43 suggestive META-P-Werte (rs12059546; chr1:239970097; META-P = $4,37 \times 10^{-7}$; OR[G] = 1,45; 95%-KI: 1,28-1,65 // rs1110615; chr1:239974170; META-P = $4,22 \times 10^{-7}$; OR[A] = 1,46; 95%-KI: 1,28-1,65).

Alle Ergebnisse sind in den Tabellen 3-9 und 3-10 zusammengefasst.

GRP	SNP	CHR	POS	GWAS	Replikation	Meta-Analyse	Effekt	Heterogenität	
			NCBI 37.3	LR-P _{GC}	CMH-P	META-P		BD-P	M-Het-P
IGE	rs12059546	1	239970097	3,22E-05	0,037	5,63E-06	++	0,128	0,317
IGE	rs1110615	1	239974170	7,82E-05	0,074	2,78E-05	++	0,177	0,266
IGE	rs388556	2	13392540	1,75E-06	0,289	2,52E-03	+-	0,141	1,16E-04
IGE	rs2953442	2	58083420	4,08E-05	2,67E-03	3,83E-07	++	0,762	0,783
IGE	rs17741930	2	145362155	6,82E-04	0,081	1,88E-04	--	0,081	0,426
IGE	rs289034	5	114122165	2,56E-06	0,796	1,42E-04	++	0,656	5,48E-03
IGE	rs2268330	14	23795976	1,64E-06	0,425	1,37E-03	+-	0,225	2,56E-04
IGE	rs12904369	15	97335674	5,84E-03	0,159	0,214	+-	0,421	4,58E-03
IGE	rs11650615	17	46123698	5,77E-06	0,012	3,26E-07	--	0,934	0,357
IGE	rs41336845	17	46215110	8,10E-06	2,54E-04	7,90E-09	--	0,980	0,995
IGE	rs16953235	17	46244948	8,96E-06	7,92E-04	2,67E-08	--	0,992	0,829
IGE	rs7236787	18	25800169	2,01E-04	0,934	3,44E-03	--	0,247	0,022
IGE	rs10853686	18	25870524	1,02E-04	0,359	3,36E-04	--	0,348	0,079
IGE	rs4892034	18	70399988	6,72E-06	0,582	1,29E-04	++	0,339	0,015
IAE	rs12059546	1	239970097	0,1572	0,290	0,078	++	0,552	0,928
IAE	rs1110615	1	239974170	0,2087	0,733	0,237	++	0,441	0,585
IAE	rs388556	2	13392540	3,54E-05	0,992	1,54E-03	--	0,167	7,83E-03
IAE	rs2953442	2	58083420	8,20E-07	0,025	1,82E-07	++	0,401	0,147
IAE	rs17741930	2	145362155	1,97E-06	0,013	1,53E-07	--	0,226	0,255
IAE	rs289034	5	114122165	2,72E-05	0,911	1,67E-03	+-	0,665	5,41E-03
IAE	rs2268330	14	23795976	1,81E-04	0,659	9,84E-03	+-	0,626	6,01E-03
IAE	rs12904369	15	97335674	7,21E-06	0,451	3,15E-03	+-	0,469	5,37E-04
IAE	rs11650615	17	46123698	6,40E-04	0,028	5,62E-05	--	0,965	0,608
IAE	rs41336845	17	46215110	1,28E-03	1,24E-03	5,53E-06	--	0,950	0,685
IAE	rs16953235	17	46244948	1,50E-03	2,42E-03	1,18E-05	--	0,967	0,779
IAE	rs7236787	18	25800169	8,68E-06	0,916	5,20E-04	--	0,819	5,37E-03
IAE	rs10853686	18	25870524	5,05E-06	0,358	4,46E-05	--	0,584	0,025
IAE	rs4892034	18	70399988	3,10E-03	0,504	7,08E-03	++	0,325	0,164
JME	rs12059546	1	239970097	5,15E-07	0,071	4,37E-07	++	0,752	0,087
JME	rs1110615	1	239974170	6,30E-07	0,062	4,22E-07	++	0,810	0,100
JME	rs388556	2	13392540	0,01337	0,472	0,139	+-	0,532	0,035
JME	rs2953442	2	58083420	0,1403	0,063	0,021	++	0,608	0,596
JME	rs17741930	2	145362155	0,2808	0,410	0,175	--	0,667	0,978
JME	rs289034	5	114122165	1,69E-04	0,734	1,64E-03	++	0,242	0,037
JME	rs2268330	14	23795976	1,46E-04	0,885	4,06E-03	+-	0,288	0,013
JME	rs12904369	15	97335674	0,929	0,213	0,398	--	0,453	0,359
JME	rs11650615	17	46123698	5,90E-05	0,048	1,23E-05	--	0,810	0,336
JME	rs41336845	17	46215110	3,17E-04	7,66E-03	7,55E-06	--	0,988	0,867
JME	rs16953235	17	46244948	3,09E-04	0,011	1,05E-05	--	0,995	0,790
JME	rs7236787	18	25800169	1,90E-03	0,375	0,062	+-	0,313	8,42E-03
JME	rs10853686	18	25870524	9,38E-04	0,302	0,053	+-	0,486	4,04E-03
JME	rs4892034	18	70399988	1,92E-06	0,485	3,32E-05	++	0,435	0,015

Tabelle 3-9 Replikation der GWAS und Meta-Analyse von 14 SNPs

Legende -> GPR: Analysegruppe; Effekt R.: Effektrichtung; BD-P / M-Het-P: Heterogenitäts-P-Werte der Replikationsanalyse bzw. der Meta-Analyse. Rot gekennzeichnet sind ungleiche Effektrichtungen und auffällige M-Het-P < 0,05. Die genomweit signifikant assoziierten SNPs sind fett hervorgehoben.

GRP	SNP	CHR	POS NCBI 37.3	Allele A1/A2	GWAS OR (95%KI)	Replikation OR (95%KI)	Meta-Analyse OR (95%KI)	Effekt R.
IGE	rs12059546	1	239970097	G/A	1,30 (1,16-1,46)	1,16 (1,01-1,34)	1,23 (1,13-1,35)	++
IGE	rs1110615	1	239974170	A/G	1,28 (1,14-1,44)	1,14 (0,99-1,31)	1,21 (1,11-1,33)	++
IGE	rs388556	2	13392540	G/C	0,76 (0,69-0,85)	1,07 (0,95-1,20)	0,88 (0,81-0,95)	+-
IGE	rs2953442	2	58083420	C/T	1,25 (1,13-1,37)	1,20 (1,07-1,35)	1,22 (1,13-1,31)	++
IGE	rs17741930	2	145362155	C/T	0,78 (0,69-0,90)	0,87 (0,75-1,02)	0,83 (0,75-0,92)	--
IGE	rs289034	5	114122165	T/C	1,27 (1,16-1,40)	1,02 (0,91-1,14)	1,15 (1,07-1,23)	++
IGE	rs2268330	14	23795976	G/C	1,33 (1,19-1,49)	0,95 (0,84-1,08)	1,14 (1,05-1,24)	+-
IGE	rs12904369	15	97335674	C/A	1,17 (1,05-1,30)	0,91 (0,80-1,04)	1,07 (0,98-1,16)	+-
IGE	rs11650615	17	46123698	G/C	0,77 (0,70-0,86)	0,85 (0,75-0,96)	0,80 (0,74-0,87)	--
IGE	rs41336845	17	46215110	T/C	0,75 (0,66-0,84)	0,76 (0,65-0,88)	0,76 (0,70-0,84)	--
IGE	rs16953235	17	46244948	G/A	0,75 (0,67-0,85)	0,78 (0,67-0,90)	0,77 (0,70-0,85)	--
IGE	rs7236787	18	25800169	G/A	0,82 (0,74-0,90)	0,99 (0,88-1,12)	0,89 (0,82-0,96)	--
IGE	rs10853686	18	25870524	T/C	0,82 (0,75-0,90)	0,95 (0,85-1,06)	0,87 (0,81-0,94)	--
IGE	rs4892034	18	70399988	T/A	1,32 (1,18-1,47)	1,04 (0,90-1,21)	1,22 (1,11-1,33)	++
IAE	rs12059546	1	239970097	G/A	1,12 (0,96-1,31)	1,11 (0,92-1,34)	1,16 (1,02-1,31)	++
IAE	rs1110615	1	239974170	A/G	1,11 (0,95-1,30)	1,03 (0,86-1,25)	1,11 (0,98-1,26)	++
IAE	rs388556	2	13392540	G/C	0,74 (0,65-0,85)	1,00 (0,86-1,17)	0,84 (0,76-0,94)	--
IAE	rs2953442	2	58083420	C/T	1,39 (1,23-1,58)	1,19 (1,02-1,39)	1,31 (1,19-1,45)	++
IAE	rs17741930	2	145362155	C/T	0,62 (0,52-0,75)	0,77 (0,63-0,95)	0,69 (0,60-0,80)	--
IAE	rs289034	5	114122165	T/C	1,31 (1,16-1,48)	0,99 (0,85-1,15)	1,18 (1,07-1,30)	+-
IAE	rs2268330	14	23795976	G/C	1,32 (1,15-1,52)	0,96 (0,82-1,14)	1,17 (1,05-1,30)	+-
IAE	rs12904369	15	97335674	C/A	1,38 (1,20-1,57)	0,94 (0,79-1,11)	1,20 (1,07-1,33)	+-
IAE	rs11650615	17	46123698	G/C	0,78 (0,68-0,90)	0,83 (0,70-0,98)	0,82 (0,74-0,91)	--
IAE	rs41336845	17	46215110	T/C	0,77 (0,65-0,90)	0,72 (0,59-0,88)	0,77 (0,68-0,87)	--
IAE	rs16953235	17	46244948	G/A	0,77 (0,66-0,90)	0,74 (0,61-0,90)	0,78 (0,69-0,88)	--
IAE	rs7236787	18	25800169	G/A	0,73 (0,64-0,84)	0,99 (0,85-1,16)	0,83 (0,75-0,92)	--
IAE	rs10853686	18	25870524	T/C	0,74 (0,66-0,84)	0,93 (0,80-1,08)	0,82 (0,74-0,90)	--
IAE	rs4892034	18	70399988	T/A	1,26 (1,09-1,45)	1,07 (0,88-1,30)	1,18 (1,05-1,33)	++
JME	rs12059546	1	239970097	G/A	1,52 (1,29-1,77)	1,21 (0,98-1,48)	1,45 (1,28-1,65)	++
JME	rs1110615	1	239974170	A/G	1,51 (1,29-1,77)	1,21 (0,99-1,49)	1,46 (1,28-1,65)	++
JME	rs388556	2	13392540	G/C	0,83 (0,71-0,96)	1,07 (0,89-1,28)	0,92 (0,82-1,03)	+-
JME	rs2953442	2	58083420	C/T	1,12 (0,97-1,29)	1,18 (0,99-1,41)	1,13 (1,01-1,26)	++
JME	rs17741930	2	145362155	C/T	0,90 (0,75-1,08)	0,91 (0,73-1,14)	0,94 (0,81-1,09)	--
JME	rs289034	5	114122165	T/C	1,31 (1,14-1,50)	1,03 (0,87-1,22)	1,20 (1,08-1,33)	++
JME	rs2268330	14	23795976	G/C	1,36 (1,17-1,59)	0,99 (0,82-1,19)	1,19 (1,06-1,35)	+-
JME	rs12904369	15	97335674	C/A	0,99 (0,85-1,16)	0,89 (0,73-1,07)	0,97 (0,85-1,09)	--
JME	rs11650615	17	46123698	G/C	0,73 (0,62-0,84)	0,82 (0,68-1,00)	0,78 (0,69-0,89)	--
JME	rs41336845	17	46215110	T/C	0,72 (0,60-0,86)	0,74 (0,59-0,92)	0,75 (0,65-0,86)	--
JME	rs16953235	17	46244948	G/A	0,72 (0,60-0,85)	0,75 (0,60-0,94)	0,75 (0,65-0,87)	--
JME	rs7236787	18	25800169	G/A	0,79 (0,69-0,91)	1,08 (0,91-1,29)	0,88 (0,78-0,98)	+-
JME	rs10853686	18	25870524	T/C	0,79 (0,69-0,91)	1,09 (0,92-1,29)	0,88 (0,79-0,97)	+-
JME	rs4892034	18	70399988	T/A	1,47 (1,26-1,72)	1,08 (0,87-1,35)	1,32 (1,17-1,50)	++

Tabelle 3-10 Odds Ratios der GWAS, der Replikation und der Meta-Analyse

Legende -> GPR: Analysegruppe; Effekt R.: Effektrichtung; OR (95%KI): Odds Ratio mit 95%-Konfidenzintervall bezogen auf Allel A1. Ungleiche Effektrichtungen sind rot gekennzeichnet. Die zwei genomweit signifikant assoziierten SNPs auf 17q21.32 sind fett hervorgehoben.

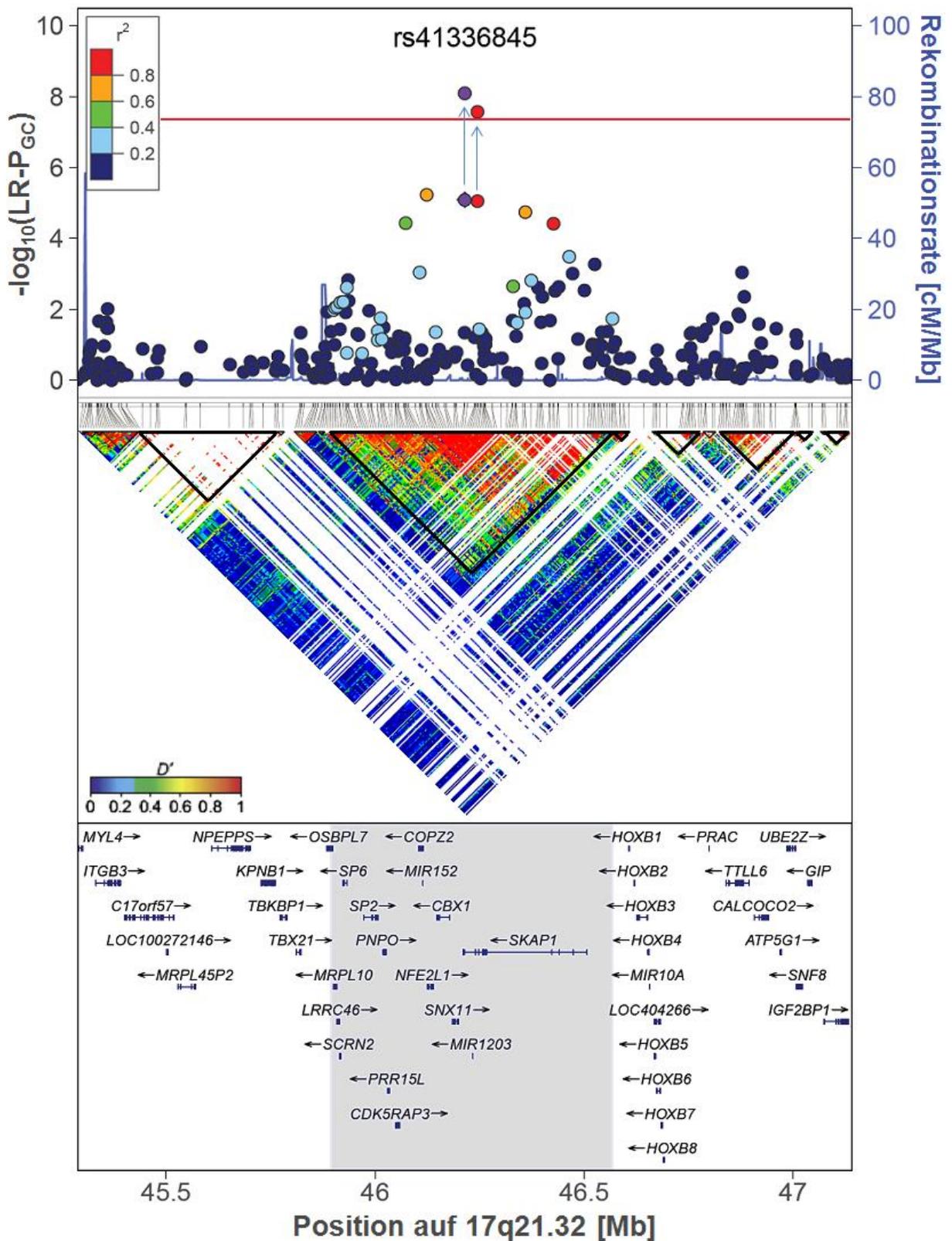


Abb. 3-8 Struktur der signifikant assoziierten Region auf 17q21.32

Gezeichnet $-\log_{10}(\text{LR-P}_{\text{GC}})$ -Werte der GWAS im IGE-Gesamtkollektiv gegen die phys. Position (NCBI 37.3), LD-Struktur und Gene im Bereich der Assoziation. Für zwei SNPs sind zusätzlich META-P-Werte eingezeichnet (Pfeile). Die rote Linie zeichnet die Signifikanzschwelle bei $P = 5 \times 10^{-8}$. Die Farbkodierung der SNPs richtet sich nach dem r^2 -Wert zu rs41336845. In den LD-Blöcken (schwarze Dreiecke) sind min. 30% aller $D' > 0,7$ basierend auf 3988 Personen der GWAS. Die Gene im Bereich des signifikanten LD-Blocks sind grau unterlegt.

Nach Überprüfung der regionalen LD-Struktur konnte die mit IGE signifikant assoziierte 17q21.23 Region auf einem 676.357 bp großen Bereich zwischen rs2010952 und rs8072250 eingegrenzt werden (chr17:45892859-46569216) (Abb. 3-8). Der Bereich wird durch einen einzigen LD-Block mit variierendem LD ausgemacht, bei mindestens 30% aller paarweisen $D' > 0,7$. Bereiche mit erhöhten Rekombinationsraten flankieren den LD-Block und bestätigen die Begrenzung der assoziierten Region („downstream“: 27 cM/Mb; „upstream“: 10 cM/Mb; normal 1,13 cM/Mb nach Kong et al., 2002).

In diesem Bereich befinden sich 16 Gene. Unter diesen ist *PNPO* (pyridoxamine 5'-phosphate oxidase) ein hochrangiges Kandidatengen für IGE. Mutationen im *PNPO*-Gen verursachen die PNPO-abhängige neonatale epileptische Enzephalopathie (Mills et al., 2005; Musayev et al., 2009). Ein weiteres Kandidatengen ist *CDK5RAP3* (CDK5 regulatory subunit associated protein 3), dessen Genprodukt als Bindungspartner von CDK5R1 (cyclin-dependent kinase 5, regulatory subunit 1 (p35)) eine Rolle in der Neurogenese spielen könnte (Ching et al., 2000).

4 DISKUSSION

Für die genetisch mitbedingten, komplexen Erkrankungen sind verschiedene Hypothesen entwickelt worden, als Erklärungsmodelle für die Architektur der genetischen Faktoren für die beobachteten klinischen Syndrome. Die „common disease/common variant“- (CD/CV-) Hypothese geht davon aus, dass genetische Variationen mit niedrigen Effektstärken ($OR < 2$) häufig in der allgemeinen Bevölkerung vorkommen (Freq. $> 1\%$) und gemeinsam einen bedeutenden Teil der Heritabilität einer Erkrankung erklären können (Lohmueller et al., 2003). Häufige krankheitsrelevante Variationen werden meist ab einer Frequenz $> 5\%$ von den üblichen Assoziationsstudien erfasst (Abb. 4-1; Hindorff et al., 2009). Alternativ dazu postuliert die „common disease multiple rare variant“- (CD/MV-) Hypothese, dass ein bedeutender Teil der Heritabilität durch eine große Anzahl von sehr seltenen Variationen mit hohen Effektstärken (Freq. $< 1\%$; $OR > 2$) erklärt werden kann (Pritchard und Cox, 2002). Genetische Variationen die zu monogenen oder oligogenen Formen einer Erkrankung führen, können als Extremfälle der CD/MV-Hypothese betrachtet werden. Seltene krankheitsrelevante Variationen können durch Kopplungsstudien (wenn $OR > 3$; Anderson et al., 2011) oder durch Sequenzierung erfasst werden.

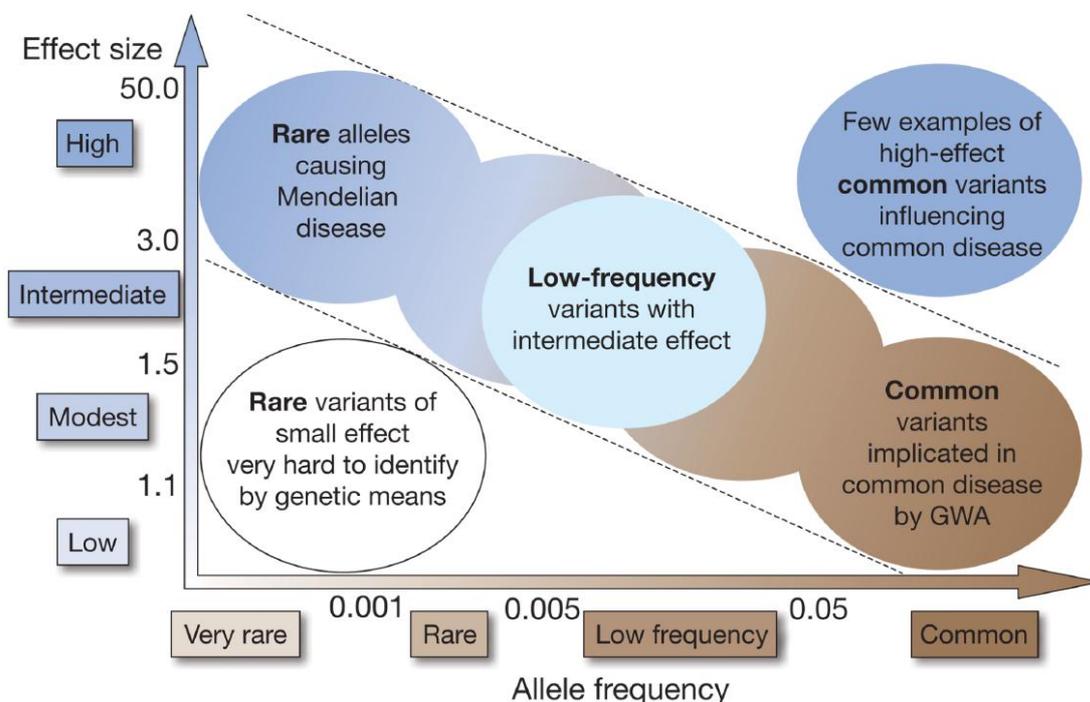


Abb. 4-1 Identifizierbarkeit von Variationen nach Frequenz und Effektstärke

Dargestellt sind Bereiche aus der möglichen Streuung von genetischen Variationen, in Abhängigkeit der Allelfrequenz und der Effektstärke. Die gestrichelten Linien grenzen die Bereiche ein, die mit gegenwärtigen Methoden erfasst werden können (nach Manolio et al., 2009).

Aufgrund der hohen genetischen Heterogenität die bei Epilepsie zu beobachten ist, können die CD/CV- und CD/MV-Hypothesen als Teilaspekte eines komplexeren, multifaktoriell-polygenen-Modells betrachtet werden (Mulley et al., 2005; Dibbens et al., 2007b). Nach diesem Modell werden alle genetischen Variationen, sowohl häufige als auch seltene, in unterschiedlichen Kombinationen zu der Entstehung von Epilepsie prädisponieren und die verschiedenen Untersyndrome hervorrufen. Dabei werden genetische Varianten mit hohen Effektstärken in Multiplexfamilien angereichert sein, während Einzelfälle eine breitere Streuung der Effektstärken sowie eine höhere Anzahl an prädisponierende Genvarianten aufweisen werden (Winawer et al., 2002; Ottman, 2005). Zur Dissektion der genetischen Faktoren bei den häufigen IGE-Syndromen mit genetisch komplexer Disposition werden derzeit drei molekulargenetische Strategien angewandt, um die unterschiedlichen genetischen Komponenten des multifaktoriellen polygenen Modells zu erfassen. Bei familiären IGE-Syndromen mit vermuteter oligogener Disposition sind genomweite Kopplungsstudien eine aussichtsreiche Strategie, um einzelne genetische Faktoren mit relativ starken Effektstärken zu lokalisieren und positionell zu identifizieren. Komplementär sind genomweite Assoziationsanalysen darauf ausgerichtet, häufige Sequenzvarianten mit meist geringen disponierenden Effekten (CD/CV-Hypothese) zu erfassen. Zur Ermittlung zahlreicher seltener Sequenzvariationen mit ausgeprägter Effektstärke (CD/MV-Hypothese) werden derzeit Sequenzanalysen der kompletten Exome bzw. Genome durchgeführt. Zusammengenommen haben diese drei molekulargenetischen Ansätze ein aussichtsreiches Potential einen Grossteil der Heritabilität zu erklären.

In der vorliegenden Studie sollte ein möglichst breites Spektrum von genetischen Varianten, mittels Kopplungsanalysen in familiären IGE-Formen sowie Assoziationsanalysen in häufigen IGE-Formen, erfasst werden. Die Grundlage für diese Studien sind die derzeit mit Abstand größten Kollektive an IGE-Multiplexfamilien und IGE-Einzelfällen, die eine bisher nicht erreichte Teststärke gewährleisten, um genetische Faktoren trotz einer ausgeprägten genetischen Heterogenität und einer derzeit noch weitgehend unverstandenen genetischen Architektur zu identifizieren.

Unsere genomweite Kopplungsanalyse identifizierte zwei krankheitsassoziierte Loci in den Chromosomenabschnitten 2q32.1-q36.1 und 13q31.1-q32.1, die differentiell zum Entstehungsrisiko von Fällen familiärer JME bzw. IAE beitragen. Für die beiden Loci auf den Chromosomen 2 und 13 gibt es aus früheren Studien suggestive Kopplungshinweise mit IGE (2q36.1: Sander et al., 2000; 13q22-q31: Hempelmann et al., 2006). Allerdings können aufgrund der Überlappung dieser früheren Studien mit den in dieser Meta-Kopplungsanalyse

eingeschlossen Familien (28% bzw. 26%) beide Kopplungsbefunde nicht als eigenständige Replikationen betrachtet werden.

Unsere weltweit erste GWAS für IGE, diese führte zu einem neuen, replizierbaren Locus in der chromosomalen Region 17q21.32, der in unserer europäischen Kohorte signifikant mit den häufigen Formen von IGE assoziiert ist.

4.1 Genomweite Kopplungsstudie

Die genomweite Kopplungsanalyse in 379 IGE-Multiplexfamilien wurde mit dem Ziel durchgeführt, Krankheitsloci zu identifizieren, die zur Ausprägung eines breiten Spektrums von familiären IGE-Syndromen prädisponieren. Während keine signifikanten Kopplungen ermittelt werden konnten, fanden sich suggestive Kopplungshinweise für sechs Regionen (1p36.22, 3p21.1-p13, 5q31.3-q35.1, 13q12.11-q12.13, 13q31.1-q32.1, 19q13.32-q13.42). Statistisch ist bei einer genomweiten Analyse zu erwarten, dass die Schwelle für eine suggestive Kopplung nur einmal zufällig erreicht wird. Dieses Ergebnis unterstützt damit das oligogene Vererbungsmodell für die familiären IGE-Syndrome.

Die genomweiten Kopplungsanalysen in den zwei Subgruppen mit 235 IAE-Multiplexfamilien bzw. mit 118 JME-Multiplexfamilien führten zu signifikanten Kopplungsbefunden in den Chromosomenabschnitten 13q31.1-q32.1 bzw. 2q32.1-q36.1. Dieses Ergebnis deutet an, dass durch die Subklassifizierung nach Anfallstyp in Kombination mit einem strikten Betroffenheitsmodell, in dem ausschließlich Familienmitglieder mit IAE oder JME als „betroffen“ klassifiziert wurden, die genetische Heterogenität der familiären IGE-Syndromen reduziert werden konnte. Gleichzeitig konnten in den Subgruppen genetische Determinanten angereichert werden, die differentiell zum Entstehungsrisiko von Absencen oder Myoklonien beitragen. Neben anfallsspezifischen Determinanten unterstützen beide Subgruppenanalysen die Kopplungsregion auf 5q31.3-q35.1, eine Region in der sich ein Gen-Cluster mit den Genen für die GABA_A-Rezeptor-β2-, -α6-, -α1- und -γ2-Untereinheiten (*GABRB2*, *GABRA6*, *GABRA1*, *GABRG2*) befindet. Mutationen in den Genen *GABRA1* und *GABRG2* sind bei mehreren Formen von Epilepsie beschrieben worden (CAE, JME, GESF+, SMEI; Tabelle 1-3). Der Kopplungshinweis deutet darauf hin, dass auch bei einigen der untersuchten IGE-Familien Mutationen in den Genen *GABRA1* und *GABRG2* vorliegen könnten.

Funktionell relevante Kandidatengene rechtfertigen auch eine weitere Betrachtung der suggestiven Kopplungshinweise in den Chromosomenabschnitten 3p21.1-p13 und 19q13.32-q13.42. Innerhalb der Kopplungsregion auf 3p21.1-p13 befindet sich das Gen *PRICKLE2* (prickle homolog 2; OMIM: *608501). Heterozygote Mutationen in dem *PRICKLE2*-Gen füh-

ren zu einer progressiven Myoklonusepilepsie (EPM5; OMIM: #613832; Tao et al., 2011). Innerhalb der Kopplungsregion auf 19q13.32-q13.42 befindet sich das Gen *PNKP* (polynucleotide kinase 3'-phosphatase; OMIM: *605610). Homozygote Mutationen in dem *PNKP*-Gen führen zu einer frühinfantilen epileptischen Enzephalopathie (EIEE10; OMIM: #613402; Shen et al., 2010). In derselben Kopplungsregion befindet sich auch ein Gen-Cluster mit den Genen für die Ca²⁺-Kanal- γ 7-, - γ 8- und - γ 6-Untereinheiten (CACNG7, CACNG8, CACNG6). Die Expression der Ca²⁺-Kanal- γ 2-Untereinheit (CACNG2), die funktionell mit γ 7, γ 8 und γ 6 verwandt ist, ist in der „stargazer“-Maus durch eine Mutation unterbrochen (Letts et al., 1998). Diese Mausmutante ist ein Modell für die Absence-Epilepsie. Eine weitere Ca²⁺-Kanal-Untereinheit (γ 3; CACNG3) prädisponiert für die familiäre Form von CAE (Everett et al., 2007). Generell sind Calciumkanäle und ihre Regulation durch verschiedene Hilfsuntereinheiten wie β und γ ausschlaggebend für die Depolarisation von Nervenzellen.

4.2 Genomweite Assoziationsstudie

Die genomweite Assoziationsstudie von 1527 IGE-Einzelfällen und 2461 Populationskontrollen (Ursprung Deutschland und Anreinerstaaten) wurde mit dem Ziel durchgeführt, genetische Varianten zu identifizieren, die häufig in der Allgemeinbevölkerung vorkommen und zur Ausprägung von IGE prädisponieren. Nach der Replikationsrunde in einem unabhängigen Kollektiv von 1299 IGE-Einzelfällen und 1240 Kontrollen (breitgefächerter, europäischer Ursprung) konnte in der Meta-Analyse eine genomweit signifikante Assoziation mit IGE auf 17q21.32 und eine suggestive Assoziation auf 2p16.1 festgestellt werden. Obwohl die vorliegende Studie mit Abstand das weltweit größte IGE-Kollektiv einschloss (2826 IGE-Probanden und 3701 Kontrollen) hatte das IGE-Gesamtkollektiv nur eine moderate Power von 80%, um Risikoallele mit einem OR = 1,33 zu identifizieren. Ein OR von 1,33 entspricht dem Median aller ORs von Risikoallelen, die bei häufigen Erkrankungen bisher identifiziert worden sind (Hindorff et al., 2009). Aufgrund der niedrigen Power des genomweiten Scans (Power = 80% bei OR = 1,46 bzw. Power = 27% bei OR = 1,33) ist davon auszugehen, dass nur ein geringer Anteil aller relevanten genetischen Varianten in dem genomweiten Scan erkannt und für die Replikationsanalyse ausgewählt worden sind.

Zielsetzung der anfallstypbezogenen Subgruppenanalysen war es, genetische Varianten zu identifizieren, die differentiell zum Entstehungsrisiko von entweder IAE oder JMEs beitragen. Die Meta-Analysen hatten jeweils eine Power von 80%, um Risikoallele mit einem OR = 1,45 in der IAE-Subgruppe bzw. OR = 1,51 in der JME-Subgruppe zu identifizieren. In keiner der Subgruppenanalysen konnten signifikante Assoziationen festgestellt werden. In der IAE-

Subgruppe fanden sich suggestive Assoziationen auf 2p16.1 und 2q22.3. In der JME-Subgruppe konnte eine suggestive Assoziation auf 1q43 festgestellt werden. Der Verlust an Power durch die Reduzierung der Stichprobengröße konnte offenbar durch die erhöhte klinische und genetische Homogenität durch die Sortierung nach Anfallstyp nicht aufgewogen werden. Entsprechend ist auch bei den Subgruppenanalysen davon auszugehen, dass die Mehrheit der relevanten genetischen Varianten nicht in die Replikationsanalyse eingegangen ist.

Die suggestiv assoziierten Regionen enthalten plausible funktionelle Kandidatengene für IGE. In der Region auf Chromosom 2p16.1 befindet sich das Gen *VRK2* (vaccinia related kinase 2; OMIM: *602169). *VRK2* ist ein Kandidatengen für Schizophrenie (Steinberg et al., 2011) und für das 2p16.1-p15-Deletionssyndrom (OMIM: #612513; Chabchoub et al., 2008). Patienten mit einem 2p16.1-p15-Deletionssyndrom weisen einen schweren Phänotyp mit geistiger Behinderung, Autismus-Spektrum-Störung, Mikrozephalie, kortikaler Dysplasie und therapieresistenten epileptischen Anfällen auf (Chandler et al., 2006; Rajcan-Separovic et al., 2007). In der direkten Umgebung des suggestiven Assoziationssignals auf Chromosom 2q22.3 (84 kb „downstream“) befindet sich das Gen *ZEB2* (zinc finger E-box binding homeobox 2; OMIM: *605802). Heterozygote Mutationen oder Deletionen des *ZEB2*-Gens sind verantwortlich für alle Fälle des Mowat-Wilson-Syndroms (OMIM: #235730; Wakamatsu et al., 2001; Yamada et al., 2001). Das Mowat-Wilson-Syndrom ist gekennzeichnet von einer typischen Gesichtsdysmorphie mit geistiger Behinderung und Mikrozephalie. Epilepsie kommt bei 74% der Patienten vor (Garavelli et al., 2009). Bis auf wenige Ausnahmen von Keimzellmosaizismus, sind alle *ZEB2*-Mutationen oder Deletionen „de novo“ Erscheinungen (Garavelli und Mainardi, 2007). Das suggestive Assoziationssignal auf 1q43 befindet sich in dem intronischen Bereich des Gens *CHRM3* (cholinergic receptor, muscarinic 3; OMIM: *118494). Ein erhöhtes Expressionsniveau von *CHRM3* ist in epileptogenen Hirngewebe von Patienten mit einer mesialer Temporallappenepilepsie ohne FS festgestellt worden (Bando et al., 2011). In Cortexproben von Patienten mit einer bipolaren affektiven Störung ist postmortem ein erniedrigtes Expressionsniveau von *CHRM3* festgestellt worden (Gibbons et al., 2009). Diese Studien unterstützen eine mögliche Rolle von *CHRM3* in der Epileptogenese, auch wenn sich in Mausmodellen keine Hinweise für eine Beteiligung von *Chrm3* in der Anfallsentstehung ergaben (Bymaster et al., 2003).

4.3 JME-Locus auf 2q32.1-q36.1

Der Locus für JME fand sich als signifikanter Kopplungsbefund in der Analyse von 118 JME-Multiplexfamilien in der chromosomalen Region 2q34 (max. $LOD_{NPL} = 3,43$ bei D2S143,

chr2:214973108). Die Kandidatenregion wurde anhand der suggestiven Kopplungshinweise auf den Bereich zwischen den Markern rs1569135 und rs2028383 eingegrenzt (2q32.1-q36.1, chr2:188115398-223823132). Der JME-Locus auf Chromosom 2 wird von einem signifikantem Kopplungsbefund in einer indischen Familie mit sieben JME-betroffenen Familienmitgliedern unterstützt (2q33.1-q36.2: Ratnapriya et al., 2010). In direkter Nähe zu dem maximalen LOD_{NPL} -Wert in der Kopplungsregion ist interessanterweise auch ein Locus für die Alzheimer-Krankheit mit frühem Beginn (Erkrankungsmanifestation: 50 bis 60J.) kartiert worden (2q34, $LOD_{NPL} = 3,20$ bei D2S2944, chr2:214579949, 31 Familien; Scott et al., 2003).

In direkter Nachbarschaft des signifikanten Kopplungssignals auf 2q34 befindet sich das Gen *ERBB4* (v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian); OMIM: *600543), ein hochrangiges Kandidatengen für IGE. Das Gen *ERBB4* kodiert für eine Rezeptor-Tyrosinkinase, die an der postsynaptischen Dichte angereichert ist (Huang et al., 2000). Die postsynaptische Dichte (postsynaptic density, PSD) ist eine Spezialisierung des Zytoskeletts an der synaptischen Verbindung, die für die Lokalisation der Transmitterzeptoren und die korrekte Signalweiterleitung erforderlich ist. Die Rezeptor-Tyrosinkinase *ERBB4* interagiert über PSD95, einer Hauptkomponente der PSD, mit dem NDMA-Rezeptor (Garcia et al., 2000). Der Ligand der Rezeptor-Tyrosinkinase *ERBB4* ist *NRG1* (neuregulin 1; OMIM: *142445). Die exzessive Aktivierung des *ERBB4*-Rezeptors durch *NRG1* ist für die verminderte Aktivität des NDMA-Rezeptors bei Schizophrenie verantwortlich (Hahn et al., 2006; Pitcher et al., 2011). Die *NRG1*-*ERBB4*-Signalkaskade ist auch an der Synaptogenese der inhibitorischen GABA-Neuronen beteiligt (Yau et al., 2003; Ting et al., 2011). „Knockout“ von *ErbB4* führt in der Maus zu spontanen Krampfanfällen (Tan et al., 2011). Ein um 40% erniedrigtes Expressionsniveau von *ERBB4* ist in epileptogenen Hirngewebe von Patienten mit einer Temporallappenepilepsie festgestellt worden (Li et al., 2011). Bei einem Patienten mit einer frühkindlichen myoklonischen Enzephalopathie ist der Bruch des Gens *ERBB4* durch eine Translokation berichtet worden (Backx et al., 2009).

Weitere interessante Kandidatengene innerhalb der Kopplungsregion auf Chromosom 2q32.1-q36.1 sind *CREB1*, *CDK5R2* und *SLC4A3*. Das Gen *CREB1* (cAMP responsive element binding protein 1; OMIM: *123810) kodiert für einen cAMP-abhängigen Transkriptionsfaktor, welcher die intrinsische und synaptische Plastizität von Nervenzellen steuert. Zusätzlich ist es relevant für das Überleben von Nervenzellen während der neuronalen Entwicklung. Sowohl ein „Knockout“ als auch eine induzierte konstitutive Aktivität von *Creb1* führen in Mausmodellen zur progressiver Neurodegeneration im Hippocampus und im Striatum (Mantamadiotis et al., 2002; Lopez de Armentia et al., 2007). Dabei führt die Überaktivierung

von *Creb1* dosisabhängig zu sporadischen epileptischen Krampfanfällen, unabhängig vom Zeitpunkt der Induktion. Die Inhibierung der *Creb1*-Aktivität führt dagegen zu einer reduzierten neuronalen Erregbarkeit (Jancic et al., 2009). Das Gen *CDK5R2* (cyclin-dependent kinase 5, regulatory subunit 2 (p39); OMIM: *603764) wird ausschließlich in Neuronen exprimiert (Tang et al., 1995) und weist eine synaptische Lokalisation auf (Humbert et al., 2000). *CDK5R2* steuert zusammen mit *CDK5R1* die Aktivität von *CDK5* (cyclin-dependent kinase 5) während der neuronalen Entwicklung (Wu et al., 2000; Ko et al., 2001). Das Gen *SLC4A3* (solute carrier family 4, anion exchanger, member 3; OMIM: *106195) kodiert für einen $\text{Cl}^-/\text{HCO}_3^-$ -Austauscher und reguliert die intra- und extrazellulären pH-Werte von Nervenzellen. Eine nicht-synonyme Sequenzvariation im *SLC4A3*-Gen (rs635311, c.2869A>C, p.A867D) wurde mit IGE assoziiert (Sander et al., 2002). Für diese Mutation ist eine Reduzierung der Austausch-Aktivität um 50% festgestellt worden (Vilas et al., 2009). „Knockout“ von *Slc4a3* führt in der Maus zu einer reduzierten Schwelle für induzierte Krampfanfälle (Hentschke et al., 2006). In der Studie von Ratnapriya et al. (2010) konnten keine weiteren Mutationen im *SLC4A3*-Gen festgestellt werden.

4.4 IAE-Locus auf 13q31.1-q32.1

Der Locus für IAE fand sich als signifikanter Kopplungsbefund in der Analyse von 235 IAE-Multiplexfamilien in der chromosomalen Region 13q31.3 (max. HLOD = 5,02 bei rs1332470, chr13:91417190, $\alpha = 0,22$). Die Kandidatenregion wurde anhand der suggestiven Kopplungshinweise auf den Bereich zwischen den Markern rs768218 und rs766606 eingegrenzt (13q31.1-q32.1, chr13:82259163-95654167).

Das aussichtsreichste Kandidatengen in der Kopplungsregion ist *GPC5* (glypican 5; OMIM: *602446), dessen Genprodukt als Modulator von Zell-Zell-Wechselwirkungen in die Neurogenese involviert ist (Luxardi et al., 2007). *GPC5* ist mit Multipler Sklerose (Baranzini et al., 2009; Lorentzen et al., 2010) und mit dem nephrotischen Syndrom (Okamoto et al., 2011) assoziiert worden. Ein weiteres Kandidatengen ist *SOX21* (SRY-related HMG-box gene 21; OMIM: *604974) dessen Genprodukt als neuronal exprimierter Transkriptionsfaktor bei der Neurogenese beteiligt ist (Sandberg et al., 2005). *SOX21* gehört zusammen mit *SOX14*, *SOX1*, *SOX2*, und *SOX3* zu der Gruppe B der *SOX*-Transkriptionsfaktoren (Kamachi et al., 2000). Die *SOX*-Transkriptionsfaktoren der Gruppe B weisen überlappende Zielgene auf. *SOX1*, *SOX2* und *SOX3* unterdrücken dabei als Transkriptionsaktivatoren die Neurogenese, während *SOX21* als Transkriptionsrepressor die neuronale Differenzierung fördert (Uchikawa et al., 1999; Bylund et al., 2003; Sandberg et al., 2005). „Knockout“ von *Sox1* führt in der

Maus zu einer Fehlbildung des Striatums und schweren epileptischen Anfällen bis hin zum Status epilepticus (Malas et al., 2003).

4.5 IGE-Locus auf 17q21.32

Der Locus für IGE fand sich als genomweit signifikante Assoziation auf 17q21.32 in der Meta-Analyse von 2826 IGE-Einzelfällen und 3701 Kontrollen europäischen Ursprungs (META-P = $7,90 \times 10^{-9}$ bei rs41336845; OR[T] = 0,76; 95%-KI: 0,70-0,84). Die Kandidatenregion wurde anhand der LD-Struktur auf den Bereich zwischen den SNPs rs2010952 und rs8072250 eingegrenzt (chr17:45892859-46569216).

Innerhalb der 676 kb großen Kandidatenregion befindet sich das Gen *PNPO* (pyridoxamine 5'-phosphate oxidase; OMIM: *603287), ein hochrangiges Kandidatengen für IGE. Das Gen *PNPO* kodiert für eine Oxidase, die Pyridoxin (Vitamin B₆) in die aktive Form überführt (Pyridoxal-5'-Phosphat, PLP). Homozygote Mutationen im *PNPO*-Gen führen zu einer neonatalen epileptischen Enzephalopathie (OMIM: #610090; Mills et al., 2005; Musayev et al., 2009). Die *PNPO*-abhängige neonatale epileptische Enzephalopathie ist eine Form der Pyridoxin-abhängigen Epilepsie, bei der die Anfälle nicht durch die Gabe von Pyridoxin kontrolliert werden können, sondern ausschließlich durch die Gabe von PLP. Der klinische Phänotyp zeichnet sich durch schwere, therapieresistente epileptische Anfälle aus, die sofort nach der Geburt auftreten und aus denen sich häufig ein Status epilepticus entwickelt. Ohne PLP-Therapie überleben die betroffenen Säuglinge nicht die Neugeborenenphase. Der epileptogene Mechanismus des PLP-Mangels wird in einer reduzierten Umsetzung des exzitatorischen Neurotransmitters L-Glutamat zu dem inhibitorischen Neurotransmitter GABA vermutet (Gospe et al., 1994). Im Gegensatz zu dem schweren Phänotyp verursacht durch einen vollständigen Verlust der Oxidase-Aktivität, könnten moderate Veränderungen des Pyridoxin-Stoffwechselwegs durch Sequenzvariationen in dem *PNPO*-Gen die Anfallsschwelle reduzieren und damit für IGE prädisponieren.

Ein weiteres Kandidatengen ist *CDK5RAP3* (CDK5 regulatory subunit associated protein 3; OMIM: *608202). Das Gen *CDK5RAP3* kodiert für einen Bindungspartner von CDK5R1 (cyclin-dependent kinase 5, regulatory subunit 1 (p35); OMIM: *603460) (Wang et al., 2000; Ching et al., 2000). Das von dem Gen *CDK5R1* kodierte Protein p35 ist eine aktivierende Untereinheit von CDK5 (cyclin-dependent kinase 5) (Ko et al., 2001). Eine Deregulation von CDK5 kann über eine Hippokampussklerose (HS; Sen et al., 2006; Tobias et al., 2009) oder eine fokale kortikale Dysplasie (FCD; Sisodiya et al., 2002) zu der Entstehung von fokalen Formen von Epilepsie führen. „Knockout“ von *Cdk5r1* in der Maus führt zu einer massiven

Fehlbildung der kortikalen Entwicklung und spontanen epileptischen Anfällen (Chae et al., 1997).

4.6 Ausblick

Die drei signifikanten positionellen Befunde dieser Studie in den Chromosomenabschnitten 2q32.1-q36.1, 13q31.1-q32.1 und 17q21.32 stellen eine Grundlage für die Identifizierung neuer Epilepsiegene dar. In einem ersten Schritt sollten die Kandidatenregionen vollständig sequenziert werden. Durch den Einsatz von hochparallelen Sequenzierungstechnologien („Next Generation Sequencing“, NGS) ist es möglich die Suche nach Mutationen auch auf die nicht kodierenden Bereiche auszuweiten. Damit sollte sichergestellt werden, dass Varianten in regulatorischen Elementen, die in großer Entfernung zu ihren Zielgenen lokalisiert sein können (Jeong et al., 2006), nicht übersehen werden. Zudem bietet die NGS-Technologie die Möglichkeit zur systematischen Erfassung von Genkopiezahlvarianten („copy number variations“, CNVs) sowie DNA-Methylierungsmustern.

Die erhobenen Affymetrix SNP 6.0-Genotypdaten der IGE-GWAS können auch als Grundlage für internationale GWAS-Meta-Analysen dienen. Erst in einer ausreichend großen Kohorte wird es möglich sein, die Mehrheit der häufigen genetischen Varianten zu identifizieren, die für Epilepsie prädisponieren.

Eine weitere aussichtsreiche „in silico“-Analyse der erhobenen GWAS-Daten ist die Ausweitung der Assoziationsanalyse zu einer genomweiten Interaktionsanalyse (GWIA) aller SNP-Paaren. Die Interaktion zwischen genetischen Varianten (Epistase) könnte einen weiteren Teil der Heritabilität von Epilepsie aufdecken. Eine GWIA ist als Machbarkeitsstudie bereits erfolgreich mit einer Teilmenge unserer verfügbaren IGE-Kohorte durchgeführt worden (Steffens et al., 2010).

Die CNV-Informationen der Affymetrix SNP 6.0-Daten können auch zur Erweiterung früherer Studien dienen. Die bisher in unserem Datensatz aufgefundenen Deletionen auf den Chromosomen 15q11.2, 15q13.3 und 16p13.11 (Helbig et al., 2009; de Kovel et al., 2010) stellen die derzeit bedeutendsten genetischen Faktoren dar, die bei den häufigen IGE-Syndromen identifiziert worden sind (Mulley und Mefford, 2011).

Inwiefern Therapieansätze für eine breite Mehrheit der Patienten aus den molekulargenetischen Erkenntnissen resultieren werden, bleibt allerdings weiterhin unklar. Die Identifizierung von neuen Epilepsiegenen wird jedoch unabhängig von der klinischen Relevanz weitere Einblicke in epileptogene Mechanismen erlauben und zur Klärung der genetischen Architektur von Epilepsien beitragen.

5 ANHANG

5.1 Kandidatenregionen aus Kopplungsstudien

Locus	Region	OMIM	Syndrom	Referenz
BCECTS	11p13	%117100	BCECTS	Strug et al., 2009
—	15q13	%117100	BCECTS	Neubauer et al., 1998
ENFL2	15q25.1-q25.3	%603204	ADNFLE	Phillips et al., 1998
BFIS4	1p36.12-p35.1	%612627	BFIS	Li et al., 2008
BFIS2	16p12-q12.2	#605751	BFIS	Szepetowski et al., 1997
BFIS1	19q12-q13.1	%601764	BFIS	Guipponi et al., 1997
ETL4	9q21-q22	%611631	EPOLM	Deprez et al., 2007
FPEVF	22q11-q12	%604364	FPEVF	Callenbach et al., 2003
ETL2	12q22-q23.3	%608096	FTLE	Claes et al., 2004
ICCA	16p12.1-q12.1	#602066	ICCA	Szepetowski et al., 1997
ETL3	4q13.2-q21.3	%611630	MFTLE	Hedera et al., 2007
PEPS	4p16.1-p15.32	%607221	PEPS	Kinton et al., 2002
FAME2	2p11.1-q12.2	%607876	BAFME	Striano et al., 2004
FAME3	5p15.31-p15	%613608	BAFME	Depienne et al., 2010
FAME1	8q23.3-q24.11	%601068	BAFME	Mikami et al., 1999
EIG11	3q26	#607628	IGE	Sander et al., 2000
—	8p12	—	IGE	Durner et al., 2001
EIG1	8q24	%600669	IGE	Zara et al., 1995
EIG3	9q32-q33.3	%608762	IGE	Baykan et al., 2004
EIG5	10p11	%611934	IGE	Kinirons et al., 2008b
EIG4	10q25-q26	%609750	IGE	Puranam et al., 2005
EIG2	14q23	%606972	IGE	Sander et al., 2000
EIG7	15q13.3	%604827	IGE	Helbig et al., 2009
PPR2	13q31.3	%609572	IGE mit PPR	Tauer et al., 2005
EJM9	2q33-q36.2	%614280	JME	Ratnapriya et al., 2010
EJM4	5q12-q14	%611364	JME	Kapoor et al., 2007
EJM3	6q21	%608816	JME	Greenberg et al., 2000
EJM2	15q13.2-q14	%604827	JME	Elmslie et al., 1997
ECA1	8q24	%600131	CAE	Fong et al., 1998
GEFSP4	2p24.2-p24.1	%609800	GEFS+	Audenaert et al., 2005
GEFSP8	6q16.3-q22.31	%613828	GEFS+	Poduri et al., 2009
GEFSP6	8p23.1-p21	%612279	GEFS+	Baulac et al., 2008
FEB9	3p24.2-p23	%611634	FS/CAE	Nabbout et al., 2007
FEB10	3q26.2-q26.33	%612637	FS/GEFS+	Dai et al., 2008
FEB4	5q14-q15	#604352	FS/GEFS+	Nakayama et al., 2000
FEB7	21q22.11-q22.13	%611515	FS/GEFS+	Hedera et al., 2006
FEB1	8q13-q21	%121210	FS	Wallace et al., 1996
FEB6	18p11.2	%609253	FS	Nakayama et al., 2004
FEB2	19p13.3-p13.2	%602477	FS	Johnson et al., 1998
FEB5	6q22-q24	%609255	FS alleine	Nabbout et al., 2002

Tabelle 5-1 Kandidatenregionen aus Kopplungsstudien zu Epilepsie

Die Liste ist sortiert nach dem Syndrom und der chromosomalen Position. Als Legende für die jeweiligen Syndrome dient Tabelle 1-1. Legende -> #-OMIM-ID: krankheitsassoziierte Mutationen sind beschrieben; %-OMIM-ID: keine Mutationen bekannt.

5.2 Qualitätsstatistiken von 21 SNPstream-SNPs

SNP	CHR	POS NCBI 37.3	CR F/K	MAF F/K	HWE-P F/K	Mendel- Fehler	Konkordanz	
							Affy 6.0	SNPstream
rs12059546	1	239970097	0,98/0,97	0,21/0,18	0,01/0,25	0	0,9974	1
rs1110615	1	239974170	0,99/0,99	0,20/0,18	0,12/0,12	0,0018	1	1
rs388556	2	13392540	1,00/1,00	0,32/0,31	0,95/0,10	0	0,9905	1
rs11885093	2	57941185	0,51/0,39	0,42/0,45	0,20/0,85	0	0,9971	NA
rs11684106	2	58075701	NA	NA	NA	NA	NA	NA
rs2953442	2	58083420	0,99/1,00	0,35/0,31	0,58/0,51	0	1	1
rs2717040	2	58169494	NA	NA	NA	NA	NA	NA
rs17741930	2	145362155	0,99/0,99	0,15/0,17	0,28/0,61	0,0018	1	1
rs1035837	2	145379270	0,81/0,72	0,16/0,18	0,13/0,91	0,0069	0,9444	1
rs289034	5	114122165	0,99/0,99	0,41/0,41	0,07/0,60	0	1	1
rs1110146	9	140777370	0,95/0,97	0,28/0,29	0,03/0,89	0,0036	0,9804	0,9828
rs2268330	14	23795976	1,00/1,00	0,26/0,27	0,09/0,77	0	1	1
rs1447579	14	42125681	0,97/0,97	0,30/0,32	0,07/0,89	0,0123	0,9700	0,9091
rs12904369	15	97335674	0,99/0,99	0,24/0,25	1,00/0,41	0	1	1
rs11650615	17	46123698	1,00/1,00	0,24/0,27	0,88/0,47	0	0,9974	1
rs41336845	17	46215110	0,99/0,99	0,15/0,19	0,83/0,85	0,0018	1	1
rs16953235	17	46244948	1,00/0,99	0,15/0,19	0,83/0,85	0	0,9974	1
rs7236787	18	25800169	1,00/1,00	0,30/0,30	0,55/0,69	0	1	1
rs10853686	18	25870524	1,00/1,00	0,39/0,40	1,00/0,95	0	1	1
rs12963674	18	55978595	NA	NA	NA	NA	NA	NA
rs4892034	18	70399988	0,99/0,99	0,18/0,17	0,04/0,36	0,0018	1	1

Tabelle 5-2 Genotypisierte SNPs aus der Replikationsanalyse

Legende -> CHR: Chromosom; POS NCBI 37.3: physikalische Position nach der NCBI Version 37.3; CR, MAF, HWE-P F/K: CR, MAF und HWE-P-Werte jeweils getrennt nach Fälle und Kontrollen; Konkordanz Affy 6.0: Konkordanz zu *Affymetrix Genome Wide Human SNP Array 6.0*-Genotypen von 383 Personen; Konkordanz SNPstream: Konkordanz von 79 duplizierten Personen in der Replikationsstufe. Sieben ausgefallene SNPs sowie die jeweilige Grundlage für den Ausschluß sind rot beschriftet.

6 LITERATURVERZEICHNIS

Abecasis, G.R., Cherny, S.S., Cookson, W.O., and Cardon, L.R. (2001). GRR: graphical representation of relationship errors. *Bioinformatics* 17, 742-743.

Abecasis, G.R., Cherny, S.S., Cookson, W.O., and Cardon, L.R. (2002). Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30, 97-101.

Abreu, P.C., Hodge, S.E., and Greenberg, D.A. (2002). Quantification of type I error probabilities for heterogeneity LOD scores. *Genet Epidemiol* 22, 156-169.

Affymetrix (2008). Quality Control Assessment in Genotyping Console™. Affymetrix White Paper Revision 1.

Affymetrix (2009). Genome-Wide Human SNP Array 6.0. Affymetrix P/N 702509 Rev. 2.

Anderson, C.A., Soranzo, N., Zeggini, E., and Barrett, J.C. (2011). Synthetic associations are unlikely to account for many common disease genome-wide association signals. *PLoS Biol* 9, e1000580.

Antonacci, F., Kidd, J.M., Marques-Bonet, T., Ventura, M., Siswara, P., Jiang, Z., and Eichler, E.E. (2009). Characterization of six human disease-associated inversion polymorphisms. *Hum Mol Genet* 18, 2555-2566.

Aridon, P., Marini, C., Di Resta, C., Brilli, E., De Fusco, M., Politi, F., Parrini, E., Manfredi, I., Pisano, T., Pruna, D., et al. (2006). Increased sensitivity of the neuronal nicotinic receptor alpha 2 subunit causes familial epilepsy with nocturnal wandering and ictal fear. *Am J Hum Genet* 79, 342-350.

Audenaert, D., Claes, L., Claeys, K.G., Deprez, L., Van Dyck, T., Goossens, D., Del-Favero, J., Van Paesschen, W., Van Broeckhoven, C., and De Jonghe, P. (2005). A novel susceptibility locus at 2p24 for generalised epilepsy with febrile seizures plus. *J Med Genet* 42, 947-952.

Audenaert, D., Schwartz, E., Claeys, K.G., Claes, L., Deprez, L., Suls, A., Van Dyck, T., Lagae, L., Van Broeckhoven, C., Macdonald, R.L., et al. (2006). A novel GABRG2 mutation associated with febrile seizures. *Neurology* 67, 687-690.

Backx, L., Ceulemans, B., Vermeesch, J.R., Devriendt, K., and Van Esch, H. (2009). Early myoclonic encephalopathy caused by a disruption of the neuregulin-1 receptor ErbB4. *Eur J Hum Genet* 17, 378-382.

Bando, S.Y., Alegro, M.C., Amaro, E., Jr., Silva, A.V., Castro, L.H., Wen, H.T., Lima Lde, 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.

Bansal, V., Bashir, A., and Bafna, V. (2007). Evidence for large inversion polymorphisms in the human genome from HapMap data. *Genome Res* 17, 219-230.

Baranzini, S.E., Wang, J., Gibson, R.A., Galwey, N., Naegelin, Y., Barkhof, F., Radue, E.W., Lindberg, R.L., Uitdehaag, B.M., Johnson, M.R., et al. (2009). Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 18, 767-778.

Baulac, S., Gourfinkel-An, I., Couarch, P., Depienne, C., Kaminska, A., Dulac, O., Baulac, M., LeGuern, E., and Nabbout, R. (2008). A novel locus for generalized epilepsy with febrile seizures plus in French families. *Arch Neurol* 65, 943-951.

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.

- Baykan, B., Madia, F., Bebek, N., Gianotti, S., Guney, A.I., Cine, N., Bianchi, A., Gokyigit, A., and Zara, F. (2004). Autosomal recessive idiopathic epilepsy in an inbred family from Turkey: identification of a putative locus on chromosome 9q32-33. *Epilepsia* 45, 479-487.
- Beck-Mannagetta, G., and Janz, D. (1991). Syndrome-related genetics in generalized epilepsy. *Epilepsy Res Suppl* 4, 105-111.
- Becker, T., Flaquer, A., Brockschmidt, F.F., Herold, C., and Steffens, M. (2009). Evaluation of potential power gain with imputed genotypes in genome-wide association studies. *Hum Hered* 68, 23-34.
- Bell, P.A., Chaturvedi, S., Gelfand, C.A., Huang, C.Y., Kochersperger, M., Kopla, R., Modica, F., Pohl, M., Varde, S., Zhao, R., et al. (2002). SNPstream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery. *Biotechniques Suppl*, 70-72, 74, 76-77.
- 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.
- Berg, A.T., and Scheffer, I.E. (2011). New concepts in classification of the epilepsies: entering the 21st century. *Epilepsia* 52, 1058-1062.
- 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., Heron, S.E., Giordano, L., Marini, C., Guerrini, R., Kaplan, R.E., Gambardella, A., Steinlein, O.K., Grinton, B.E., Dean, J.T., et al. (2004a). Benign familial neonatal-infantile seizures: characterization of a new sodium channelopathy. *Ann Neurol* 55, 550-557.
- Berkovic, S.F., Izzillo, P., McMahon, J.M., Harkin, L.A., McIntosh, A.M., Phillips, H.A., Briellmann, R.S., Wallace, R.H., Mazarib, A., Neufeld, M.Y., et al. (2004b). LGI1 mutations in temporal lobe epilepsies. *Neurology* 62, 1115-1119.
- Bianchi, A., Viaggi, S., and Chioffi, E. (2003). Family study of epilepsy in first degree relatives: data from the Italian Episcree Study. *Seizure* 12, 203-210.
- Bickeböller, H., and Fischer, C. (2007). Einführung in die Genetische Epidemiologie. 1. Auflage. Springer Verlag Berlin Heidelberg.
- Blake, J.A., Richardson, J.E., Davisson, M.T., and Eppig, J.T. (1997). The Mouse Genome Database (MGD). A comprehensive public resource of genetic, phenotypic and genomic data. The Mouse Genome Informatics Group. *Nucleic Acids Res* 25, 85-91.
- Bland, J.M., and Altman, D.G. (1995). Multiple significance tests: the Bonferroni method. *BMJ* 310, 170.
- Bryc, K., Velez, C., Karafet, T., Moreno-Estrada, A., Reynolds, A., Auton, A., Hammer, M., Bustamante, C.D., and Ostrer, H. (2010). Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc Natl Acad Sci U S A* 107 Suppl 2, 8954-8961.
- Bylund, M., Andersson, E., Novitsch, B.G., and Muhr, J. (2003). Vertebrate neurogenesis is counteracted by Sox1-3 activity. *Nat Neurosci* 6, 1162-1168.

- 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. *Eur J Neurosci* 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.
- Callenbach, P.M., van den Maagdenberg, A.M., Hottenga, J.J., van den Boogerd, E.H., de Coo, R.F., Lindhout, D., Frants, R.R., Sandkuijl, L.A., and Brouwer, O.F. (2003). Familial partial epilepsy with variable foci in a Dutch family: clinical characteristics and confirmation of linkage to chromosome 22q. *Epilepsia* 44, 1298-1305.
- 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* 50(6), 1645-1656.
- Carranza Rojo, D., Hamiwka, L., McMahon, J.M., Dibbens, L.M., Arsov, T., Suls, A., Stodberg, T., Kelley, K., Wirrell, E., Appleton, B., et al. (2011). De novo SCN1A mutations in migrating partial seizures of infancy. *Neurology* 77, 380-383.
- Cavalleri, G.L., Weale, M.E., Shianna, K.V., Singh, R., Lynch, J.M., Grinton, B., Szoeker, C., Murphy, K., Kinirons, P., O'Rourke, D., et al. (2007a). Multicentre search for genetic susceptibility loci in sporadic epilepsy syndrome and seizure types: a case-control study. *Lancet Neurol* 6, 970-980.
- Cavalleri, G.L., Walley, N.M., Soranzo, N., Mulley, J., Doherty, C.P., Kapoor, A., Depondt, C., Lynch, J.M., Scheffer, I.E., Heils, A., et al. (2007b). A multicenter study of BRD2 as a risk factor for juvenile myoclonic epilepsy. *Epilepsia* 48, 706-712.
- Chabchoub, E., Vermeesch, J.R., de Ravel, T., de Cock, P., and Fryns, J.P. (2008). The facial dysmorphism in the newly recognised microdeletion 2p15-p16.1 refined to a 570 kb region in 2p15. *J Med Genet* 45, 189-192.
- Chae, T., Kwon, Y.T., Bronson, R., Dikkes, P., Li, E., and Tsai, L.H. (1997). Mice lacking p35, a neuronal specific activator of Cdk5, display cortical lamination defects, seizures, and adult lethality. *Neuron* 18, 29-42.
- Chai, H.S., Therneau, T.M., Bailey, K.R., and Kocher, J.P. (2010). Spatial normalization improves the quality of genotype calling for Affymetrix SNP 6.0 arrays. *BMC Bioinformatics* 11, 356.
- 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.
- Charlier, C., Singh, N.A., Ryan, S.G., Lewis, T.B., Reus, B.E., Leach, R.J., and Leppert, M. (1998). A pore mutation in a novel KQT-like potassium channel gene in an idiopathic epilepsy family. *Nat Genet* 18, 53-55.
- Chen, Y., Lu, J., Pan, H., Zhang, Y., Wu, H., Xu, K., Liu, X., Jiang, Y., Bao, X., Yao, Z., et al. (2003). Association between genetic variation of CACNA1H and childhood absence epilepsy. *Ann Neurol* 54, 239-243.
- Ching, Y.P., Qi, Z., and Wang, J.H. (2000). Cloning of three novel neuronal Cdk5 activator binding proteins. *Gene* 242, 285-294.
- Chioza, B., Osei-Lah, A., Wilkie, H., Nashef, L., McCormick, D., Asherson, P., and Makoff, A.J. (2002). Suggestive evidence for association of two potassium channel genes with different idiopathic generalised epilepsy syndromes. *Epilepsy Res* 52, 107-116.

- Chokkalingam, A.P., Aldrich, M.C., Bartley, K., Hsu, L., Metayer, C., Barcellos, L.F., Wiemels, J.L., Wiencke, J.K., Buffler, P.A., and Selvin, S. (2011). Matching on Race and Ethnicity in Case-Control Studies as a Means of Control for Population Stratification. *Epidemiol* 1:101.
- Claes, L., Audenaert, D., Deprez, L., Van Paesschen, W., Depondt, C., Goossens, D., Del-Favero, J., Van Broeckhoven, C., and De Jonghe, P. (2004). Novel locus on chromosome 12q22-q23.3 responsible for familial temporal lobe epilepsy associated with febrile seizures. *J Med Genet* 41, 710-714.
- 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.
- Crossett, A., Kent, B.P., Klei, L., Ringquist, S., Trucco, M., Roeder, K., and Devlin, B. (2010). Using ancestry matching to combine family-based and unrelated samples for genome-wide association studies. *Stat Med* 29, 2932-2945.
- Crunelli, V., and Leresche, N. (2002). Childhood absence epilepsy: genes, channels, neurons and networks. *Nat Rev Neurosci* 3, 371-382.
- Curatolo, P., Moavero, R., Lo Castro, A., and Cerminara, C. (2009). Pharmacotherapy of idiopathic generalized epilepsies. *Expert Opin Pharmacother* 10, 5-17.
- Dai, X.H., Chen, W.W., Wang, X., Zhu, Q.H., Li, C., Li, L., Liu, M.G., Wang, Q.K., and Liu, J.Y. (2008). A novel genetic locus for familial febrile seizures and epilepsy on chromosome 3q26.2-q26.33. *Hum Genet* 124, 423-429.
- de Andrade, M., Atkinson, E.J., Bamlet, W.R., Matsumoto, M.E., Maharjan, S., Slager, S.L., Vachon, C.M., Cunningham, J.M., and Kardia, S.L. (2011). Evaluating the influence of quality control decisions and software algorithms on SNP calling for the affymetrix 6.0 SNP array platform. *Hum Hered* 71, 221-233.
- de Bakker, P.I., Ferreira, M.A., Jia, X., Neale, B.M., Raychaudhuri, S., and Voight, B.F. (2008). Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum Mol Genet* 17, R122-128.
- de Kovel, C.G., 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.
- de Nijs, L., Leon, 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.
- Deng, L., Zhang, Y., Kang, J., Liu, T., Zhao, H., Gao, Y., Li, C., Pan, H., Tang, X., Wang, D., et al. (2008). An unusual haplotype structure on human chromosome 8p23 derived from the inversion polymorphism. *Hum Mutat* 29, 1209-1216.
- Depienne, C., Magnin, E., Bouteiller, D., Stevanin, G., Saint-Martin, C., Vidailhet, M., Apartis, E., Hirsch, E., LeGuern, E., Labauge, P., et al. (2010). Familial cortical myoclonic tremor with epilepsy: the third locus (FCMTE3) maps to 5p. *Neurology* 74, 2000-2003.
- Deprez, L., Peeters, K., Van Paesschen, W., Claeys, K.G., Claes, L.R., Suls, A., Audenaert, D., Van Dyck, T., Goossens, D., Del-Favero, J., et al. (2007). Familial occipitotemporal lobe epilepsy and migraine with visual aura: linkage to chromosome 9q. *Neurology* 68, 1995-2002.
- 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., Feng, H.J., Richards, M.C., Harkin, L.A., Hodgson, B.L., Scott, D., Jenkins, M., Petrou, S., Sutherland, G.R., Scheffer, I.E., et al. (2004). GABRD encoding a protein for extra- or peri-synaptic GABAA receptors is a susceptibility locus for generalized epilepsies. *Hum Mol Genet* 13, 1315-1319.
- Dibbens, L.M., Ekberg, J., Taylor, I., Hodgson, B.L., Conroy, S.J., Lensink, I.L., Kumar, S., Zielinski, M.A., Harkin, L.A., Sutherland, G.R., et al. (2007a). NEDD4-2 as a potential candidate susceptibility gene for epileptic photosensitivity. *Genes Brain Behav* 6, 750-755.
- Dibbens, L.M., Heron, S.E., and Mulley, J.C. (2007b). A polygenic heterogeneity model for common epilepsies with complex genetics. *Genes Brain Behav* 6, 593-597.
- Douglas, J.A., Boehnke, M., and Lange, K. (2000). A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am J Hum Genet* 66, 1287-1297.
- 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.
- Elmslie, F.V., Rees, M., Williamson, M.P., Kerr, M., Kjeldsen, M.J., Pang, K.A., Sundqvist, A., Friis, M.L., Chadwick, D., Richens, A., et al. (1997). Genetic mapping of a major susceptibility locus for juvenile myoclonic epilepsy on chromosome 15q. *Hum Mol Genet* 6, 1329-1334.
- 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.
- Engel, J., Jr. (2011). The etiologic classification of epilepsy. *Epilepsia* 52, 1195-1197; discussion 1205-1199.
- Eppig, J.T., Blake, J.A., Bult, C.J., Kadin, J.A., and Richardson, J.E. (2012). The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. *Nucleic Acids Res* 40, D881-886.
- Escayg, A., De Waard, M., Lee, D.D., Bichet, D., Wolf, P., Mayer, T., Johnston, J., Baloh, R., Sander, T., and Meisler, M.H. (2000a). 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.
- Escayg, A., MacDonald, B.T., Meisler, M.H., Baulac, S., Huberfeld, G., An-Gourfinkel, I., Brice, A., LeGuern, E., Moulard, B., Chaigne, D., et al. (2000b). Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2. *Nat Genet* 24, 343-345.
- Everett, K.V., Chioza, B., Aicardi, J., Aschauer, H., Brouwer, O., Callenbach, P., Covanis, A., Dulac, O., Eeg-Olofsson, O., Feucht, M., et al. (2007). Linkage and association analysis of CACNG3 in childhood absence epilepsy. *Eur J Hum Genet* 15, 463-472.
- Ewens, W.J., and Spielman, R.S. (1995). The transmission/disequilibrium test: history, subdivision, and admixture. *Am J Hum Genet* 57, 455-464.
- Fellay, J., Shianna, K.V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., Zhang, K., Gumbs, C., Castagna, A., Cossarizza, A., et al. (2007). A whole-genome association study of major determinants for host control of HIV-1. *Science* 317, 944-947.
- Ferrie, C.D. (2010). Terminology and organization of seizures and epilepsies: radical changes not justified by new evidence. *Epilepsia* 51, 713-714.
- Fong, G.C., Shah, P.U., Gee, M.N., Serratosa, J.M., Castroviejo, I.P., Khan, S., Ravat, S.H., Mani, J., Huang, Y., Zhao, H.Z., et al. (1998). Childhood absence epilepsy with tonic-clonic seizures and elec-

- troencephalogram 3-4-Hz spike and multispikes-slow wave complexes: linkage to chromosome 8q24. *Am J Hum Genet* 63, 1117-1129.
- Freedman, M.L., Reich, D., Penney, K.L., McDonald, G.J., Mignault, A.A., Patterson, N., Gabriel, S.B., Topol, E.J., Smoller, J.W., Pato, C.N., et al. (2004). Assessing the impact of population stratification on genetic association studies. *Nat Genet* 36, 388-393.
- Freitag, C.M., May, T.W., Pfafflin, M., Konig, 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.
- Garavelli, L., and Mainardi, P.C. (2007). Mowat-Wilson syndrome. *Orphanet J Rare Dis* 2, 42.
- Garavelli, L., Zollino, M., Mainardi, P.C., Gurreri, 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.
- Garcia, R.A., Vasudevan, K., and Buonanno, A. (2000). The neuregulin receptor ErbB-4 interacts with PDZ-containing proteins at neuronal synapses. *Proc Natl Acad Sci U S A* 97, 3596-3601.
- Gastaut, H. (1970). Clinical and electroencephalographical classification of epileptic seizures. *Epilepsia* 11, 102-113.
- Gibbons, A.S., Scarr, E., McLean, C., Sundram, S., and Dean, B. (2009). Decreased muscarinic receptor binding in the frontal cortex of bipolar disorder and major depressive disorder subjects. *J Affect Disord* 116, 184-191.
- Glasscock, E., Qian, J., Yoo, J.W., and Noebels, J.L. (2007). Masking epilepsy by combining two epilepsy genes. *Nat Neurosci* 10, 1554-1558.
- Goodman, S. (2008). A dirty dozen: twelve p-value misconceptions. *Semin Hematol* 45, 135-140.
- Gospe, S.M., Jr., Olin, K.L., and Keen, C.L. (1994). Reduced GABA synthesis in pyridoxine-dependent seizures. *Lancet* 343, 1133-1134.
- 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.
- Greenberg, D.A., Cayanis, E., Strug, L., Marathe, S., Durner, M., Pal, D.K., Alvin, G.B., Klotz, I., Dickler, E., Shinnar, S., et al. (2005). Malic enzyme 2 may underlie susceptibility to adolescent-onset idiopathic generalized epilepsy. *Am J Hum Genet* 76, 139-146.
- Greenberg, D.A., and Subaran, R. (2011). Blindness, phenotype, and fashionable genetic analysis: a critical examination of the current state of epilepsy genetic studies. *Epilepsia* 52, 1-9.
- Gudbjartsson, D.F., Thorvaldsson, T., Kong, A., Gunnarsson, G., and Ingólfssdóttir, A. (2005). Allegro version 2. *Nat Genet* 37, 1015-1016.
- Guipponi, M., Rivier, F., Vigeveno, F., Beck, C., Crespel, A., Echenne, B., Lucchini, P., Sebastianelli, R., Baldy-Moulinier, M., and Malafosse, A. (1997). Linkage mapping of benign familial infantile convulsions (BFIC) to chromosome 19q. *Hum Mol Genet* 6, 473-477.
- Gunderson, K.L., Kruglyak, S., Graige, M.S., Garcia, F., Kermani, B.G., Zhao, C., Che, D., Dickinson, T., Wickham, E., Bierle, J., et al. (2004). Decoding randomly ordered DNA arrays. *Genome Res* 14, 870-877.

- Guo, C.Y., Lunetta, K.L., DeStefano, A.L., and Cupples, L.A. (2009). Combined haplotype relative risk (CHRR): a general and simple genetic association test that combines trios and unrelated case-controls. *Genet Epidemiol* 33, 54-62.
- Guo, Y., Baum, L.W., Sham, P.C., Wong, V., Ng, P.W., Lui, C.H., Sin, N.C., Tsoi, T.H., Tang, C.S., Kwan, J.S., et al. (2011). 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 Hadera, P. (2007). New ideas in epilepsy genetics: novel epilepsy genes, copy number alterations, and gene regulation. *Arch Neurol* 64, 324-328.
- Hahn, C.G., Wang, H.Y., Cho, D.S., Talbot, K., Gur, R.E., Berrettini, W.H., Bakshi, K., Kamins, J., Borgmann-Winter, K.E., Siegel, S.J., et al. (2006). Altered neuregulin 1-erbB4 signaling contributes to NMDA receptor hypofunction in schizophrenia. *Nat Med* 12, 824-828.
- Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., and McKusick, V.A. (2005). Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* 33, D514-517.
- Hardy, G.H. (1908). Mendelian Proportions in a Mixed Population. *Science* 28, 49-50.
- 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.
- 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.
- Hadera, P., Ma, S., Blair, M.A., Taylor, K.A., Hamati, A., Bradford, Y., Abou-Khalil, B., and Haines, J.L. (2006). Identification of a novel locus for febrile seizures and epilepsy on chromosome 21q22. *Epilepsia* 47, 1622-1628.
- Hadera, P., Blair, M.A., Andermann, E., Andermann, F., D'Agostino, D., Taylor, K.A., Chahine, L., Pandolfo, M., Bradford, Y., Haines, J.L., et al. (2007). Familial mesial temporal lobe epilepsy maps to chromosome 4q13.2-q21.3. *Neurology* 68, 2107-2112.
- 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.
- Hentschke, M., Wiemann, M., Hentschke, S., Kurth, I., Hermans-Borgmeyer, I., Seidenbecher, T., Jentsch, T.J., Gal, A., and Hubner, C.A. (2006). Mice with a targeted disruption of the Cl⁻/HCO₃⁻ exchanger AE3 display a reduced seizure threshold. *Mol Cell Biol* 26, 182-191.
- Heron, S.E., Khosravani, H., Varela, D., Bladen, C., Williams, T.C., Newman, M.R., Scheffer, I.E., Berkovic, S.F., Mulley, J.C., and Zamponi, G.W. (2007). Extended spectrum of idiopathic generalized epilepsies associated with CACNA1H functional variants. *Ann Neurol* 62, 560-568.
- Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., and Manolio, T.A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 106, 9362-9367.

- Hodge, S.E., Vieland, V.J., and Greenberg, D.A. (2002). HLODs remain powerful tools for detection of linkage in the presence of genetic heterogeneity. *Am J Hum Genet* 70, 556-559.
- Holliday, E., Mowry, B., Chant, D., and Nyholt, D. (2005). The importance of modelling heterogeneity in complex disease: application to NIMH Schizophrenia Genetics Initiative data. *Hum Genet* 117, 160-167.
- 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.
- Huang, Y.Z., Won, S., Ali, D.W., Wang, Q., Tanowitz, M., Du, Q.S., Pelkey, K.A., Yang, D.J., Xiong, W.C., Salter, M.W., et al. (2000). Regulation of neuregulin signaling by PSD-95 interacting with ErbB4 at CNS synapses. *Neuron* 26, 443-455.
- Humbert, S., Lanier, L.M., and Tsai, L.H. (2000). Synaptic localization of p39, a neuronal activator of cdk5. *Neuroreport* 11, 2213-2216.
- 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.
- Illumina (2006). Infinium II Assay Workflow. Illumina Pub. No. 370-2006-027.
- Illumina (2010). Infinium Genotyping Data Analysis. Illumina Pub. No. 970-2007-005.
- Jallon, P., Loiseau, P., and Loiseau, J. (2001). Newly diagnosed unprovoked epileptic seizures: presentation at diagnosis in CAROLE study. Coordination Active du Reseau Observatoire Longitudinal de l' Epilepsie. *Epilepsia* 42, 464-475.
- Jancic, D., Lopez de Armentia, M., Valor, L.M., Olivares, R., and Barco, A. (2009). Inhibition of cAMP response element-binding protein reduces neuronal excitability and plasticity, and triggers neurodegeneration. *Cereb Cortex* 19, 2535-2547.
- Jansen, F.E., Sadleir, L.G., Harkin, L.A., Vadlamudi, L., McMahon, J.M., Mulley, J.C., Scheffer, I.E., and Berkovic, S.F. (2006). Severe myoclonic epilepsy of infancy (Dravet syndrome): recognition and diagnosis in adults. *Neurology* 67, 2224-2226.
- Janz, D. (1997). The idiopathic generalized epilepsies of adolescence with childhood and juvenile age of onset. *Epilepsia* 38, 4-11.
- Jeffreys, A.J., Kauppi, L., and Neumann, R. (2001). Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet* 29, 217-222.
- Jensen, F.E. (2011). Epilepsy as a spectrum disorder: Implications from novel clinical and basic neuroscience. *Epilepsia* 52 Suppl 1, 1-6.
- Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M., and Epstein, D.J. (2006). A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development* 133, 761-772.
- Jin, Y., Birlea, S.A., Fain, P.R., Mailloux, C.M., Riccardi, S.L., Gowan, K., Holland, P.J., Bennett, D.C., Wallace, M.R., McCormack, W.T., et al. (2010). Common variants in FOXP1 are associated with generalized vitiligo. *Nat Genet* 42, 576-578.

- Johnson, E.W., Dubovsky, J., Rich, S.S., O'Donovan, C.A., Orr, H.T., Anderson, V.E., Gil-Nagel, A., Ahmann, P., Dokken, C.G., Schneider, D.T., et al. (1998). Evidence for a novel gene for familial febrile convulsions, FEB2, linked to chromosome 19p in an extended family from the Midwest. *Hum Mol Genet* 7, 63-67.
- 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.
- 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; discussion 193-196.
- 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.
- Kamachi, Y., Uchikawa, M., and Kondoh, H. (2000). Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet* 16, 182-187.
- Kapoor, A., Ratnapriya, R., Kuruttukulam, G., and Anand, A. (2007). A novel genetic locus for juvenile myoclonic epilepsy at chromosome 5q12-q14. *Hum Genet* 121, 655-662.
- Kasperaviciute, 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., et al. (2010). Common genetic variation and susceptibility to partial epilepsies: a genome-wide association study. *Brain* 133, 2136-2147.
- Kennedy, G.C., Matsuzaki, H., Dong, S., Liu, W.M., 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.
- Kinirons, P., Rabinowitz, D., Gravel, M., Long, J., Winawer, M., Senechal, G., Ottman, R., and Cossette, P. (2008a). Phenotypic concordance in 70 families with IGE-implications for genetic studies of epilepsy. *Epilepsy Res* 82, 21-28.
- Kinirons, P., Verlaan, D.J., Dube, M.P., Poirier, J., Deacon, C., Lortie, A., Clement, J.F., Desbiens, R., Carmant, L., Cieuta-Walti, C., et al. (2008b). A novel locus for idiopathic generalized epilepsy in French-Canadian families maps to 10p11. *Am J Med Genet A* 146A, 578-584.
- Kinton, L., Johnson, M.R., Smith, S.J., Farrell, F., Stevens, J., Rance, J.B., Claudino, A.M., Duncan, J.S., Davis, M.B., Wood, N.W., et al. (2002). Partial epilepsy with pericentral spikes: a new familial epilepsy syndrome with evidence for linkage to chromosome 4p15. *Ann Neurol* 51, 740-749.
- 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.
- Ko, J., Humbert, S., Bronson, R.T., Takahashi, S., Kulkarni, A.B., Li, E., and Tsai, L.H. (2001). p35 and p39 are essential for cyclin-dependent kinase 5 function during neurodevelopment. *J Neurosci* 21, 6758-6771.
- Kong, A., and Cox, N.J. (1997). Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 61, 1179-1188.
- Kong, A., Gudbjartsson, D.F., Sainz, J., Jonsdottir, G.M., Gudjonsson, S.A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., et al. (2002). A high-resolution recombination map of the human genome. *Nat Genet* 31, 241-247.
- Krawczak, M., Nikolaus, S., von Eberstein, H., Croucher, P.J., 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.

- Kruglyak, L., Daly, M.J., Reeve-Daly, M.P., and Lander, E.S. (1996). Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58, 1347-1363.
- 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 alpha1 and gamma2 GABAA receptor subunit mutations in families with idiopathic generalized epilepsy. *Eur J Neurosci* 34, 237-249.
- Lander, E., and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11, 241-247.
- Lao, O., Lu, T.T., Nothnagel, M., Junge, O., Freitag-Wolf, S., Caliebe, A., Balascakova, M., Bertranpetit, J., Bindoff, L.A., Comas, D., et al. (2008). Correlation between genetic and geographic structure in Europe. *Curr Biol* 18, 1241-1248.
- Lenzen, K.P., Heils, A., Lorenz, S., Hempelmann, A., Hofels, S., Lohoff, F.W., Schmitz, B., and Sander, T. (2005). Supportive evidence for an allelic association of the human KCNJ10 potassium channel gene with idiopathic generalized epilepsy. *Epilepsy Res* 63, 113-118.
- Leon, 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.
- Letts, V.A., Felix, R., Biddlecome, G.H., Arikath, J., Mahaffey, C.L., Valenzuela, A., Bartlett, F.S., 2nd, Mori, Y., Campbell, K.P., and Frankel, W.N. (1998). The mouse stargazer gene encodes a neuronal Ca²⁺-channel gamma subunit. *Nat Genet* 19, 340-347.
- Leu, C., de Kovel, C.G., Zara, F., Striano, P., Pezzella, M., Robbiano, A., Bianchi, A., Bisulli, F., Coppola, A., Giallonardo, A.T., et al. (2012). Genome-wide linkage meta-analysis identifies susceptibility loci at 2q34 and 13q31.3 for genetic generalized epilepsies. *Epilepsia* 53, 308-318.
- Lewontin, R.C. (1988). On measures of gametic disequilibrium. *Genetics* 120, 849-852.
- Li, H.Y., Li, N., Jiang, H., Shen, L., Guo, J.F., Zhang, R.X., Xia, K., Pan, Q., Zi, X.H., and Tang, B.S. (2008). A novel genetic locus for benign familial infantile seizures maps to chromosome 1p36.12-p35.1. *Clin Genet* 74, 490-492.
- Li, K.X., Lu, Y.M., Xu, Z.H., Zhang, J., Zhu, J.M., Zhang, J.M., Cao, S.X., Chen, X.J., Chen, Z., Luo, J.H., et al. (2011). Neuregulin 1 regulates excitability of fast-spiking neurons through Kv1.1 and acts in epilepsy. *Nat Neurosci* 15, 267-273.
- Lin, D.Y., and Zeng, D. (2010). Meta-analysis of genome-wide association studies: no efficiency gain in using individual participant data. *Genet Epidemiol* 34, 60-66.
- Lohmueller, K.E., Pearce, C.L., Pike, M., Lander, E.S., and Hirschhorn, J.N. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 33, 177-182.
- Lopez de Armentia, M., Jancic, D., Olivares, R., Alarcon, J.M., Kandel, E.R., and Barco, A. (2007). cAMP response element-binding protein-mediated gene expression increases the intrinsic excitability of CA1 pyramidal neurons. *J Neurosci* 27, 13909-13918.
- Lorentzen, A.R., Melum, E., Ellinghaus, E., Smestad, C., Mero, I.L., Aarseth, J.H., Myhr, K.M., Celius, E.G., Lie, B.A., Karlsen, T.H., et al. (2010). Association to the Glypican-5 gene in multiple sclerosis. *J Neuroimmunol* 226, 194-197.
- Luxardi, G., Galli, A., Forlani, S., Lawson, K., Maina, F., and Dono, R. (2007). Glypicans are differentially expressed during patterning and neurogenesis of early mouse brain. *Biochem Biophys Res Commun* 352, 55-60.
- Macdonald, R.L., Kang, J.Q., and Gallagher, M.J. (2010). Mutations in GABAA receptor subunits associated with genetic epilepsies. *J Physiol* 588, 1861-1869.

- Malas, S., Postlethwaite, M., Ekonomou, A., Whalley, B., Nishiguchi, S., Wood, H., Meldrum, B., Constanti, A., and Episkopou, V. (2003). Sox1-deficient mice suffer from epilepsy associated with abnormal ventral forebrain development and olfactory cortex hyperexcitability. *Neuroscience* 119, 421-432.
- Maljevic, S., Krampfl, K., Cobilanschi, J., Tilgen, N., Beyer, S., Weber, Y.G., Schlesinger, F., Ursu, D., Melzer, W., Cossette, P., et al. (2006). A mutation in the GABA(A) receptor alpha(1)-subunit is associated with absence epilepsy. *Ann Neurol* 59, 983-987.
- Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747-753.
- Mantamadiotis, T., Lemberger, T., Bleckmann, S.C., Kern, H., Kretz, O., Martin Villalba, A., Tronche, F., Kellendonk, C., Gau, D., Kapfhammer, J., et al. (2002). Disruption of CREB function in brain leads to neurodegeneration. *Nat Genet* 31, 47-54.
- Mantegazza, M., Gambardella, A., Rusconi, R., Schiavon, E., Annesi, F., Cassulini, R.R., Labate, A., Carrideo, S., Chifari, R., Canevini, M.P., et al. (2005). Identification of an Nav1.1 sodium channel (SCN1A) loss-of-function mutation associated with familial simple febrile seizures. *Proc Natl Acad Sci U S A* 102, 18177-18182.
- Marchini, J., Cardon, L.R., Phillips, M.S., and Donnelly, P. (2004). The effects of human population structure on large genetic association studies. *Nat Genet* 36, 512-517.
- Marchini, J., and Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nat Rev Genet* 11, 499-511.
- Martin, M.S., Tang, B., Papale, L.A., Yu, F.H., Catterall, W.A., and Escayg, A. (2007). The voltage-gated sodium channel Scn8a is a genetic modifier of severe myoclonic epilepsy of infancy. *Hum Mol Genet* 16, 2892-2899.
- Martinez-Juarez, I.E., Alonso, M.E., Medina, M.T., Duron, R.M., Bailey, J.N., Lopez-Ruiz, M., Ramos-Ramirez, R., Leon, L., Pineda, G., Castroviejo, I.P., et al. (2006). Juvenile myoclonic epilepsy subsyndromes: family studies and long-term follow-up. *Brain* 129, 1269-1280.
- Matise, T.C., Chen, F., Chen, W., De La Vega, F.M., Hansen, M., He, C., Hyland, F.C., Kennedy, G.C., Kong, X., Murray, S.S., et al. (2007). A second-generation combined linkage physical map of the human genome. *Genome Res* 17, 1783-1786.
- 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., and Hirschhorn, J.N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 9, 356-369.
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12, 153-157.
- Mefford, H.C., and Mulley, J.C. (2010). Genetically complex epilepsies, copy number variants and syndrome constellations. *Genome Med* 2, 71.
- 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.

- Mikami, M., Yasuda, T., Terao, A., Nakamura, M., Ueno, S., Tanabe, H., Tanaka, T., Onuma, T., Goto, Y., Kaneko, S., et al. (1999). Localization of a gene for benign adult familial myoclonic epilepsy to chromosome 8q23.3-q24.1. *Am J Hum Genet* 65, 745-751.
- Mills, P.B., Surtees, R.A., Champion, M.P., Beesley, C.E., Dalton, N., Scambler, P.J., Heales, S.J., Briddon, A., Scheimberg, I., Hoffmann, G.F., et al. (2005). Neonatal epileptic encephalopathy caused by mutations in the PNPO gene encoding pyridox(am)ine 5'-phosphate oxidase. *Hum Mol Genet* 14, 1077-1086.
- Moeller, F., LeVan, P., Muhle, H., Stephani, U., Dubeau, F., Siniatchkin, M., and Gotman, J. (2010). Absence seizures: individual patterns revealed by EEG-fMRI. *Epilepsia* 51, 2000-2010.
- 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.
- Moore, T., Hecquet, S., McLellann, A., Ville, D., Grid, D., Picard, F., Moulard, B., Asherson, P., Makoff, A.J., McCormick, D., et al. (2001). Polymorphism analysis of JRK/JH8, the human homologue of mouse jerky, and description of a rare mutation in a case of CAE evolving to JME. *Epilepsy Res* 46, 157-167.
- Morton, N.E. (1955). Sequential tests for the detection of linkage. *Am J Hum Genet* 7, 277-318.
- Mullen, S.A., Suls, A., De Jonghe, P., Berkovic, S.F., and Scheffer, I.E. (2010). Absence epilepsies with widely variable onset are a key feature of familial GLUT1 deficiency. *Neurology* 75, 432-440.
- 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.
- Mulley, J.C., and Mefford, H.C. (2011). Epilepsy and the new cytogenetics. *Epilepsia* 52, 423-432.
- 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.
- Musayev, F.N., Di Salvo, M.L., Saavedra, M.A., Contestabile, R., Ghatge, M.S., Haynes, A., Schirch, V., and Safo, M.K. (2009). Molecular basis of reduced pyridoxine 5'-phosphate oxidase catalytic activity in neonatal epileptic encephalopathy disorder. *J Biol Chem* 284, 30949-30956.
- Nabbout, R., Prud'homme, J.F., Herman, A., Feingold, J., Brice, A., Dulac, O., and LeGuern, E. (2002). A locus for simple pure febrile seizures maps to chromosome 6q22-q24. *Brain* 125, 2668-2680.
- Nabbout, R., Baulac, S., Desguerre, I., Bahi-Buisson, N., Chiron, C., Ruberg, M., Dulac, O., and LeGuern, E. (2007). New locus for febrile seizures with absence epilepsy on 3p and a possible modifier gene on 18p. *Neurology* 68, 1374-1381.
- Nagelkerke, N.J., Hoebee, B., Teunis, P., and Kimman, T.G. (2004). Combining the transmission disequilibrium test and case-control methodology using generalized logistic regression. *Eur J Hum Genet* 12, 964-970.
- Nakayama, J., Hamano, K., Iwasaki, N., Nakahara, S., Horigome, Y., Saitoh, H., Aoki, T., Maki, T., Kikuchi, M., Migita, T., et al. (2000). Significant evidence for linkage of febrile seizures to chromosome 5q14-q15. *Hum Mol Genet* 9, 87-91.
- Nakayama, J., Fu, Y.H., Clark, A.M., Nakahara, S., Hamano, K., Iwasaki, N., Matsui, A., Arinami, T., and Ptacek, L.J. (2002). A nonsense mutation of the MASS1 gene in a family with febrile and afebrile seizures. *Ann Neurol* 52, 654-657.

- Nakayama, J., Yamamoto, N., Hamano, K., Iwasaki, N., Ohta, M., Nakahara, S., Matsui, A., Noguchi, E., and Arinami, T. (2004). Linkage and association of febrile seizures to the IMPA2 gene on human chromosome 18. *Neurology* 63, 1803-1807.
- Neubauer, B.A., Fiedler, B., Himmelein, B., Kampfer, F., Lassker, U., Schwabe, G., Spanier, I., Tams, D., Bretscher, C., Moldenhauer, K., et al. (1998). Centrottemporal spikes in families with rolandic epilepsy: linkage to chromosome 15q14. *Neurology* 51, 1608-1612.
- Nikiforov, T.T., Rendle, R.B., Goelet, P., Rogers, Y.H., Kotewicz, M.L., Anderson, S., Trainor, G.L., and Knapp, M.R. (1994). Genetic Bit Analysis: a solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Res* 22, 4167-4175.
- O'Connell, J.R., and Weeks, D.E. (1998). PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63, 259-266.
- Ogiwara, I., Ito, K., Sawaishi, Y., Osaka, H., Mazaki, E., Inoue, I., Montal, M., Hashikawa, T., Shike, T., Fujiwara, T., et al. (2009). De novo mutations of voltage-gated sodium channel alpha1 gene SCN2A in intractable epilepsies. *Neurology* 73, 1046-1053.
- Okamoto, K., Tokunaga, K., Doi, K., Fujita, T., Suzuki, H., Katoh, T., Watanabe, T., Nishida, N., Mabuchi, A., Takahashi, A., et al. (2011). Common variation in GPC5 is associated with acquired nephrotic syndrome. *Nat Genet* 43, 459-463.
- 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. (2005). Analysis of genetically complex epilepsies. *Epilepsia* 46 Suppl 10, 7-14.
- Panayiotopoulos, C.P. (2011). The new ILAE report on terminology and concepts for organization of epileptic seizures: a clinician's critical view and contribution. *Epilepsia* 52, 2155-2160.
- Pandolfo, M. (2011). Genetics of epilepsy. *Semin Neurol* 31, 506-518.
- Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genet* 2, e190.
- Phillips, H.A., Scheffer, I.E., Crossland, K.M., Bhatia, K.P., Fish, D.R., Marsden, C.D., Howell, S.J., Stephenson, J.B., Tolmie, J., Plazzi, G., et al. (1998). Autosomal dominant nocturnal frontal-lobe epilepsy: genetic heterogeneity and evidence for a second locus at 15q24. *Am J Hum Genet* 63, 1108-1116.
- Phillips, H.A., Favre, I., Kirkpatrick, M., Zuberi, S.M., Goudie, D., Heron, S.E., Scheffer, I.E., Sutherland, G.R., Berkovic, S.F., Bertrand, D., et al. (2001). CHRNA2 is the second acetylcholine receptor subunit associated with autosomal dominant nocturnal frontal lobe epilepsy. *Am J Hum Genet* 68, 225-231.
- Pitcher, G.M., Kalia, L.V., Ng, D., Goodfellow, N.M., Yee, K.T., Lambe, E.K., and Salter, M.W. (2011). Schizophrenia susceptibility pathway neuregulin 1-ErbB4 suppresses Src upregulation of NMDA receptors. *Nat Med* 17, 470-478.
- Pluzhnikov, A., Below, J.E., Konkashbaev, A., Tikhomirov, A., Kistner-Griffin, E., Roe, C.A., Nicolae, D.L., and Cox, N.J. (2010). Spoiling the whole bunch: quality control aimed at preserving the integrity of high-throughput genotyping. *Am J Hum Genet* 87, 123-128.

- Poduri, A., Wang, Y., Gordon, D., Barral-Rodriguez, S., Barker-Cummings, C., Ulgen, A., Chitsazzadeh, V., Hill, R.S., Risch, N., Hauser, W.A., et al. (2009). Novel susceptibility locus at chromosome 6q16.3-22.31 in a family with GEFS+. *Neurology* 73, 1264-1272.
- 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.
- Price, A.L., Weale, M.E., Patterson, N., Myers, S.R., Need, A.C., Shianna, K.V., Ge, D., Rotter, J.I., Torres, E., Taylor, K.D., et al. (2008). Long-range LD can confound genome scans in admixed populations. *Am J Hum Genet* 83, 132-135; author reply 135-139.
- Pritchard, J.K., Stephens, M., Rosenberg, N.A., and Donnelly, P. (2000). Association mapping in structured populations. *Am J Hum Genet* 67, 170-181.
- Pritchard, J.K., and Cox, N.J. (2002). The allelic architecture of human disease genes: common disease-common variant...or not? *Hum Mol Genet* 11, 2417-2423.
- Puranam, R.S., Jain, S., Kleindienst, A.M., Saxena, S., Kim, M.K., Kelly Changizi, B., Padma, M.V., Andrews, I., Elston, R.C., Tiwari, H.K., et al. (2005). A locus for generalized tonic-clonic seizure susceptibility maps to chromosome 10q25-q26. *Ann Neurol* 58, 449-458.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., 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., Holden, J.J., 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.
- Reid, C.A., Berkovic, S.F., and Petrou, S. (2009). Mechanisms of human inherited epilepsies. *Prog Neurobiol* 87, 41-57.
- Ruschendorf, F., and Nurnberg, P. (2005). ALOHOMORA: a tool for linkage analysis using 10K SNP array data. *Bioinformatics* 21, 2123-2125.
- Salanti, G., Amountza, G., Ntzani, E.E., and Ioannidis, J.P. (2005). Hardy-Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. *Eur J Hum Genet* 13, 840-848.
- Salek-Haddadi, A., Mayer, T., Hamandi, K., Symms, M., Josephs, O., Fluegel, D., Woermann, F., Richardson, M.P., Noppeney, U., Wolf, P., et al. (2009). Imaging seizure activity: a combined EEG/EMG-fMRI study in reading epilepsy. *Epilepsia* 50, 256-264.
- Salzmann, A., Guipponi, M., Lyons, P.J., Fricker, L.D., Sapio, M., Lambercy, C., Buresi, C., Ouled Amar Bencheikh, B., Lahjouji, F., Ouazzani, R., et al. (2012). Carboxypeptidase A6 gene (CPA6) mutations in a recessive familial form of febrile seizures and temporal lobe epilepsy and in sporadic temporal lobe epilepsy. *Hum Mutat* 33, 124-135.
- Sandberg, M., Kallstrom, M., and Muhr, J. (2005). Sox21 promotes the progression of vertebrate neurogenesis. *Nat Neurosci* 8, 995-1001.
- 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.

- Sander, T., Toliat, M.R., Heils, A., Leschik, G., Becker, C., Ruschendorf, F., Rohde, K., Mundlos, S., and Nurnberg, P. (2002). Association of the 867Asp variant of the human anion exchanger 3 gene with common subtypes of idiopathic generalized epilepsy. *Epilepsy Res* 51, 249-255.
- Sander, T., and Gaus, V. (2007). Genetik der idiopathischen Epilepsien/ The genetics of idiopathic epilepsy *Medizinische Genetik* 19, 325-329.
- Schaffer, A.A., Lemire, M., Ott, J., Lathrop, G.M., and Weeks, D.E. (2011). Coordinated conditional simulation with SLINK and SUP of many markers linked or associated to a trait in large pedigrees. *Hum Hered* 71, 126-134.
- 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.
- Scott, W.K., Hauser, E.R., Schmechel, D.E., Welsh-Bohmer, K.A., Small, G.W., Roses, A.D., Saunders, A.M., Gilbert, J.R., Vance, J.M., Haines, J.L., et al. (2003). Ordered-subsets linkage analysis detects novel Alzheimer disease loci on chromosomes 2q34 and 15q22. *Am J Hum Genet* 73, 1041-1051.
- 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.
- Seelow, D., Schwarz, J.M., and Schuelke, M. (2008). GeneDistiller--distilling candidate genes from linkage intervals. *PLoS One* 3, e3874.
- Sen, A., Thom, M., Martinian, L., Jacobs, T., Nikolic, M., and Sisodiya, S.M. (2006). Deregulation of cdk5 in Hippocampal sclerosis. *J Neuropathol Exp Neurol* 65, 55-66.
- 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.
- Shorvon, S.D. (2011). The etiologic classification of epilepsy. *Epilepsia* 52, 1052-1057.
- Simino, J., Shi, G., Arnett, D., Broeckel, U., Hunt, S.C., and Rao, D.C. (2011). Variants on chromosome 6p22.3 associated with blood pressure in the HyperGEN study: follow-up of FBPP quantitative trait loci. *Am J Hypertens* 24, 1227-1233.
- Singh, N.A., Charlier, C., Stauffer, D., DuPont, B.R., Leach, R.J., Melis, R., Ronen, G.M., Bjerre, I., Quattlebaum, T., Murphy, J.V., et al. (1998). A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. *Nat Genet* 18, 25-29.
- Singh, N.A., Pappas, C., Dahle, E.J., Claes, L.R., Pruess, T.H., De Jonghe, P., Thompson, J., Dixon, M., Gurnett, C., Peiffer, A., et al. (2009). A role of SCN9A in human epilepsies, as a cause of febrile seizures and as a potential modifier of Dravet syndrome. *PLoS Genet* 5, e1000649.
- Sisodiya, S.M., Thom, M., Lin, W.R., Bajaj, N.P., Cross, J.H., and Harding, B.N. (2002). Abnormal expression of cdk5 in focal cortical dysplasia in humans. *Neurosci Lett* 328, 217-220.
- Skol, A.D., Scott, L.J., Abecasis, G.R., and Boehnke, M. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* 38, 209-213.
- Slatkin, M. (2008). Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* 9, 477-485.
- 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.
- Stark, A.E. (2006). A clarification of the Hardy-Weinberg law. *Genetics* 174, 1695-1697.

- Steemers, F.J., Chang, W., Lee, G., Barker, D.L., Shen, R., and Gunderson, K.L. (2006). Whole-genome genotyping with the single-base extension assay. *Nat Methods* 3, 31-33.
- Steffens, M., Becker, T., Sander, T., Fimmers, R., Herold, C., Holler, D.A., Leu, C., Herms, S., Cichon, S., Bohn, B., et al. (2010). Feasible and successful: genome-wide interaction analysis involving all 1.9 x 10¹¹ pair-wise interaction tests. *Hum Hered* 69, 268-284.
- Steinberg, S., de Jong, S., Andreassen, O.A., Werge, T., Borglum, A.D., Mors, O., Mortensen, P.B., Gustafsson, O., Costas, J., Pietilainen, O.P., et al. (2011). Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet* 20, 4076-4081.
- Steinlein, O.K., Mulley, J.C., Propping, P., Wallace, R.H., Phillips, H.A., Sutherland, G.R., Scheffer, I.E., and Berkovic, S.F. (1995). A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. *Nat Genet* 11, 201-203.
- Stogmann, E., Lichtner, P., Baumgartner, C., Bonelli, S., Assem-Hilger, E., Leutmezer, F., Schmied, M., Hotzy, C., Strom, T.M., Meitinger, T., et al. (2006). Idiopathic generalized epilepsy phenotypes associated with different EFHC1 mutations. *Neurology* 67, 2029-2031.
- Stouffer, S.A., Suchman, E.A., DeVinney, L.C., Star, S.A., and Jr., W.R.M. (1949). *Studies in Social Psychology in World War II: The American Soldier Volume 1, Adjustment During Army Life*. Princeton: Princeton University Press, 125.
- Strachan, T., and Read, A.P. (2005). *Molekulare Humangenetik*. Spektrum Akademischer Verlag 3. Auflage.
- Striano, P., Chifari, R., Striano, S., de Fusco, M., Elia, M., Guerrini, R., Casari, G., and Canevini, M.P. (2004). A new benign adult familial myoclonic epilepsy (BAFME) pedigree suggesting linkage to chromosome 2p11.1-q12.2. *Epilepsia* 45, 190-192.
- Striano, P., Bordo, L., Lispi, M.L., Specchio, N., Minetti, C., Vigevano, F., and Zara, F. (2006). A novel SCN2A mutation in family with benign familial infantile seizures. *Epilepsia* 47, 218-220.
- Striano, P., Mancardi, M.M., Biancheri, R., Madia, F., Gennaro, E., Paravidino, R., Beccaria, F., Capovilla, G., Dalla Bernardina, B., Darra, F., et al. (2007). Brain MRI findings in severe myoclonic epilepsy in infancy and genotype-phenotype correlations. *Epilepsia* 48, 1092-1096.
- Striano, P., Weber, Y.G., Toliat, M.R., Schubert, J., Leu, C., Chaimana, R., Baulac, S., Guerrero, R., Leguern, E., Lehesjoki, A.E., et al. (2012). GLUT1 mutations are a rare cause of familial idiopathic generalized epilepsy. *Neurology* 78, 557-562.
- Strug, L.J., Clarke, T., Chiang, T., Chien, M., Baskurt, Z., Li, W., Dorfman, R., Bali, B., Wirrell, E., Kugler, S.L., et al. (2009). Centrotemporal sharp wave EEG trait in rolandic epilepsy maps to Elongator Protein Complex 4 (ELP4). *Eur J Hum Genet* 17, 1171-1181.
- Sugawara, T., Tsurubuchi, Y., Agarwala, K.L., Ito, M., Fukuma, G., Mazaki-Miyazaki, E., Nagafuji, H., Noda, M., Imoto, K., Wada, K., et al. (2001). A missense mutation of the Na⁺ channel alpha II subunit gene Na(v)1.2 in a patient with febrile and afebrile seizures causes channel dysfunction. *Proc Natl Acad Sci U S A* 98, 6384-6389.
- Suls, A., Mullen, S.A., Weber, Y.G., Verhaert, K., Ceulemans, B., Guerrini, R., Wuttke, T.V., Salvó-Vargas, A., Deprez, L., Claes, L.R., 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.

- Szepetowski, P., Rochette, J., Berquin, P., Piussan, C., Lathrop, G.M., and Monaco, A.P. (1997). Familial infantile convulsions and paroxysmal choreoathetosis: a new neurological syndrome linked to the pericentromeric region of human chromosome 16. *Am J Hum Genet* 61, 889-898.
- 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, G.H., Liu, Y.Y., Hu, X.L., Yin, D.M., Mei, L., and Xiong, Z.Q. (2011). Neuregulin 1 represses limbic epileptogenesis through ErbB4 in parvalbumin-expressing interneurons. *Nat Neurosci* 15, 258-266.
- Tanaka, M., Olsen, R.W., Medina, M.T., Schwartz, E., Alonso, M.E., Duron, R.M., Castro-Ortega, R., Martinez-Juarez, I.E., Pascual-Castroviejo, I., Machado-Salas, J., et al. (2008). Hyperglycosylation and reduced GABA currents of mutated GABRB3 polypeptide in remitting childhood absence epilepsy. *Am J Hum Genet* 82, 1249-1261.
- Tang, D., Yeung, J., Lee, K.Y., Matsushita, M., Matsui, H., Tomizawa, K., Hatase, O., and Wang, J.H. (1995). An isoform of the neuronal cyclin-dependent kinase 5 (Cdk5) activator. *J Biol Chem* 270, 26897-26903.
- Tao, H., Manak, J.R., Sowers, L., Mei, X., Kiyonari, H., Abe, T., Dahdaleh, N.S., Yang, T., Wu, S., Chen, S., et al. (2011). Mutations in prickle orthologs cause seizures in flies, mice, and humans. *Am J Hum Genet* 88, 138-149.
- 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.
- Terwilliger, J.D., and Ott, J. (1992). A haplotype-based 'haplotype relative risk' approach to detecting allelic associations. *Hum Hered* 42, 337-346.
- 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.
- Thomson, G. (1995). Mapping disease genes: family-based association studies. *Am J Hum Genet* 57, 487-498.
- Ting, A.K., Chen, Y., Wen, L., Yin, D.M., Shen, C., Tao, Y., Liu, X., Xiong, W.C., and Mei, L. (2011). Neuregulin 1 promotes excitatory synapse development and function in GABAergic interneurons. *J Neurosci* 31, 15-25.
- Tobias, A., Saxena, M., and Lelievre, V. (2009). CDK5: the "pathfinder" for new born neurons in adult hippocampus? *Cell Adh Migr* 3, 319-321.
- Uchikawa, M., Kamachi, Y., and Kondoh, H. (1999). Two distinct subgroups of Group B Sox genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken. *Mech Dev* 84, 103-120.
- Vanmolkot, K.R., Kors, E.E., Hottenga, J.J., Terwindt, G.M., Haan, J., Hoefnagels, W.A., Black, D.F., Sandkuijl, L.A., Frants, R.R., Ferrari, M.D., et al. (2003). Novel mutations in the Na⁺, K⁺-ATPase pump gene ATP1A2 associated with familial hemiplegic migraine and benign familial infantile convulsions. *Ann Neurol* 54, 360-366.
- Vilas, G.L., Johnson, D.E., Freund, P., and Casey, J.R. (2009). Characterization of an epilepsy-associated variant of the human Cl⁻/HCO₃⁻ exchanger AE3. *Am J Physiol Cell Physiol* 297, C526-536.

- Wakamatsu, N., Yamada, Y., Yamada, K., Ono, T., Nomura, N., Taniguchi, H., Kitoh, H., Mutoh, N., Yamanaka, T., Mushiake, K., et al. (2001). Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. *Nat Genet* 27, 369-370.
- Wallace, R.H., Berkovic, S.F., Howell, R.A., Sutherland, G.R., and Mulley, J.C. (1996). Suggestion of a major gene for familial febrile convulsions mapping to 8q13-21. *J Med Genet* 33, 308-312.
- Wallace, R.H., Wang, D.W., Singh, R., Scheffer, I.E., George, A.L., Jr., Phillips, H.A., Saar, K., Reis, A., Johnson, E.W., Sutherland, G.R., et al. (1998). Febrile seizures and generalized epilepsy associated with a mutation in the Na⁺-channel beta1 subunit gene SCN1B. *Nat Genet* 19, 366-370.
- 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.
- Wallace, R.H., Hodgson, B.L., Grinton, B.E., Gardiner, R.M., Robinson, R., Rodriguez-Casero, V., Sadleir, L., Morgan, J., Harkin, L.A., Dibbens, L.M., et al. (2003). Sodium channel alpha1-subunit mutations in severe myoclonic epilepsy of infancy and infantile spasms. *Neurology* 61, 765-769.
- Wang, X., Ching, Y.P., Lam, W.H., Qi, Z., Zhang, M., and Wang, J.H. (2000). Identification of a common protein association region in the neuronal Cdk5 activator. *J Biol Chem* 275, 31763-31769.
- 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, H., Haiman, C.A., Kolonel, L.N., Henderson, B.E., Wilkens, L.R., Le Marchand, L., and Stram, D.O. (2010). Self-reported ethnicity, genetic structure and the impact of population stratification in a multiethnic study. *Hum Genet* 128, 165-177.
- Wang, J., and Shete, S. (2012). Testing departure from Hardy-Weinberg proportions. *Methods Mol Biol* 850, 77-102.
- Weinberg, W. (1908). Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg* 64, 368-382.
- Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Edgar, R., Federhen, S., et al. (2007). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 35, D5-12.
- Whittemore, A.S., and Halpern, J. (1994). Probability of gene identity by descent: computation and applications. *Biometrics* 50, 109-117.
- Whittemore, A.S., and Halpern, J. (2001). Problems in the definition, interpretation, and evaluation of genetic heterogeneity. *Am J Hum Genet* 68, 457-465.
- 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.
- Wigginton, J.E., Cutler, D.J., and Abecasis, G.R. (2005). A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 76, 887-893.
- Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 26, 2190-2191.
- Wiltshire, S., Cardon, L.R., and McCarthy, M.I. (2002). Evaluating the results of genomewide linkage scans of complex traits by locus counting. *Am J Hum Genet* 71, 1175-1182.
- 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.
- Wolf, P. (2011). Networks and systems, conceptualizations, and research. *Epilepsia* 52, 1198-1200; discussion 1205-1199.
- Wong, M. (2010). Juvenile myoclonic epilepsy: is it an idiopathic epilepsy caused by a malformation of cortical development? *Epilepsy Curr* 10, 69-71.
- Wong, M. (2011). Epilepsy is both a symptom and a disease: a proposal for a two-tiered classification system. *Epilepsia* 52, 1201-1203; discussion 1205-1209.
- Wright, F.A., Strug, L.J., Doshi, V.K., Commander, C.W., Blackman, S.M., Sun, L., Berthiaume, Y., Cutler, D., Cojocar, A., Collaco, J.M., et al. (2011). Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2. *Nat Genet* 43, 539-546.
- Wu, D.C., Yu, Y.P., Lee, N.T., Yu, A.C., Wang, J.H., and Han, Y.F. (2000). The expression of Cdk5, p35, p39, and Cdk5 kinase activity in developing, adult, and aged rat brains. *Neurochem Res* 25, 923-929.
- Yalcin, O. (2012). Genes and molecular mechanisms involved in the epileptogenesis of idiopathic absence epilepsies. *Seizure* 21, 79-86.
- Yamada, K., Yamada, Y., Nomura, N., Miura, K., Wakako, R., Hayakawa, C., Matsumoto, A., Kumagai, T., Yoshimura, I., Miyazaki, S., et al. (2001). Nonsense and frameshift mutations in ZFX1B, encoding Smad-interacting protein 1, cause a complex developmental disorder with a great variety of clinical features. *Am J Hum Genet* 69, 1178-1185.
- Yau, H.J., Wang, H.F., Lai, C., and Liu, F.C. (2003). Neural development of the neuregulin receptor ErbB4 in the cerebral cortex and the hippocampus: preferential expression by interneurons tangentially migrating from the ganglionic eminences. *Cereb Cortex* 13, 252-264.
- 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.
- Zara, F., Bianchi, A., Avanzini, G., Di Donato, S., Castellotti, B., Patel, P.I., and Pandolfo, M. (1995). Mapping of genes predisposing to idiopathic generalized epilepsy. *Hum Mol Genet* 4, 1201-1207.
- Zhu, X., Zhang, S., Zhao, H., and Cooper, R.S. (2002). Association mapping, using a mixture model for complex traits. *Genet Epidemiol* 23, 181-196.
- Zou, G.Y., and Donner, A. (2006). The merits of testing Hardy-Weinberg equilibrium in the analysis of unmatched case-control data: a cautionary note. *Ann Hum Genet* 70, 923-933.

DANKSAGUNG

Herrn Prof. Dr. Peter Nürnberg danke ich für seine exzellente Betreuung und Unterstützung meiner Dissertation, die Bereitstellung eines hochmodernen Labors mit seinem kompetenten Team. Ich danke Herrn Prof. Nürnberg für die Ermöglichung der vorliegenden Arbeit.

Frau Prof. Dr. Brunhilde Wirth danke ich für die freundliche Übernahme des Koreferats.

Herrn Dr. habil. Thomas Sander danke ich für die Bereitstellung des hochinteressanten Themas und seiner intensiven Betreuung, seinem Interesse und Unterstützung. Seiner Kompetenz und Diskussionsbereitschaft verdanke ich mein tiefgreifendes Verständnis für die Methodik genetischer Studien, sowie deren Planung und Optimierung. Darüber hinaus hat seine Kooperationsfähigkeit mit vielen Epilepsiezentren mir die Arbeit an diesem weltweit einmaligen Patientenkollektiv ermöglicht.

Ich möchte mich für die Unterstützung bei den vielen kompetenten und freundlichen Personen bedanken, ohne die diese Arbeit nicht möglich gewesen wäre.

Herrn Prof. Dr. Thomas Wienker danke ich für seine Unterstützung und seinem Vertrauen. Er verdient meinen höchsten Respekt.

Herrn Dr. Michael Steffens danke ich für seine hochkompetenten Antworten in Statistikfragen. Seine Anleitung der statistischen Methodik hat dieser Arbeit den Weg geebnet.

Herrn Dr. Mohammad Toliat danke ich für seine stets freundliche und erfahrene Unterstützung im Labor.

Mein besonderer Dank gilt Herrn Sebastian Fey für seine freundliche und teamfähige Persönlichkeit im Labor, sowie seine exzellente technische Assistenz und Hilfestellung.

Für die freundliche und warme Arbeitsatmosphäre möchte ich mich beim Team der Epilepsiegenetik, insbesondere bei Holger Trucks und Dennis Lal bedanken. Die fachliche Diskussion mit Ihnen war nicht nur fruchtbar, sondern auch die Zusammenarbeit mit ihnen hat mir großen Spass gemacht.

Für die technische Unterstützung danke ich Herrn Wilfried Gunia.

Herrn Dr. Holger Thiele danke ich für seine Verbindlichkeit und viele richtungsweisende Gespräche.

Herrn Dr. Markus Leber danke ich für seine bereitwillige Unterstützung und die unvergessliche Rundreise durch Istanbul 2010.

Mein spezieller Dank gilt Herrn Thomas Alef und Herrn Sebastian Fey für ihre freundschaftliche Unterstützung und Begleitung zu jeder Zeit.

Frau Dr. Breitenkamp, Herrn Alexander Münchenhangen, meiner Familie und allen meinen Freunden danke ich zutiefst für das unermüdliche Vertrauen und den Zuspruch in den vergangenen Jahren meiner Promotion.

Zusätzlich entschuldige ich mich bei allen, die hier nicht Erwähnung gefunden haben. Ich hoffe, sie haben Verständnis dafür, dass solche letzten Seiten, wie immer, unter enormen Zeitdruck entstehen.

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."

13.06.2013

Datum



Unterschrift

CURRICULUM VITAE

Persönliche Daten

Name	Costin Leu
Adresse	Zülpicherstr. 265 in 50937 Köln
Geburtsdatum	26.05.1978
Geburtsort	Bukarest
Staatsangehörigkeit	deutsch
Familienstand	ledig

Bildung und wissenschaftliche Tätigkeiten

08/2007 – 07/2012	Promotion bei Prof. Dr. Nürnberg im Bereich Epilepsiegenetik am „Cologne Center for Genomics“ der Universität zu Köln. <u>Thema:</u> „ <i>Molekulargenetische Exploration der idiopathisch generalisierten Epilepsien</i> “.
04/2007 – 07/2007	Asthma-Studie von GH. Koppelman, Universität Groningen
10/1997 – 04/2007	Studium der Biologie an der Universität zu Köln, Abschluss: Diplom Biologe, Gesamtnote „sehr gut“. Hauptfach: Genetik; Nebenfächer: Biochemie, Organische Chemie Diplomarbeit am „Cologne Center for Genomics“ der Universität zu Köln, im Bereich Neurogenetik, <u>Thema:</u> „ <i>Die molekulargenetischen Grundlagen einer neuen Form der familiären Opticus Atrophie</i> “

Publikationen

EPICURE Consortium, Leu C., de Kovel C. G., Zara F., Striano P., Pezzella M., Robbiano A., Bianchi A., Bisulli F., Coppola A., Giallonardo A. T., Beccaria F., Trenité D. K., Lindhout D., Gaus V., Schmitz B., Janz D., Weber Y. G., Becker F., Lerche H., Kleefuss-Lie A. A., Hallman K., Kunz W. S., Elger C. E., Muhle H., Stephani U., Møller R. S., Hjalgrim H., Mullen S., Scheffer I. E., Berkovic S. F., Everett K. V., Gardiner M. R., Marini C., Guerrini R., Lehesjoki A. E., Siren A., Nabbout R., Baulac S., Leguern E., Serratos J. M., 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., Rüschenhoff F., Trucks H., Nürnberg P., Avanzini G., Koeleman B. P., Sander T. (2012). Genome-wide linkage meta-analysis identifies susceptibility loci at 2q34 and 13q31.3 for genetic generalized epilepsies. *Epilepsia*, 53(2), 308-18.

Striano P., Weber Y. G., Toliat M. R., Schubert J., Leu C., Chaimana R., Baulac S., Guerrero R., LeGuern E., Lehesjoki A. E., Polvi A., Robbiano A., Serratos J. M., Guerrini R., Nuernberg P., Sander T., Zara F., Lerche H. & Marini C. (2012). GLUT1-mutations are a rare cause of familial idiopathic generalized epilepsy. *Neurology*, 78(8), 557-562.

Steffens M., Becker T., Sander T., Fimmers R., Herold C., Holler D. A., Leu C., Herms S., Cichon S., Bohn B., Gerstner T., Griebel M., Nöthen M. M., Wienker T. F., Baur M. P. (2010). Feasible and successful: genome-wide interaction analysis involving all 1.9×10^{11} pair-wise interaction tests. *Human Heredity*, 69(4), 268-284.

de Kovel C. G., Pinto D., Tauer U., Lorenz S., Muhle H., Leu C., Neubauer B. A., Hempelmann A., Callenbach P. M., Scheffer I. E., Berkovic S. F., Rudolf G., Striano P., Siren A., Baykan B., Sander T., Lindhout D., Kasteleijn-Nolst Trenité D. G., Stephani U., Koeleman B. P. (2010). Whole-genome linkage scan for epilepsy-related photosensitivity: a mega-analysis. *Epilepsy Research*, 89(2-3), 286-294.

de Kovel C. G., Trucks H., Helbig I., Mefford H. C., Baker C., Leu C., Kluck C., Muhle H., von Spiczak S., Ostertag P., Obermeier T., Kleefuss-Lie A. A., Hallmann K., Steffens M., Gaus V., Klein K. M., Hamer H. M., Rosenow F., Brilstra E. H., Trenité D. K., Swinkels M. E., Weber Y. G., Unterberger I., Zimprich F., Urak L., Feucht M., Fuchs K., Møller R. S., Hjalgrim H., De Jonghe P., Suls A., Rückert I. M., Wichmann H. E., Franke A., Schreiber S., Nürnberg P., Elger C. E., Lerche H., Stephani U., Koeleman B. P., Lindhout D., Eichler E. E., Sander T. (2010). Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain*, 133(Pt 1), 23-32.

Dibbens L. M., Mullen S., Helbig I., Mefford H. C., Bayly M. A., Bellows S., Leu C., Trucks H., Obermeier T., Wittig M., Franke A., Caglayan H., Yapici Z., EPICURE Consortium, Sander T., Eichler E. E., Scheffer I. E., Mulley J. C., Berkovic S. F. (2009). Familial and sporadic 15q13.3 microdeletions in idiopathic generalized epilepsy: precedent for disorders with complex inheritance. *Human Molecular Genetics*, 18(19), 3626-3631.

Helbig I., Mefford H. C., Sharp A. J., Guipponi M., Fichera M., Franke A., Muhle H., de Kovel C., Baker C., von Spiczak S., Kron K. L., Steinich I., Kleefuss-Lie A. A., Leu C., Gaus V., Schmitz B., Klein K. M., Reif P. S., Rosenow F., Weber Y., Lerche H., Zimprich F., Urak L., Fuchs K., Feucht M., Genton P., Thomas P., Visscher F., de Haan G. J., Møller R. S., Hjalgrim H., Luciano D., Wittig M., Nothnagel M., Elger C. E., Nürnberg P., Romano C., Malafosse A., Koeleman B. P., Lindhout D., Stephani U., Schreiber S., Eichler E. E., Sander T. (2009). 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. *Nature Genetics*, 41(2), 160-162.

te Meerman G., de Vries A., Oosterom E., Niens M., Altmüller J., Leu C., Hofstra R., Wijmenga C., Leal S., Nürnberg P., Postma D., Koppelman G. (2009). Haplotype sharing analysis of null alleles at 7,650 loci in a 474 trio genome-wide screen demonstrates a recent origin for mutations in the primer site. In A. R. de Vries (Ed.), *The value of haplotypes* (pp. 25-36).

Eigene Posterbeiträge und Abstracts

Leu C., Steffens M., de Kovel C. G. F., Koeleman B. P. C., Weber Y. G., Lerche H., Surges R., Kunz W., Elger C. E., Muhle H., Stephani U., Rückert I. M., Wichmann H. E., Franke A., Schreiber S., Leber M., Trucks H., Toliat M. R., Wienker T. F., Nürnberg P., Sander T., EPICURE Consortium. (2010). Genome-wide association mapping of susceptibility alleles predisposing to idiopathic generalized epilepsy. 3rd Annual Meeting of NGFN-Plus and NGFN-Transfer 2010, FU-Berlin, Deutschland.

Leu C., Steffens M., de Kovel C. G. F., Koeleman B. P. C., Lindhout D., Weber Y. G., Ler-

che H., Kleefuß-Lie A. A., Hallmann K., Elger C. E., Muhle H., Stephani U., Rückert I. M., Wichmann H. E., Franke A., Schreiber S., Trucks H., Toliat M., Nürnberg P., Sander T., EPICURE Consortium. (2010). Genome-wide association mapping of susceptibility alleles predisposing to idiopathic generalised epilepsy. GfH Abstract, Medizinische Genetik, 1, 175.

Leu C. & EPICURE Consortium. (2010). Genome-wide linkage mapping of susceptibility genes for idiopathic generalized epilepsy. Fourth Annual EPICURE Meeting, INMED, Marseille, Frankreich.

Leu C., Steffens M., de Kovel C. G. F., Koeleman B. P. C., Lindhout D., Weber Y. G., Lerche H., Kleefuß-Lie A. A., Hallmann K., Elger C. E., Muhle H., Stephani U., Rückert I. M., Wichmann H. E., Franke A., Schreiber S., Trucks H., Toliat M. R., Nürnberg P., Sander T., EPICURE Consortium. (2009). Genome-wide association mapping of susceptibility alleles predisposing to idiopathic generalized epilepsy. 2-nd Annual Meeting of NGFN-Plus and NGFN-Transfer 2009, FU-Berlin, Deutschland.

Mattheisen M., Steffens M., Becker T., Sander T., Fimmers R., Herold C., Holler D. A., Leu C., Herms S., Cichon S., Bohn B., Gerstner T., Griebel M., Nöthen M. M., Baur M. P., Wienker T. F. (2009). Lessons to be Learned From Genome-wide Interaction Analysis (GWIA). IGES Abstract, Genetic Epidemiology, 33, 752-835.

Leu C., de Kovel C. G. F., Trucks H., Helbig I., Mefford H. C., Nürnberg P., Elger C. E., Lerche H., Stephani U., Koeleman B. P. C., Lindhout D., Eichler E.E. & Sander T. (2009). Recurrent microdeletions increase risk for idiopathic generalized epilepsy. Genomics of Common Diseases Conference 2009, Wellcome Trust, Cambridge, Vereinigtes Königreich.

Leu C., Helbig I., Sander T & EPICURE Consortium. (2008). WP1.2 Genome-wide Association Mapping of Common Gene Variants predisposing to Idiopathic Generalized Epilepsy. EPICURE Second Annual Meeting, Malta.

Teilnahme an Kongressen

25 – 27 Nov 2010	„3-rd Annual Meeting of NGFN-Plus and NGFN-Transfer“, FU Berlin
20 – 24 Okt 2010	„IntenC Kick off Meeting“, DETAE, Universität Istanbul, Türkei
19 Mrz 2010	„Genomics Day 2010“, CCG / CECAD, Universität zu Köln
02 – 04 Mrz 2010	„21. Jahrestagung der Gesellschaft für Humangenetik (GfH)“, Universität Hamburg
29 – 30 Jan 2010	„Fourth Annual EPICURE Meeting“, INMED-INSERM, Marseille, Frankreich
26 – 28 Nov 2009	„2-nd Annual Meeting of NGFN-Plus and NGFN-Transfer“, FU Berlin
23 – 26 Sep 2009	„Genomics of Common Diseases Conference 2009“, Wellcome Trust, Cambridge, UK
18 – 20 Mrz 2009	„Cologne Spring Meeting 2009, The Variable Genome“,

Universität zu Köln

- 06 – 08 Feb 2009 EPICURE Third Annual Meeting, “Functional genomics and neurobiology of epilepsy: a basis for new therapeutic strategies“, Charité, Berlin
- 21 – 25 Sep 2008 „8th European Congress on Epileptology (ECE)“, ICC Berlin
- 11 – 14 Jun 2008 „EPICURE Workshop“, Madrid, Spanien
- 09 – 11 Nov 2007 „6th NGFN Meeting 2007“ und „NGFN Qualitätsmanagement Workshop“, DKFZ, Heidelberg