

**Identifizierung von Hochpenetranzvarianten
an genomweit signifikanten Genloci
für die Bipolare Störung**

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Hohen Medizinischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

Katharina Drexler, geb. Karhoff

aus Münster

2024

Angefertigt mit der Genehmigung
der Medizinischen Fakultät der Universität Bonn

1. Gutachter: Prof. Dr. Markus M. Nöthen
2. Gutachter: Prof. Dr. Andreas Mayr

Tag der Mündlichen Prüfung: 27.08.2024

Aus dem Institut für Humangenetik
Direktor: Prof. Dr. med. Markus M. Nöthen

Inhaltsverzeichnis

	Abkürzungsverzeichnis	5
1.	Einleitung	9
1.1	Die Bipolare Störung	9
1.2	Genetische Grundlagen	10
1.3	Stand der Forschung	15
1.4	Genetische Überlappungen bei psychiatrischen Erkrankungen	18
1.5	Fragestellung dieser Arbeit	18
2.	Material und Methoden	20
2.1	Auswahl der DNA-Proben	20
2.2	Auswahl der Kandidatengene	20
2.3	Resequenzierung mittels <i>Molecular-Inversion-Probes</i> -Methode	29
2.4	Anpassung der smMIP-Primer (<i>Balancing</i> und <i>Rebalancing</i>)	36
2.5	Vorbereitung der DNA-Proben	39
2.6	Sequenzierung	39
2.7	Bioinformatische Auswertung und Qualitätsanalyse	41
2.8	Genetische <i>Burdenanalyse</i>	42
3.	Ergebnisse	44
3.1	Identifizierte Varianten nach Qualitätsanalyse	44
3.1.1	<i>ADD3</i>	47
3.1.2	<i>ANK3</i>	47
3.1.3	<i>GRIN2A</i>	48
3.2	Genetische <i>Burdenanalyse</i>	48
3.3	<i>Residual Variation Intolerance Score (RVIS)</i>	50
4.	Diskussion	51
4.1	Vielversprechende Kandidatengene und -varianten	51
4.1.1	<i>GRIN2A</i>	54
4.1.2	<i>ANK3</i>	55
4.1.3	<i>ADD3</i>	56
4.2	Limitationen	56

4.3	Ausblick	57
5.	Zusammenfassung	58
6.	Anhang	59
7.	Abbildungsverzeichnis	70
8.	Tabellenverzeichnis	71
9.	Literaturverzeichnis	72
10.	Danksagung	93

Abkürzungsverzeichnis

AF	Allelfrequenz
ANNOVAR	<i>ANNOtate VARiation</i> (Annotation von Varianten)
ATP	Adenosintriphosphat
BAM	<i>Binary Alignment Map</i> (Komprimierte, binäre Version einer Datei mit Informationen über Reads und deren Mapping im Genom)
BCL	<i>Binary Base Call</i> (Datei, welche die Sequenzrohdaten enthält)
BipEx	<i>Bipolar Exome</i> (Sequenzierprojekt für die Bipolare Störung)
BP/bp	Basenpaare
BWA-MEM	<i>Burrows-Wheeler Alignment Maximum Exact Matches</i> (Algorithmus zum Angleichen der Reads an das Referenzgenom)
CADD	<i>Combined Annotation Dependent Depletion</i> (Score zur Beurteilung der funktionellen Relevanz von Varianten)
cAMP	cyklisches Adenosinmonophosphat
CCDS	<i>Consensus CoDing Sequence</i> (Projekt zur Identifizierung einheitlicher, hochqualitativer Konsensussequenzen)
CHR	Chromosom
CNV	<i>Copy Number Variant</i> (Kopienzahlvariante)
ddA	DiDesoxyAdenosintriphosphat
ddC	DiDesoxyCytidintriphosphat

ddG	DiDesoxyGuanosintriphosphat
ddT	DiDesoxyThymidintriphosphat
(ds)DNA	<i>double-stranded DeoxyriboNucleic Acid</i> (doppelsträngige Desoxyribonukleinsäure)
dNTP	Desoxyribonucleosidtriphosphat
DSM	<i>Diagnostic and Statistical Manual of Mental Disorders</i> (Diagnostisches und Statistisches Manual Psychischer Störungen)
EB	<i>Elution Buffer</i> (Verdünnungspuffer)
EDTA	EthylenDiaminTetraAcetat
eQTL	<i>expression Quantitative Trait Locus</i> (Effekte von genetischen Varianten auf die Genexpression)
FDR	<i>False Discovery Rate</i> (Falscherkennungsrate)
FU	<i>Fluorescence Unit</i> (Fluoreszierende Einheit)
GATK	<i>Genome Analysis ToolKit</i> (Werkzeug zur Analyse von genetischen Daten)
GBD	<i>Global Burden of Disease</i> (Studie zur globalen Krankheitslast)
gnomAD	<i>Genome Aggregation Database</i> (Browser für aggregierte Genomdatensätze)
GTE _x	<i>Genotype-Tissue Expression</i> (Projekt zur Untersuchung der Genexpression und ihres Zusammenhangs mit genetischer Variation in verschiedenen Geweben)
GWAS	<i>Genome-Wide Association Study</i> (Genomweite Assoziationsstudie)

hg	<i>Human Genome</i> (Menschliches Genom)
HT	<i>HybridizaTion Buffer</i> (Hybridisierungspuffer)
ICD	<i>International Classification of Diseases</i> (Internationale Klassifikation der Krankheiten)
ID	Identifikationsnummer
IDT	<i>Integrated DNA Technologies</i> (Biotechnologieunternehmen für DNA-Synthesierung)
InDel	Insertion/Deletion
k.A.	keine Angabe
MAF	<i>Minor Allele Frequency</i> (Frequenz des selteneren Allels)
(sm)MIP	<i>(single molecule) Molecular Inversion Probe</i> (Einzelsträngiges DNA-Molekül, welches zur Zielsequenz komplementäre DNA enthält)
NaOH	Natriumhydroxid
NGS	<i>Next Generation Sequencing</i> (Sequenzierung der nächsten Generation)
NMDA	N-Methyl-D-Aspartat
OH	Hydroxid
P	Phosphat
PCR	<i>Polymerase Chain Reaction</i> (Polymerasekettenreaktion)
PEAR	<i>Paired-End reAd MergeR</i> (Software zur Vereinigung von Reads)

PGC	<i>Psychiatric Genomics Consortium</i> (Wissenschaftliches Konsortium mit dem Forschungsschwerpunkt psychiatrische Genetik)
(m)RNA	(messenger-) <i>Ribonucleic Acid</i> (Messenger-) Ribonukleinsäure)
RS/rs	<i>Reference SNP</i> (Referenz-SNP)
RVIS	<i>Residual Variation Intolerance Score</i> (Score zur Prüfung der Relevanz von Genen)
SCHEMA	<i>SCHizophrenia Exome sequencing Meta-Analysis</i> (Konsortium zur Analyse von Sequenzdaten in der Schizophrenie)
SKAT	<i>Sequence Kernel Association Test</i> (Test für Varianten-Phänotyp-Assoziationen)
SNP	<i>Single Nucleotide Polymorphism</i> (Einzelbasenpolymorphismus)
SNV	<i>Single Nucleotide Variant</i> (Einzelbasenvariante)
SOCS	<i>Suppressor Of Cytokine Signaling</i> (Unterdrücker des Zytokinsignalwegs)
UCSC	<i>University of California, Santa Cruz</i> (Universität von Kalifornien, Santa Cruz)
VCF	<i>Variant Call Format</i> (Format zur Beschreibung von genetischen Varianten)
WES	<i>Whole Exome Sequencing</i> (Exomsequenzierung)

1. Einleitung

1.1 Die Bipolare Störung

Die Bipolare Störung gehört mit einer Lebenszeitprävalenz von etwa 1-2 % zu den häufigsten psychiatrischen Erkrankungen weltweit (Merikangas et al., 2011). Etwa 50 Millionen Menschen sind auf der ganzen Welt von dieser Erkrankung betroffen (O'Connell und Coombes, 2021). Das mittlere Erkrankungsalter liegt dabei um das 20. Lebensjahr (O'Connell und Coombes, 2021), wobei ein niedrigeres Erkrankungsalter mit einer schlechteren Prognose assoziiert ist (Joslyn et al., 2016). Die Erkrankung besteht häufig ein Leben lang (Budde et al., 2017). Laut dem Bericht des *Global Burden of Disease* gehört die Bipolare Störung zu den 30 führenden Gründen für eine Einschränkung im Alltag (GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016). Zusammen mit der Unipolaren Depression wird die Bipolare Störung zu den affektiven Störungen gezählt, ein wesentliches Unterscheidungsmerkmal besteht dabei in dem wiederholten Wechsel von (hypo-)manischen Episoden und depressiven oder euthymischen Phasen, der die Bipolare Störung typischerweise kennzeichnet (O'Connell und Coombes, 2021). Die Krankheit kann klinisch in die Bipolar-I-Störung und die Bipolar-II-Störung unterteilt werden (American Psychiatric Association, 2013). Bei der Bipolar-I-Störung besteht häufig ein Wechsel zwischen depressiven und manischen Phasen. Die Bipolar-II-Störung hingegen ist durch das Auftreten von mindestens einer depressiven und einer hypomanischen Phase (Episoden einer mildereren Ausprägung der Manie) im Leben charakterisiert, ohne dass eine manische Episode auftritt (American Psychiatric Association, 2013, s. Abb. 1, S. 10). Klinische und diagnostische Parallelen bestehen bei der Bipolaren Störung laut DSM-V (*Diagnostic and Statistical Manual of Mental Disorders V*, Diagnostisches und Statistisches Manual Psychischer Störungen V, American Psychiatric Association, 2013) zu der Schizoaffectiven Störung, bei der Betroffene sowohl unter psychotischen Symptomen als auch unter manischen oder depressiven Phasen leiden können (American Psychiatric Association, 2013). Darüber hinaus bestehen bei der Bipolaren Störung häufig Komorbiditäten zu anderen psychiatrischen Erkrankungen wie verschiedenen Angststörungen (Yapici Eser et al., 2018), der Borderline-Persönlichkeitsstörung (Frías et al., 2016) oder Alkohol- und Drogenmissbrauch (Salloum und Brown, 2017).

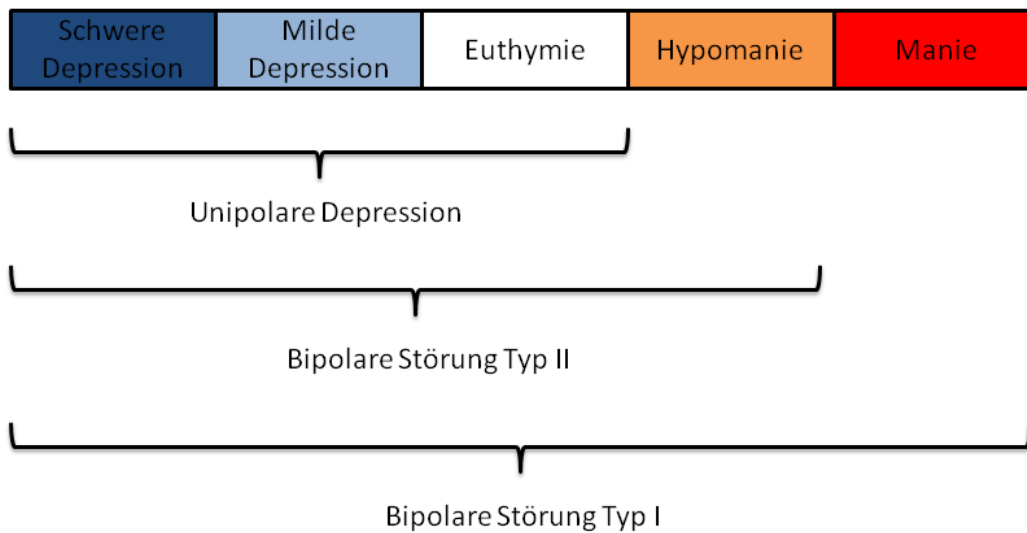


Abb. 1: Symptome der affektiven Störungen. Adaptiert von O'Connell und Coombes, 2021

1.2 Genetische Grundlagen

In der Humangenetik wird zwischen dem Genotypen, der Gesamtheit der Erbinformation, und dem Phänotypen, der klinischen Erscheinungsform, unterschieden (Hoffmeister, 2022a). Träger der Erbinformation sind dabei insbesondere die im Zellkern des Menschen befindlichen Chromosomen, die aus Desoxyribonukleinsäure (*Deoxyribonucleic Acid*, DNA) und unterschiedlichen Proteinen bestehen (Hoffmeister, 2022a). Das Genom eines Menschen besteht typischerweise aus 23 Chromosomenpaaren, von denen 22 Paare sogenannte Autosomen (= Körperchromosomen) und ein Paar sogenannte Gonosomen (= Geschlechtschromosomen, XY für den männlichen Phänotyp oder XX für den weiblichen Phänotyp) darstellen. Jeweils ein Chromosom pro Paar stammt aus dem väterlichen Erbgut, und eines aus dem mütterlichen Erbgut (Hoffmeister, 2022a). Ein Gen stellt einen Abschnitt der DNA dar, der für die Ausprägung eines bestimmten Merkmals verantwortlich ist (Hoffmeister, 2022a). Ist ein Gen auf beiden Chromosomen gleich ausgeprägt, spricht man von Homozygotie, das Individuum besitzt in diesem Fall zwei gleiche sogenannte Allele. Befinden sich zwei verschiedene Ausprägungen (= Allele) desselben Gens auf dem entsprechenden Ort der Chromosomen, ist das Individuum heterozygot für das Merkmal und besitzt dementsprechend zwei verschiedene Allele (Hoffmeister, 2022b; Hoffmeister, 2022c). Der Vererbungsmodus der Allele kann "rezessiv" oder "dominant" sein (Hoffmeister, 2022b). Ein dominantes Allel setzt sich bei Vorliegen zweier verschiedener

Allele durch, das rezessive Allel wird dabei durch das dominante Allel überdeckt. Ein rezessives Allel kommt dementsprechend in der Regel nur bei Homozygotie zur phänotypischen Ausprägung (Hoffmeister, 2022b). Führen verschiedenartige Merkmalsausprägungen innerhalb eines Gens oder an verschiedenen Genorten (auch "Genloci") jeweils zum identischen Phänotyp, spricht man von Heterogenie oder Heterogenität (Spektrum Akademischer Verlag, 1999). Ob ein Phänotyp sich ausprägt, hängt auch von der sogenannten Penetranz ab: Sie gibt die Wahrscheinlichkeit an, mit der eine bestimmte genetische Veränderung zum Phänotypen führt (Giudicessi und Ackerman, 2013).

Der DNA-Strang besteht im Wesentlichen aus einer Abfolge von vier verschiedenen Basen (Adenin, Cytosin, Guanin und Thymin), die sich jeweils mit einer komplementären Base (Adenin zu Thymin, Cytosin zu Guanin) über Wasserstoffbrückenbindungen verbinden und so eine Doppelhelix bilden (Hoffmeister, 2022c). Als "Rückgrat" der Basen fungieren jeweils ein Kohlenhydratmolekül (Desoxyribose mit einer freien Hydroxidgruppe am 3'-Ende) und eine Phosphatgruppe am 5. Kohlenstoffatom der Desoxyribose (5'-Ende). Die Kombination aus Base, Kohlenhydrat und Phosphat bezeichnet man als Nukleotid (Hoffmeister, 2022c). Ein neues Nukleotid kann immer nur an einem freien 3'-OH-Ende angeknüpft werden, beispielsweise während der Amplifikation der DNA mittels Polymerasekettenreaktion (*Polymerase Chain Reaction*, PCR). Ein 3'-OH (Hydroxid)-Ende und ein 5'-P (Phosphat)-Ende stehen sich dabei stets gegenüber, d.h. der Aufbau des DNA-Doppelstrangs ist antiparallel (Hoffmeister, 2022c).

Um die Basenabfolge der DNA in ein funktionsfähiges Protein zu übersetzen, wird die DNA in einem ersten Schritt mithilfe einer RNA (*Ribonucleic Acid*, Ribonukleinsäure)-Polymerase zunächst in mRNA (*messenger-RNA*) umgeschrieben. Dieser Vorgang wird als Transkription bezeichnet (Hoffmeister, 2022c). Zur Bildung der reifen mRNA werden nicht-kodierende Regionen (Introns) gezielt herausgeschnitten, sodass nur die kodierenden Regionen (Exons) verbleiben. Dieser Vorgang wird als Spleißen bezeichnet und geschieht an den Grenzen zwischen Exonen und Intronen, den sogenannten Spleißstellen (Hoffmeister, 2022c). Die fertige mRNA wird im Anschluss aus dem Zellkern zu den Ribosomen, dem Ort der eigentlichen Proteinsynthese, transportiert und dient dort als "Bauplan" für das Protein (Hoffmeister, 2022c). Im zweiten Schritt der Proteinbiosynthese, der Translation, wird die Basenabfolge der mRNA schließlich in eine Aminosäuresequenz

übersetzt. Drei aufeinanderfolgende Basen kodieren dabei jeweils für eine Aminosäure, mehrere Aminosäuren bilden schließlich das Protein (Hoffmeister, 2022c).

Veränderungen der Basenabfolge, die je nach Ausmaß zur Fehlfunktion des Proteins führen können, nennt man Mutationen bzw. Varianten. Bei den sogenannten Punktmutationen ist jeweils eine Base durch eine andere ausgetauscht (Hoffmeister, 2022d). Hierbei unterscheidet man zwischen stummen oder synonymen Mutationen, die keine Auswirkung auf die Ausbildung der Aminosäuresequenz haben, sogenannten *Missense*-Mutationen, die zum Einbau einer falschen Aminosäure in das Protein führen, und sogenannten Stoppmutationen, die die Übersetzung der Basensequenz in das Protein durch Einbau eines Stoppcodons frühzeitig beenden. Bei einer *Missense*-Mutation kann die Proteinfunktion verändert sein, bei einer Stoppmutation ist das gebildete Protein oft unvollständig und funktionslos (Hoffmeister, 2022d), alternativ wird die Bildung eines verkürzten Proteins durch den sogenannten *nonsense-mediated decay* verhindert (Lykke-Andersen und Jensen, 2015). Beim Verlust (Deletion) oder Hinzufügen (Insertion) von einer oder mehreren Basen kann das Leseraster während der Proteinbiosynthese verschoben sein (Rasterschub-Mutation). Es entsteht dabei eine komplett neue Aminosäuresequenz, die die Funktion des Proteinprodukts ebenfalls beeinflussen kann. Bei einer Insertion oder Deletion von genau drei Basen ist das Leseraster nicht gestört (*Inframe*-Mutation), es wird allerdings eine zusätzliche Aminosäure eingebaut bzw. es kommt zum Wegfall einer Aminosäure (Hoffmeister, 2022d).

Wie schädlich eine genetische Variante für die phänotypische Ausprägung ist, kann mithilfe von verschiedenen bioinformatischen Methoden dargestellt werden. Ein hilfreiches Programm ist beispielsweise PolyPhen2, welches die mögliche Auswirkung einer Aminosäuresubstitution auf die Struktur und die Funktion des Proteins vorhersagen kann (Adzhubei et al., 2013). Ein weiterer Parameter für die Einstufung von phänotypischer Relevanz stellt der CADD (*Combined Annotation Dependent Depletion*)-Score dar, der die funktionelle Relevanz von sowohl Punktmutationen als auch von Insertionen/Deletionen im menschlichen Genom anhand eines Wertes abbilden kann (Rentzsch et al., 2019).

Zusätzlich zu den kodierenden Regionen (exonische Regionen) besteht die menschliche DNA zu großen Teilen aus nichtkodierenden Regionen (intronische und intergenische Regionen). Aussagen zur funktionellen Relevanz von Varianten in diesen nichtkodierenden

Regionen sind schwierig, bisherige Studien konnten aber u. a. regulatorische Effekte auf die Methylierung und Expression benachbarter Gene zeigen (z. B. Schulz et al., 2017).

In den Jahren 1995 bis 2005 konnte man mittels DNA-Sequenziermethoden das menschliche Genom weitestgehend entschlüsseln (Hoffmeister, 2022c). Bei diesen Sequenziermethoden wird die DNA zunächst denaturiert. Mithilfe von verschiedenen Enzymen (Polymerase, Ligase, Exonuklease) können DNA-Abschnitte synthetisiert, zusammengeführt und auch abgeschnitten werden (Hoffmeister, 2022c). Als Startmolekül für die DNA-Synthese werden sogenannte DNA-Primer benötigt: Diese stellen kurze Basensequenzen dar, die zu einem Teil der Zielsequenz komplementär sind und als Startmoleküle für die Verlängerung des DNA-Strangs dienen (Hoffmeister, 2022c). Nach Zugabe der benötigten Enzyme und der einzelnen Basen entsteht ein zur Basensequenz komplementärer DNA-Strang, der auf die ursprüngliche Sequenz schließen lässt (Hoffmeister, 2022c).

Eine konventionelle Sequenziermethode stellt beispielsweise die Kettenabbruchmethode nach Sanger dar (Hoffmeister, 2022c, s. Abb. 2, S. 14). Hier werden der Reaktion modifizierte Basen zugegeben, die jeweils keine weitere Basenbindung zulassen. Das Ergebnis sind verschieden lange DNA-Fragmente, die anschließend in einer Gelelektrophorese sichtbar gemacht werden können (Hoffmeister, 2022c).

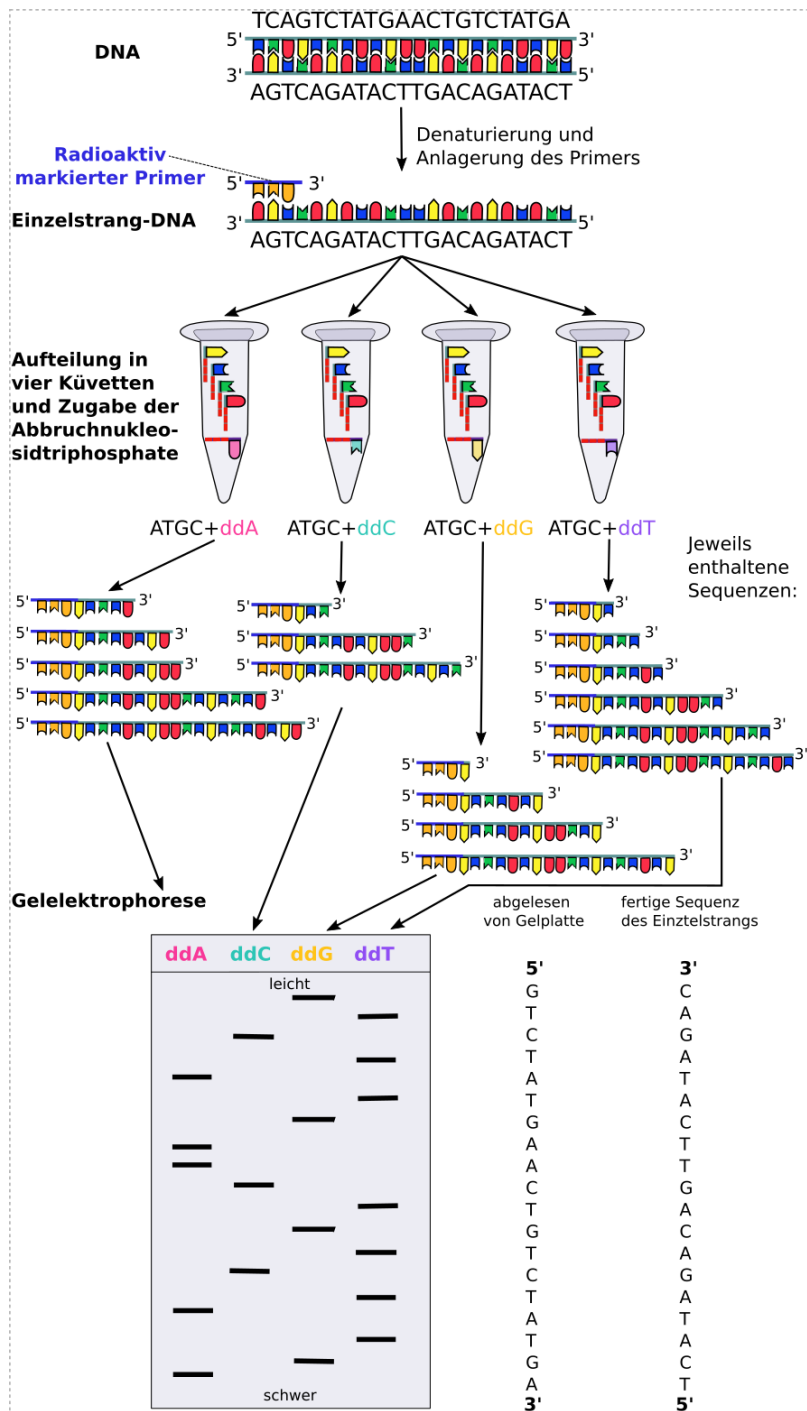


Abb. 2: Sequenziermethode nach Sanger.

Aus: Hoffmeister, 2022: Freies Lehrbuch: Einführung in die Biologie (für die Sekundarstufe 1 und 2). Kapitel 08.12: Molekulare Grundlagen der Vererbung (Molekulargenetik).

Abgedruckt nach der GNU-Lizenz für freie Dokumentation

(<https://www.gnu.org/licenses/fdl-1.3.txt>).

ddA: DiDesoxyAdenosintriphosphat, ddC: DiDesoxyCytidintriphosphat, ddG: DiDesoxyGuanosintriphosphat, ddT: DiDesoxyThymidintriphosphat

Für die Sequenzierung von großen Patienten- und Kontrollkollektiven bieten sich allerdings *Next-Generation-Sequencing* (NGS)-Technologien an (Behjati und Tarpey, 2013). Im Gegensatz zur konventionellen Sanger-Sequenzierung besitzen diese eine hohe Sequenzierkapazität und -geschwindigkeit. Die in dieser Arbeit verwendete innovative smMIPs- (*single molecule Molecular Inversion Probes*)-Technologie ist eine am Institut für Humangenetik des Universitätsklinikums Bonn etablierte Sequenziermethode und stellt durch die parallele Bearbeitung von einer großen Anzahl an DNA-Proben eine sehr effiziente und kostengünstige Möglichkeit der Resequenzierung dar. Hierdurch wird ebenfalls eine besonders hohe diagnostische Sensitivität erreicht (Hiatt et al., 2013).

1.3 Stand der Forschung

Die molekularbiologischen Vorgänge, die zur Entstehung der Bipolaren Störung beitragen, sind bisher weitestgehend ungeklärt (Forstner et al., 2020b). Bekannt ist allerdings, dass der Bipolaren Störung eine multifaktorielle Genese zugrunde liegt, bei der sowohl Umweltfaktoren als auch ein genetischer Hintergrund zur Pathophysiologie der Erkrankung beitragen (Craddock und Sklar, 2013). Aus den bisher durchgeführten formalgenetischen Studien ist bekannt, dass der Anteil des genetischen Beitrags bei der Bipolaren Störung auf bis zu 85 % geschätzt wird (Bienvenu et al., 2011; Budde et al., 2017; Lee et al., 2011). Von der Identifizierung genetischer Varianten bei der Bipolaren Störung erhofft man sich vor allem, die der Erkrankung zugrundeliegenden molekularen Mechanismen näher zu verstehen. Dies könnte dazu beitragen, mögliche Zielstrukturen für Medikamente zu identifizieren, um Patienten eine möglichst individualisierte Therapie anbieten zu können (Goes, 2016).

Durchbrüche in der Analyse der genetischen Grundlagen der Bipolaren Störung erzielte man mit sogenannten genomweiten Assoziationsstudien (*Genome-Wide Association Studies*, GWAS), in welchen man genomweit die Unterschiede zwischen Erkrankten und Gesunden auf Ebene von Einzelbasenpolymorphismen (*Single Nucleotide Polymorphisms*, SNPs) untersucht. Die SNPs stellen häufig (MAF (*Minor Allele Frequency*) > 1 %) im Genom vorkommende Veränderungen dar (Cichon et al., 2011; Mühleisen et al., 2014). Eine große GWAS des *Psychiatric Genomics Consortiums* (PGC), zu der Mitarbeiter des Instituts für Humangenetik des Universitätsklinikums Bonn beigetragen haben, fasste im Jahr 2019 Daten von insgesamt etwa 29.000 Patienten mit Bipolarer Störung und etwa 169.000

gesunden Kontrollpersonen zusammen. In dieser Studie wurden insgesamt 30 genomweit signifikante Genloci für die Bipolare Störung identifiziert, darunter 20 vorher unbekannte Loci (Stahl et al., 2019). In einer weiteren GWAS von Mullins et al. (2021) konnten in einer Untersuchung von über 40.000 Patienten und über 300.000 Kontrollpersonen insgesamt 64 assoziierte genomische Loci identifiziert werden. In dieser Studie ergab sich für die Bipolare Störung eine SNP-basierte Heritabilität von 18,6 % (Mullins et al., 2021).

Studien belegen weiterhin, dass die (meist in nicht-kodierenden Regionen befindlichen) Einzelbasenpolymorphismen Einfluss auf bestimmte Stoffwechselwege ausüben könnten. Unter anderem konnten Zusammenhänge mit der Insulinsekretion, dem Endocannabinoidsignalweg (Stahl et al., 2019), der Kalziumsignalübertragung (Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011), dem glutamatergen System (Nurnberger et al., 2014), der neuronalen Entwicklung und kleinen, nichtkodierenden Ribonukleinsäuren (microRNAs) nachgewiesen werden (Forstner et al., 2015; Mühleisen et al., 2018; Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium, 2015).

Die in den GWAS nachgewiesenen häufigen Varianten sind durch eher kleine Effektstärken gekennzeichnet und erklären lediglich etwa 20-30 % der phänotypischen Varianz der Bipolaren Störung (Mullins et al., 2021; Stahl et al., 2019). Die Identifizierung von seltenen Varianten (MAF <1 %) in proteinkodierenden Genen ist daher insofern ebenfalls interessant, da diese eventuell eine höhere Penetranz aufweisen und so ein erhöhtes individuelles Erkrankungsrisiko vermitteln könnten (Budde et al., 2017). Aufgrund der größeren Effektstärke eignen sich diese Varianten besonders gut für nachfolgende funktionelle Analysen wie z. B. Untersuchungen im zellulären System oder in Modellorganismen (Budde et al., 2017).

Besteht bei einer seltenen Variante ein Austausch von lediglich einer Base, spricht man von einer Einzelbasenvariante (*Single Nucleotide Variant*, SNV) (Budde et al., 2017). Auch Deletionen und Insertionen können seltene Varianten darstellen, man spricht von kleineren InDels oder größeren Kopienzahlvarianten (*Copy Number Variants*, CNVs), die etwa eine Kilobase bis mehrere Megabasenpaare lang sein können und häufig mehrere Gene betreffen (Budde et al., 2017). Seltene CNVs konnten beispielsweise bereits von Marshall et al. (2017) als Risikofaktoren für die Schizophrenie beschrieben werden. Für

die Bipolare Störung ergab sich in einer durchgeführten Metaanalyse von Green et al. (2016) für drei CNVs eine Assoziation, nach Korrektur für multiples Testen verblieb allerdings lediglich für die Duplikation am chromosomalen Locus 16p11.2 ein signifikantes Signal. In einer neueren Studie von Toma et al. (2018) mit 15 mehrfach betroffenen Familien wurden 17 seltene CNVs mit einer potentiellen Beteiligung an der Pathogenese der Bipolaren Störung identifiziert.

Auf der Suche nach seltenen SNVs bei Patienten mit Bipolarer Störung, die ein individuelles Erkrankungsrisiko vermitteln könnten, wurden Sequenzierstudien aller proteinkodierenden Abschnitte des Genoms (*Whole Exome Sequencing*, WES) sowohl in großen Fall-Kontroll-Kohorten (Palmer et al., 2022) als auch in mehrfach betroffenen Familien (z.B. in Forstner et al., 2020a; Maaser et al., 2018) durchgeführt. Beispielsweise ließen sich in der 2020 von Forstner et al. veröffentlichten WES-Studie bei jeweils drei betroffenen Patienten aus 27 mehrfach betroffenen Familien aus Spanien und Deutschland insgesamt 378 seltene, nichtsynonyme, potentiell funktionelle Varianten in 368 verschiedenen Genen identifizieren. 19 dieser Gene waren bereits in genomweiten Assoziationsstudien oder anderen Sequenzierstudien als Risikogene für die Bipolare Störung vorbeschrieben (Forstner et al., 2020a).

Eine große WES-Studie aus dem Jahr 2022 mit Daten von 13.933 Patienten mit Bipolarer Störung und 14.422 Kontrollprobanden aus dem *Bipolar-Exome-Projekt* (BipEx, Palmer et al., 2022) konnte die Rolle von seltenen Varianten als signifikanter Risikofaktor für die Bipolare Störung deutlich machen: Der exomweite Vergleich offenbarte eine nominal signifikante Anreicherung ($P=0,024$) von seltenen (Allelfrequenz ca. $<0,01\%$), "damaging" *Missense*-Varianten bei Patienten mit Bipolarer Störung im Verhältnis zu Kontrollpersonen (Palmer et al., 2022). Exomweite Signifikanz konnte in dieser Studie von keinem Gen erreicht werden, das stärkste Signal ergab sich für die Bipolare Störung für das Gen *AKAP11* ($P=1,15 \times 10^{-5}$). Das Produkt dieses im Gehirn exprimierten Gens scheint an der Wirkung von Lithium, einem bekannten gegen die Bipolare Störung eingesetzten Medikament, involviert zu sein (Palmer et al., 2022).

Diese Ergebnisse legen nahe, dass die Analyse weiterer Patienten- und Kontrollkollektive notwendig ist, um den Beitrag seltener Varianten zur Entstehung der Bipolaren Störung besser verstehen zu können.

1.4 Genetische Überlappungen bei psychiatrischen Erkrankungen

Interessant ist bei der Betrachtung von genetischen Varianten bei psychiatrischen Störungen, dass sich die Bipolare Störung genetische Varianten mit anderen psychiatrischen Erkrankungen teilt. In einer größeren Studie des *Brainstorm Consortiums* ließen sich geteilte genetische Hintergründe mehrerer psychiatrischer und neurologischer Erkrankungen in Bezug auf häufige Varianten nachweisen (Brainstorm Consortium et al., 2018). Eine hohe Korrelation besteht beispielsweise zwischen der Bipolaren Störung und der Schizophrenie (70 %) und eine moderate Korrelation zwischen der Bipolaren Störung und der Unipolaren Depression (35 %) (Cross-Disorder Group of the Psychiatric Genomics Consortium et al., 2019; Stahl et al., 2019).

In der Kombination der großen BipEx-Studie von Palmer et al. (2022) mit SCHEMA (*Schizophrenia Exome Sequencing Meta-Analysis*), einem Datensatz mit WES-Daten von etwa 24.000 Patienten mit Schizophrenie (Singh et al., 2022), ließ sich *AKAP11* als definitives Risikogen für sowohl die Bipolare Störung als auch die Schizophrenie identifizieren (Palmer et al., 2022). Diese Beobachtung könnte Ausdruck der geteilten klinischen Symptomatik sein, da beispielsweise psychotische Phasen einen substantiellen Teil der Bipolaren Störung darstellen können, ebenso ist die Schizophrenie teilweise durch manische Phasen charakterisiert (American Psychiatric Association, 2013).

1.5 Fragestellung dieser Arbeit

Die vorliegende Arbeit beschäftigte sich mit der Fragestellung, ob sich an den 30 genomweit signifikanten Loci der GWAS von Stahl et al. (2019) seltene, hochpenetrante funktionell relevante Varianten befinden, die zur Entstehung der Bipolaren Störung beitragen. Da die am stärksten assoziierten Varianten dieser Loci hauptsächlich intergenisch bzw. intronisch liegen, ist die zentrale Hypothese dieser Arbeit, dass sich mögliche hochpenetrante Varianten in proteinkodierenden Genen befinden, die in unmittelbarer Nähe zu den

genomweit signifikanten Loci lokalisiert sind und/oder bei denen ein möglicher funktioneller Bezug zur Bipolaren Störung besteht. Dazu wurden die 30 Loci mithilfe von verschiedenen Datenbanken systematisch untersucht, um die vielversprechendsten Kandidatengene zu identifizieren. In einigen der implizierten Kandidatengene konnten bereits seltene Varianten für die Bipolare Störung (z.B. in *ANK3* in Toma et al., 2021) oder für andere psychiatrische Erkrankungen (z.B. in *GRIN2A* in Singh et al., 2022) identifiziert werden. Die 20 implizierten Kandidatengene wurden anschließend in über 1.000 Patienten mit Bipolarer Störung und über 1.000 gesunden Kontrollpersonen mithilfe von effizienten *Next Generation Sequencing* (NGS)- Technologien resequenziert und auf das Vorhandensein von seltenen, potentiell funktionell relevanten Varianten untersucht.

2. Material und Methoden

2.1 Auswahl der DNA-Proben

Die Auswahl an Proben für die Resequenzierung umfasste 1.052 Patienten mit Bipolarer Störung und 1.247 gesunde Kontrollpersonen deutscher Ethnizität.

Die Patientenrekrutierung wurde am Zentralinstitut für Seelische Gesundheit Mannheim, Universität Heidelberg, der Klinik für Psychiatrie und Psychotherapie des Universitätsklinikums Marburg, der Klinik für Psychische Gesundheit des Universitätsklinikums Münster und kooperierenden Kliniken durchgeführt. Unter den insgesamt 1.052 Patienten befanden sich 473 männliche (45,0 %) und 579 (55,0 %) weibliche Probanden. Alle Patienten besaßen zum Zeitpunkt der Rekrutierung eine eindeutige DSM-IV-Diagnose der Bipolaren Störung.

Das Kollektiv an Kontrollpersonen umfasste 1.247 Personen, davon 772 aus der Heinz Nixdorf Recall Study (Schmermund et al., 2002) und 475 aus der Forschergruppe 2107 (Kircher et al., 2019). Bei den Kontrollpersonen aus der Forschergruppe 2107 wurde das Vorhandensein einer psychischen Störung im Vorhinein ausgeschlossen. Das Geschlechterverhältnis der Kontrollpersonen (578= 46,4 % männlich und 669= 53,6 % weiblich) entsprach ungefähr dem der Patienten (s.o.). Das Durchschnittsalter der Kontrollpersonen betrug zum Zeitpunkt der Rekrutierung 49,9 Jahre bei einer Spannweite von 18 bis zu 75 Jahren.

Alle Teilnehmer und Teilnehmerinnen gaben ihr schriftliches Einverständnis. Die Studie wurde durch die Ethikkommissionen der beteiligten Institutionen genehmigt. Die Untersuchungen wurden auf Grundlage der revidierten Deklaration von Helsinki des Weltärztebundes (1983) und den entsprechenden gesetzlichen Grundlagen durchgeführt (World Medical Association, 2023).

2.2 Auswahl der Kandidatengene

Um die vielversprechendsten Gene der 30 Genloci aus der GWAS von Stahl et al. (2019) zu identifizieren, erfolgte eine detaillierte Analyse der besagten Genloci mithilfe von diversen Datenbanken. In Abbildung 3 (S. 22) sind die 30 Genloci der GWAS (Stahl et al.,

2019) inklusive ihrer *Lead SNPs* in einem sogenannten *Manhattan Plot* dargestellt. Der *Lead SNP* ist der SNP eines Locus mit dem jeweils niedrigsten P-Wert, also der SNP, der die stärkste Assoziation zeigt (s. Abb. 3, S. 22). Jeder SNP besitzt üblicherweise eine SNP-ID beginnend mit "rs" (*Reference SNP*).

Acht genomweit signifikante Varianten der GWAS sind sogenannte InDels, also Insertionen oder Deletionen, bei denen in der Analyse durch das PGC keine SNP-ID angegeben wurde. Mithilfe der Datenbank *Genome Aggregation Database* (gnomAD, <https://gnomad.broadinstitute.org/>, Version v2.0.2) konnte fünf der acht InDels eine SNP-ID zugeordnet werden.

Die Priorisierungsstrategie der vorliegenden Studie war auf proteinkodierende Gene fokussiert, da seltene Varianten, die in diesen Genen liegen, eine höhere Penetranz aufweisen und sich damit für funktionelle Folgestudien besser eignen könnten (Budde et al., 2017). Nichtkodierende Gene wie lncRNAs, microRNAs, Pseudogene oder Antisense-Gene wurden daher nicht berücksichtigt. In einem ersten Schritt wurde mithilfe des UCSC (*University of California, Santa Cruz*) Genome Browsers (Kent et al., 2002, hg19) untersucht, welche proteinkodierenden Gene in den Regionen im Bereich von -200.000 Basenpaaren (BP) und +200.000 Basenpaaren um den jeweiligen *Lead SNP* liegen. Diese Analyse ergab zunächst für alle 30 Loci insgesamt 161 proteinkodierende Gene. Um weiter zu priorisieren, wählten wir die Gene für die Resequenzierung aus, die im Bereich von -200.000 BP und +200.000 BP um den jeweiligen *Lead SNP* die einzigen proteinkodierenden Gene darstellten. Dies traf für die acht Gene *ADCY2*, *ANK3*, *CACNA1C*, *FSTL5*, *GRIN2A*, *RIMS1*, *SHANK2* und *THSD7A* zu (s. Abb. 3, S. 22 und Tab. 1, S. 23-26).

Um weitere Gene auszuwählen, wurden diese in einem zweiten Schritt auf einen möglichen, funktionellen Bezug zur Bipolaren Störung untersucht. Dafür wurde die GTEx-Datenbank (*Genotype-Tissue Expression*, GTEx Consortium, 2013) genutzt. Die Proben des GTEx-Projekts wurden auf sogenannte eQTL-Effekte (*expression Quantitative Trait Loci*), also auf Effekte der genetischen Varianten auf die Genexpression, in Bezug auf insgesamt 53 Organgewebe getestet. Für die weitere Priorisierung von Kandidatengenen in der vorliegenden Studie wurde nun untersucht, ob die *Lead SNPs* der 30 Loci bereits bekannte eQTL-Effekte in verschiedenen Geweben des zentralen Nervensystems (d. h.

Amygdala, Vorderer Cingulärer Cortex, Nucleus Caudatus, Cortex, Frontaler Cortex, Cerebellum, Cerebelläre Hemisphäre, Hippocampus, Hypothalamus, Nucleus Accumbens, Putamen, Rückenmark, Substantia Nigra) besitzen. Diese Untersuchung ergab für insgesamt 10 der 30 Loci Ergebnisse für insgesamt 13 proteinkodierende Gene (*ADD3*, *ASB16*, *CACNA1C*, *CDAN1*, *CILP2*, *GLYCTK*, *GNL3*, *HAPLN4*, *LMAN2L*, *NMB*, *PACS1*, *PLEKHO1* und *PPM1M*).

Da das Gen *CACNA1C* das einzige proteinkodierende Gen im Bereich von +/-200.000 BP um den *Lead SNP* rs10744560 ist und sich zusätzlich ein eQTL-Effekt im Gewebe des zentralen Nervensystems fand, ergaben sich durch die zwei o. g. Strategien an den 30 Genloci insgesamt 20 Kandidatengene, die für die Resequenzierung ausgewählt wurden (s. Tab. 1, S. 23-26).

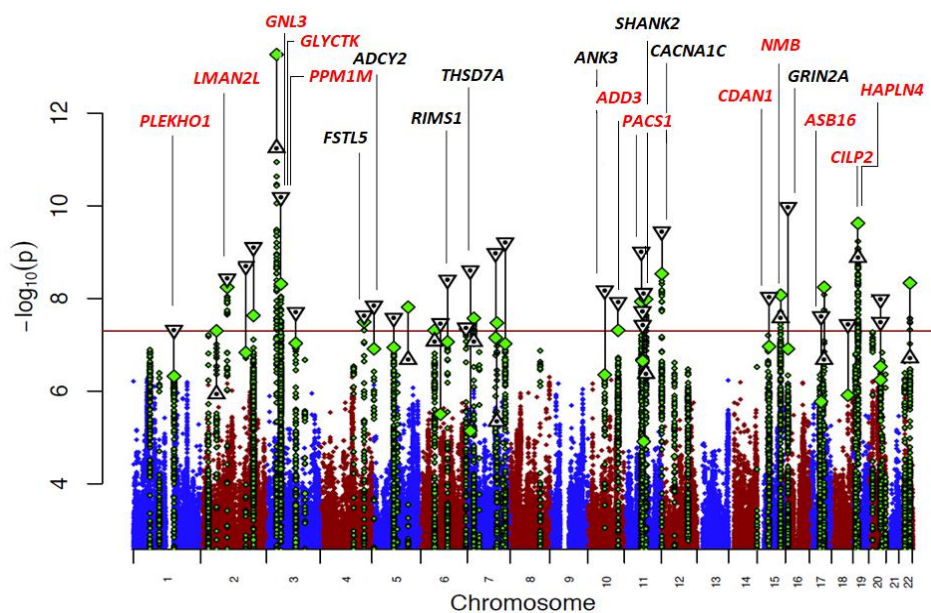


Abb. 3: Manhattan Plot der GWAS aus 2019, adaptiert von Stahl et al., 2018*, mit Darstellung der in der vorliegenden Studie ausgewählten Kandidatengene.

Die x-Achse zeigt die genomische Lokalisation, geordnet nach chromosomaler Position (aufsteigend), die y-Achse bildet den negativ dekadischen Logarithmus des P-Werts ($-\log_{10}(p)$) ab. Die rote Linie kennzeichnet die Grenze für genomweite Signifikanz ($P < 5 \times 10^{-8}$). Die einzigen proteinkodierenden Gene im Bereich von +/- 200.000 BP um den *Lead SNP* sind in schwarz, Gene mit eQTL-Effekt des Lead SNPs in Geweben des zentralen Nervensystems in rot dargestellt. *CACNA1C* wurde in beiden Ansätzen priorisiert und in der Abbildung aus graphischen Gründen in schwarzer Schrift dargestellt.

*Abgedruckt gemäß *CC-BY 4.0 International license*

(Creative Commons, <https://creativecommons.org/licenses/by/4.0/>)

Tab. 1: Identifizierte Kandidatengene.

Locus= Nummer des Locus von 1 bis 30. *Lead SNP*= SNP des jeweiligen Locus mit dem niedrigsten P-Wert (Quelle: Stahl et al., 2019). Position= Position des *Lead SNPs* +/- 200.000 BP. Proteinkodierende Gene= Im *UCSC Genome Browser* identifizierte proteinkodierende Gene im Bereich von +/- 200.000 BP des *Lead SNPs*. Gene mit eQTL-Effekt= Im *GTEX* Portal identifizierte Gene mit eQTL-Effekt des *Lead SNPs* in Geweben des zentralen Nervensystems. Chr= Chromosom. In= Insertion. Del= Deletion. SNP= *Single Nucleotide Polymorphism*. Ausgewählte Gene sind fett dargestellt.

Locus	Lead SNP	Position	Proteinkodierende Gene	Gene mit eQTL-Effekt
1	rs7544145	chr1: 149.938.699- 150.338.699	<i>ANP32E, APH1A, C1orf54, CA14, CIART, MRPS21, OTUD7B, PLEKHO1, PRPF3, RPRD2, VPS45</i>	PLEKHO1 (Cerebellum)
2	chr2_97376407_In	chr2: 97.176.407- 97.576.407	<i>ANKRD23, ANKRD39, ARID5A, CNNM3, CNNM4, FAM178B, FER1L5, KANSL3, LMAN2L, SEMA4C</i>	LMAN2L (Cerebellum)
3	rs17183814	chr2: 165.952.389- 166.352.389	<i>CSRNP3, SCN2A, SCN3A</i>	
4	chr2_194465711_Del	chr2: 194.265.711- 194.665.711		
5	rs9834970	chr3: 36.656.030- 37.056.030	<i>DCLK3, EPM2AIP1, MLH1, TRANK1</i>	
6	rs2302417	chr3: 52.614.256- 53.014.256	<i>GLT8D1, GNL3, ITIH1, ITIH3, ITIH4, MUSTN1, NEK4, PBRM1, SFMBT1, SPCS1, STIMATE, STIMATE-MUSTN1</i>	GLYCTK (Cerebellum) GNL3 (Cerebelläre Hemisphäre, Cerebellum) PPM1M (Cerebellum)

Locus	Lead SNP	Position	Proteinkodierende Gene	Gene mit eQTL-Effekt
7	rs3804640	chr3: 107.593.709- 107.993.709	<i>CD47, IFT57</i>	
8	rs11724116	chr4: 162.094.038- 162.494.038	<i>FSTL5</i>	
9	chr5_7587236_Del	chr5: 7.387.236- 7.787.236	<i>ADCY2</i>	
10	rs10035291	chr5: 80.596.368- 80.996.368	<i>ACOT12, SSBP2, ZCCHC9</i>	
11	chr6_72519394_Del	chr6: 72.319.394- 72.719.394	<i>RIMS1</i>	
12	rs2388334	chr6: 98.391.622- 98.791.622		
13	rs10455979	chr6: 166.795.260- 167.195.260	<i>MPC1, RPS6KA2</i>	
14	rs113779084	chr7: 11.671.787- 12.071.787	<i>THSD7A</i>	
15	rs73188321	chr7: 104.848.158- 105.248.158	<i>ATXN7L1, EFCAB10, PUS7, RINT1, SRPK2</i>	
16	chr7_140700006_In	chr7: 140.500.006- 140.900.006	<i>BRAF, MRPS33, TMEM178B</i>	

Locus	Lead SNP	Position	Proteinkodierende Gene	Gene mit eQTL-Effekt
17	rs10994318	chr10: 61.925.856- 62.325.856	ANK3	
18	chr10_111745562_In	chr10: 111.545.562- 111.945.562	ADD3, XPNPEP1	ADD3 (Cerebelläre Hemisphäre, Cerebellum)
19	rs12226877	chr11: 61.391.907- 61.791.907	<i>BEST1, DAGLA, FADS1, FADS2, FADS3, FEN1, FTH1, MYRF, RAB3IL1, TMEM258</i>	
20	rs10896090	chr11: 65.745.186- 66.145.186	<i>B4GAT1, BANF1, BRMS1, CATSPER1, CD248, CNIH2, CST6, EIF1AD, GAL3ST3, KLC2, PACS1, RAB1B, RIN1, SART1, SF3B2, SLC29A2, TMEM151A, YIF1A</i>	PACS1 (Cortex, Frontaler Cortex)
21	rs7122539	chr11: 66.462.731- 66.862.731	<i>C11orf80, C11orf86, LRFN4, PC, RCE1, RHOD, SPTBN2, SYT12</i>	
22	rs12575685	chr11: 70.317.927- 70.717.927	SHANK2	
23	rs10744560	chr12: 2.187.099- 2.587.099	CACNA1C	CACNA1C (Cerebelläre Hemisphäre, Cerebellum)
24	rs4447398	chr15: 42.704.904- 43.104.904	CDAN1, HAUS2, LRRC57, SNAP23, STARD9, TTBK2, ZNF106	CDAN1 (Cerebellum)

Locus	Lead SNP	Position	Proteinkodierende Gene	Gene mit eQTL-Effekt
25	chr15_85357857_In	chr15: 85.157.857- 85.557.857	<i>ALPK3</i> , <i>NMB</i> , <i>PDE8A</i> , <i>SCAND2P</i> , <i>SEC11A</i> , <i>SLC28A1</i> , <i>WDR73</i> , <i>ZNF592</i> , <i>ZSCAN2</i>	<i>NMB</i> (Nucleus Caudatus)
26	rs11647445	chr16: 9.726.966- 10.126.966	<i>GRIN2A</i>	
27	rs112114764	chr17: 42.001.041- 42.401.041	<i>ASB16</i> , <i>ATXN7L3</i> , <i>C17orf53</i> , <i>G6PC3</i> , <i>HDAC5</i> , <i>HROB</i> , <i>LSM12</i> , <i>NAGS</i> , <i>PPY</i> , <i>PYY</i> , <i>RUNDC3A</i> , <i>SLC25A39</i> , <i>SLC4A1</i> , <i>TMEM101</i> , <i>TMUB2</i> , <i>UBTF</i>	<i>ASB16</i> (Cerebelläre Hemisphäre, Cerebellum)
28	rs11557713	chr18: 60.043.876- 60.443.876	<i>PHLPP1</i> , <i>TNFRSF11A</i> , <i>ZCCHC2</i>	
29	rs111444407	chr19: 19.158.207- 19.558.207	<i>ARMC6</i> , <i>BORCS8</i> , <i>BORCS8-MEF2B</i> , <i>GATAD2A</i> , <i>HAPLN4</i> , <i>MAU2</i> , <i>MEF2B</i> , <i>NCAN</i> , <i>NR2C2AP</i> , <i>RFXANK</i> , <i>SLC25A42</i> , <i>SUGP1</i> , <i>TM6SF2</i> , <i>TMEM161A</i>	<i>CILP2</i> (Nucleus Caudatus, Nucleus Accumbens, Putamen) <i>HAPLN4</i> (Cerebellum)
30	chr20_43682549_In	chr20: 43.482.549- 43.882.549	<i>KCNS1</i> , <i>PABPC1L</i> , <i>PI3</i> , <i>SEMG1</i> , <i>SEMG2</i> , <i>SLPI</i> , <i>STK4</i> , <i>TOMM34</i> , <i>WFDC5</i> , <i>WFDC12</i> , <i>YWHAB</i>	

Über den UCSC Genome Browser (Kent et al., 2002) wurden weitere Informationen über die Funktion der exprimierten Proteine der ausgewählten Kandidatengene eingeholt (s. Tab. 2, S. 27-28).

Tab. 2: Kandidatengene inklusive Genprodukt und Funktion.
Genprodukt= Genprodukt des Gens. Funktion (UCSC)= Funktion des Genprodukts gemäß *UCSC Genome Browser*.

Gen	Genprodukt	Funktion (UCSC)
ADCY2	Adenylatcyklase 2	Bildung von cyclischem Adenosinmonophosphat (cAMP)
ADD3	Adducin 3	Calmodulin-bindendes Protein im Spektrin-Aktin-Netzwerk
ANK3	Ankyrin G	Integration von Ionenkanälen an Ranvier'schen Schnürringen
ASB16	Ankyrin-Wiederholungs- und SOCS- (<i>suppressor of cytokine signalling</i>) Box-enthaltendes Protein 16	Kopplung von SOCS-Proteinen mit dem Elongin B- und C-Komplex zur Proteindegradierung
CACNA1C	Spannungsabhängiger Kalziumkanal, Untereinheit Alpha 1C	Einstrom von Kalzium nach Membrandepolarisation
CDAN1	Codanin 1	Integrität der Kernhülle, Befestigung von Mikrotubuli
CILP2	Knorpelzwischen-schichtprotein 2	Protein des Knorpelgerüsts
FSTL5	Follistatinverwandtes Protein 5	Calciumbindung, Zelldifferenzierung

Gen	Genprodukt	Funktion (UCSC)
<i>GLYCTK</i>	Glyceratkinase	Phosphorylierung von Glycerat
<i>GNL3</i>	Guaninnukleotidbindendes Protein Like 3	Interaktion mit p53 (Tumorgenetik), Stammzellproliferation
<i>GRIN2A</i>	Ionotroper Glutamatrezeptor vom NMDA-Typ, Untereinheit 2A	Langzeitpotenzierung an exzitatorischen Synapsen
<i>HAPLN4</i>	Hyaluron- und Proteoglycan- verbindendes Protein 4	Formung der Extrazellulären Matrix
<i>LMAN2L</i>	Lectin- und Mannose-bindendes Protein 2 Like	Glykoproteintransporter am Endoplasmatischen Retikulum
<i>NMB</i>	Neuromedin B	Regulation der Kolonmuskulatur
<i>PACS1</i>	Phosphofurinsäure-Cluster sortierendes Protein 1	Transportmolekül am Golgi-Apparat
<i>PLEKHO1</i>	Pleckstrin-Homologie-Domäne, Familie O, Mitglied 1	Interaktion mit Proteinkinase CK2
<i>PPM1M</i>	Magnesiumabhängige Proteinphosphatase 1M	Phosphoprotein-Phosphatase-Aktivität
<i>RIMS1</i>	Exozytose-regulierendes Synaptisches Membranprotein 1	Regulation der Exozytose von synaptischen Vesikeln und spannungsabhängiger Calciumkanäle
<i>SHANK2</i>	SH3 und Multiple Ankyrinwiederholungs-Domänen 2	Postsynaptische Dichte von exzitatorischen Synapsen
<i>THSD7A</i>	Thrombospondin-Typ-1-Domäne 7A	Hemmung von Endothelzellmigration

2.3 Resequenzierung mittels *Molecular-Inversion-Probes*-Methode

Für die Resequenzierung der Patienten und Kontrollpersonen wählten wir die MIPs-Methode, eine am Institut für Humangenetik des Universitätsklinikums Bonn etablierte NGS-Technologie, die die parallele Sequenzierung von einer großen Anzahl an DNA-Proben ermöglicht (Hiatt et al., 2013).

Ein "MIP" stellt bei dieser Methode ein ca. 70 Basen langes einzelsträngiges DNA-Molekül dar (s. Abb. 4, S. 29). Jeweils am Anfang und am Ende des Moleküls befinden sich ca. 20 Basen lange Regionen, die sogenannten Zielarme (Extensions- und Ligationsarm). Diese Regionen stellen kurze Basensequenzen dar, die zu einem Teil der Ziel-DNA komplementär sind und als Startmoleküle für die Verlängerung des DNA-Strangs dienen. Die dazwischenliegende 30 Basen lange Linker-Sequenz ist zwischen allen MIPs identisch und beinhaltet jeweils zwei universelle PCR-Primer-Bindungsstellen. Bei der von uns angewandten smMIPs-Option (*sm= single molecule*) enthält die Linker-Sequenz zusätzlich fünf degenerierte Basen und ist somit 35 Basen lang. Diese fünf degenerierten Basen dienen als molekulare Marker und werden mitsequenziert. Sequenzen, die denselben molekularen Marker tragen, werden in eine einzige Sequenz zusammengefasst, sodass Konsensussequenzen entstehen und PCR-Artefakte reduziert werden.

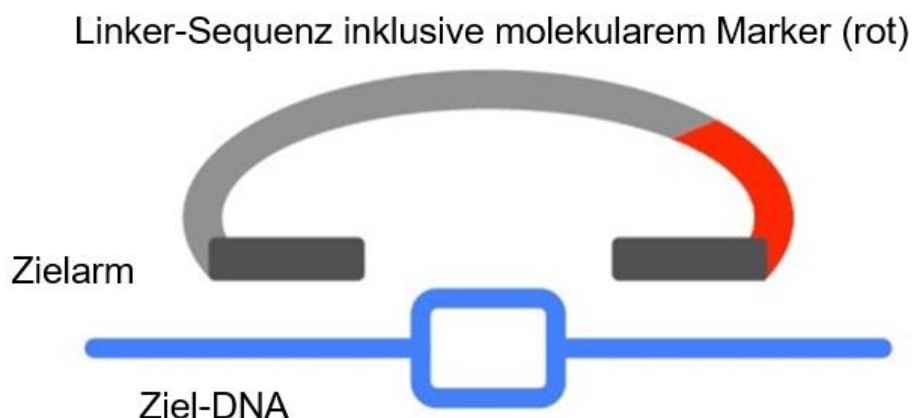


Abb. 4: Aufbau eines smMIP-Primers. Adaptiert von: Hiatt et al., 2013**

** Abgedruckt gemäß CC BY-NC 3.0

(Creative Commons, <https://creativecommons.org/licenses/by-nc/3.0/>)

Mithilfe der MIPgen-Software (Boyle et al., 2014) wurden die Sequenzen der smMIP-Primer für unsere 20 Kandidatengene erstellt. Die Zielsequenzen stellen dabei die einzelnen Exone der Gene dar. Beim Design der smMIPs wurden die Exongrenzen um jeweils sechs Basenpaare verlängert, um Spleißstellen abzudecken. Varianten in solchen Regionen können Spleißvorgänge beeinträchtigen und dadurch die Funktion des zu bildenden Proteins beeinflussen (Ward und Cooper, 2010), weshalb sie ebenfalls in die zu sequenzierende Zielregion mit aufgenommen wurden.

Tauchten häufige Einzelbasenvarianten, sogenannte *single nucleotide polymorphisms* (SNPs, definiert als Allelfrequenz >1 % in der Europäischen Nicht-Finnischen Population in gnomAD, <https://gnomad.broadinstitute.org/>, Version v2.0.2) in den Zielarmen auf, wurde die betroffene Base in der smMIP-Sequenz entsprechend kodiert. Hierzu wurde für die betroffene Base eine mehrdeutige Kodierung (R= A/G; Y= C/T; M= A/C; K= G/T; S= C/G; W= A/T) benutzt, um anzuzeigen, dass an dieser Stelle in 50 % der smMIP-Sequenzen das Referenzallel und in 50 % der smMIP-Sequenzen das alternative Allel synthetisiert werden soll. Trat mehr als ein häufiger SNP oder eine Insertion/Deletion in einem der Zielarme auf, wurde die komplette smMIP-Sequenz neu erstellt. Diese Veränderungen sollten eine zuverlässige Bindung der Primer an die DNA gewährleisten. Nach Qualitätskontrolle und ggf. Neudesign der Primer wurden insgesamt 686 smMIPs mit einer durchschnittlichen Länge von 79 Basen (zwischen 75 und 80 Basen) bei *Integrated DNA Technologies* (IDT, Leuven, Belgien) bestellt.

Nach Erhalt der bestellten smMIP-Primer erfolgte dann zunächst die Erstellung von 20 Genpools. Dabei wurden jeweils 5 µl jedes smMIPs eines Genes zusammenpipettiert. Anschließend wurde aus den einzelnen Genpools ein Gesamt-smMIP-Pool erstellt, indem von jedem Genpool das 0,1-fache der jeweiligen smMIP-Anzahl (s. Tab. 3, S. 31) entnommen wurde und in ein Eppendorf-Gefäß zusammengeführt wurde. So konnte sichergestellt werden, dass jeder Primer in einfacher Konzentration vorhanden ist. Es ergab sich ein Gesamt-smMIP-Pool von 68,6 µl (s. Tab. 3, S. 31).

Tab. 3: Anzahl der smMIP-Primer pro Gen und Erstellung von Gen- und Gesamt-smMIP-Pool.

Anzahl smMIPs= Anzahl der smMIP-Primer pro Gen. Genpool= Volumen des gesamten Pools pro Gen (= 5 µl x Anzahl der smMIPs) in µl. Anteil am Gesamt-smMIP-Pool= Volumen des smMIP-Pools pro Gen (= 0,1 µl x Anzahl der smMIPs pro Gen) in µl.

Gen	Anzahl smMIPs	Genpool	Anteil am Gesamt-smMIP-Pool
<i>ADCY2</i>	37	185	3,7
<i>ADD3</i>	21	105	2,1
<i>ANK3</i>	113	565	11,3
<i>ASB16</i>	14	70	1,4
<i>CACNA1C</i>	76	380	7,6
<i>CDAN1</i>	42	210	4,2
<i>CILP2</i>	36	180	3,6
<i>FSTL5</i>	27	135	2,7
<i>GLYCTK</i>	14	70	1,4
<i>GNL3</i>	17	85	1,7
<i>GRIN2A</i>	37	185	3,7
<i>HAPLN4</i>	15	75	1,5
<i>LMAN2L</i>	11	55	1,1
<i>NMB</i>	6	30	0,6
<i>PACS1</i>	31	155	3,1
<i>PLEKHO1</i>	11	55	1,1
<i>PPM1M</i>	17	85	1,7
<i>RIMS1</i>	56	280	5,6
<i>SHANK2</i>	55	275	5,5
<i>THSD7A</i>	50	250	5,0
Insgesamt	686	3.430	68,6

Der nächste Schritt im Protokoll bestand in der Phosphorylierung des smMIP-Pools, da die im nachfolgenden Hybridisierungsschritt verwendete DNA-Ligase ein 5'-phosphoryliertes Ende benötigt. Dazu wurde der smMIP-Pool (68,6 µl) mit 2,7 µl T4 Polynukleotidkinase (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten), 8,5 µl 10x T4 DNA Ligase Puffer mit 10 mmol/l ATP (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten) und 5,2 µl Wasser phosphoryliert. Anschließend wurde er für 45 Minuten bei 37 °C und für 20 Minuten bei 65 °C im Thermocycler (Eppendorf, Hamburg, Deutschland) inkubiert und bei 4 °C gelagert.

In einem ersten Testlauf wurden sechs Kontroll-DNA-Proben mittels des Qubit 2.0 Fluorometer und des Qubit dsDNA-Assay Kits (Thermo Fisher Scientific, Waltham, Massachusetts, Vereinigte Staaten) einzeln vermessen und anschließend auf eine einheitliche Konzentration von 8 ng/µl gebracht.

Für den nächsten Schritt, die Hybridisierung, wurde der smMIP-Pool nun 1:900 verdünnt, um ein praktikables, pipettierbares Volumen für die Hybridisierungsmischung zu erhalten. Pro DNA-Probe wurden jeweils 0,34 µl des verdünnten smMIP-Pools, 2,5 µl 10x Amplicase DNA Ligase Puffer (Lucigen, Middleton, Wisconsin, Vereinigte Staaten), 0,03 µl dNTP Mix (0,25 mmol/l, NextTec, Leverkusen, Deutschland), 0,32 µl Hemo Klentaq DNA-Polymerase 10 U/µl (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten) und 0,01 µl Amplicase DNA-Ligase 100 U/µl (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten) sowie 9,3 µl Wasser benötigt. Diese Hybridisierungsmischung (12,5 µl) wurde jeweils zu 12,5 µl der DNA-Probe gegeben ($12,5 \mu\text{l} \times 8 \text{ ng}/\mu\text{l} = 100 \text{ ng DNA}$). Nach Denaturierung bei 95 °C für 10 Minuten erfolgte die Hybridisierung bei 60°C für 22,5 h, in denen die zielkomplementären Primerenden mit den Zielbereichen der DNA hybridisieren. Die DNA-Polymerase startete am 3'-OH-Ende der Primer und verlängerte den Primer um die komplementäre Sequenz der Zielregion mithilfe der dNTPs. Die DNA-Ligase ligierte die neu eingearbeitete Zielsequenz mit dem phosphorylierten 5'-Ende des Primers und es entstand ein ringförmiges Molekül mit der Primer- und Zielsequenz.

In einer anschließenden Exonuklease-Behandlung (direkt nach der Hybridisierung durchgeführt) wurden ungebundene smMIP-Sequenzen, welche linear vorliegen sowie einzelsträngige genomische DNA entfernt, sodass lediglich die ringförmigen smMIP- Moleküle

mit der nun integrierten Zielsequenz übrig blieben. Dazu wurden zu jedem hybridisierten Produkt 2 µl einer Exonukleasemischung hinzugegeben (jeweils 0,5 µl der Exonukleasen EXO I (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten) und EXO III (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten), 0,2 µl 10x Amplicase DNA Ligase Puffer (Lucigen, Middleton, Wisconsin, Vereinigte Staaten) sowie 0,8 µl Wasser). Anschließend erfolgte die Inkubation bei 37 °C für 45 Minuten und eine Exonuklease-Deaktivierung bei 90 °C für 2 Minuten.

In der darauffolgenden PCR-Reaktion sollte die in den smMIPs inkorporierte Zielregion amplifiziert werden. Für die PCR wurden Primer verwendet, die komplementär zu den zwei universellen PCR-Primer-Bindungsstellen in der Linkersequenz sind (jeweils *Forward* ("Vorwärts")- und *Reverse* ("Rückwärts")-Primer). Die *Forward*-Primer besitzen immer dieselbe Sequenz, die *Reverse*-Primer enthalten jeweils einen molekularen Barcode. Es standen 384 *Reverse*-Primer mit verschiedenen molekularen Barcodes zur Verfügung, sodass später bis zu 384 Proben in einem Pool sequenziert werden konnten. Die PCR wurde mit einem Volumen von 25 µl durchgeführt inklusive 5 µl des Exonuklease-behandelten smMIP-Produkts; 12,5 µl 2 x iProof HF Master Mix (Bio-Rad Laboratories, Hercules, Kalifornien, Vereinigte Staaten); 0,125 µl *Forward*-Primer (100 µmol/l, IDT, Leuven, Belgien) und 1,25 µl individuellem, mit Barcode markiertem *Reverse*-Primer (10 µmol/l, IDT, Leuven, Belgien) sowie 6,125 µl reinem Wasser. Es wurden mit Illumina-Sequenzern (San Diego, Kalifornien, Vereinigte Staaten) kompatible *Forward*- und *Reverse*-Primer benutzt.

Im ersten Testlauf sollte zusätzlich die optimale Zyklenanzahl für die PCR ermittelt werden. Dazu erfolgten gleichzeitig fünf PCR-Ansätze (PCR-Programm: 98 °C für 30 s; 24 x/26 x/28 x/30 x/32 x (jeweils 98 °C für 10 s, 60 °C für 30 s, 72 °C für 30 s); 74 °C für 2 min und 4 °C bis zur weiteren Verwendung). Die Produkte aller fünf PCR-Reaktionen wurden auf Agarosegel (2 % DNA-interkalierendes Ethidiumbromid enthaltend) aufgetragen (s. Abb. 5, S. 34). So konnte aufgrund der Banden im Agarosegelbild abgeschätzt werden, welche Zyklenanzahl für die Amplifikation optimal war. In allen fünf Reaktionen zeigte sich für jeweils alle sechs Proben eine eindeutige Produktbande, die das hybridisierte und amplifizierte Produkt darstellt. Die Banden aus dem PCR-Ansatz mit 24 Zyklen

zeigten am wenigsten Nebenprodukt und besaßen gleichzeitig die stärkste Intensität (s. Abb. 5, S. 34), sodass mit diesem PCR-Produkt fortgefahren wurde.

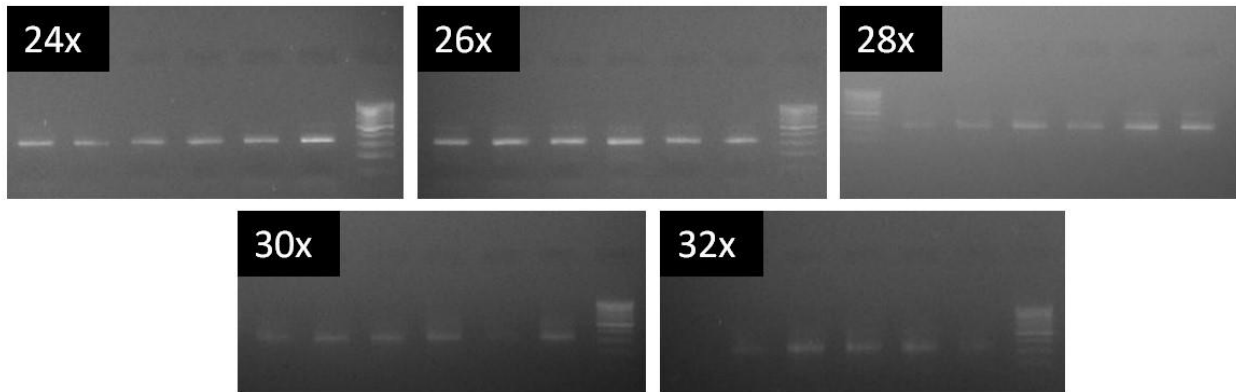


Abb. 5: Agarosegelbilder der fünf Test-PCR-Zyklen.

Zu sehen ist jeweils das PCR-Produkt von sechs DNA-Proben zwischen 300 und 400 bp neben der Leitersubstanz nach 24, 26, 28, 30 und 32 PCR-Zyklen.

Der nächste Schritt bestand in der Aufreinigung der PCR-Produkte durch AMPure XP Beads (Beckman Coulter, Brea, Kalifornien, Vereinigte Staaten). Dafür wurden von jeder DNA-Probe je 15 µl PCR-Produkt, bereits einen individuellen Barcode enthaltend, gepoolt. Diesem Pool wurden zunächst 5 µl entnommen, um diesen „unaufgereinigten“ Pool später mit dem aufgereinigten Pool zu vergleichen. Je nach Verunreinigungsgrad der DNA kann die Menge an AMPure XP Beads (Beckman Coulter, Brea, Kalifornien, Vereinigte Staaten) von 0,7-facher bis 1-facher Menge gewählt werden: Je mehr unspezifische Banden im Agarosegel zu sehen sind, desto weniger Beads werden eingesetzt. Aufgrund der vorher kaum sichtbaren unspezifischen Banden wurde ein 1:1-Verhältnis zwischen Poolvolumen und AMPure XP Beads (Beckman Coulter, Brea, Kalifornien, Vereinigte Staaten) gewählt (15 µl x 6 = 90 µl; 90 µl - 5 µl = 85 µl PCR-Produkt-Pool zu 85 µl AMPure XP Beads). Die Schritte der Aufreinigung waren wie folgt: PCR-Produkt-Pool und AMPure XP Beads in Eppendorf-Gefäß zusammenführen, vortexen, herunterzentrifugieren und 10 Minuten bei Raumtemperatur inkubieren, fünf Minuten auf dem Magneten inkubieren, Überstand abnehmen und verwerfen (PCR-Produkt ist an der magnetischen Substanz gebunden), Waschen mit 700 µl 70 % Ethanol und Eppendorf-Gefäß für 30 Sekunden invertieren, auf den Magneten zurückstellen und Überstand abnehmen, erneutes Waschen mit 700 µl

70 % Ethanol und Eppendorf-Gefäß für 30 Sekunden invertieren, restliches Ethanol abnehmen, Eppendorf-Gefäß offen für 5-10 Minuten auf dem Magneten stehen lassen, um die Beads zu trocknen, eluieren mit 25 µl Tris-EDTA (Ethylendiamintetraacetat)-Puffer (Thermo Fisher Scientific, Waltham, Massachusetts, Vereinigte Staaten), vortexen und zentrifugieren und eine Minute auf dem Magneten inkubieren. Schließlich wurde der aufgereinigte PCR-Produkt-Pool in ein frisches Eppendorf-Gefäß überführt.

Mit der TapeStation mit D1000 ScreenTapes (Agilent Technologies, Santa Clara, Kalifornien, Vereinigte Staaten) wurden nun der aufgereinigte und der unaufgereinigte Pool verglichen. Der aufgereinigte Pool zeigte weniger unspezifisches Nebenprodukt und die intensivste Bande zwischen 300 und 400 bp als Ausdruck des PCR-Produkts (s. Abb. 6, S. 35).

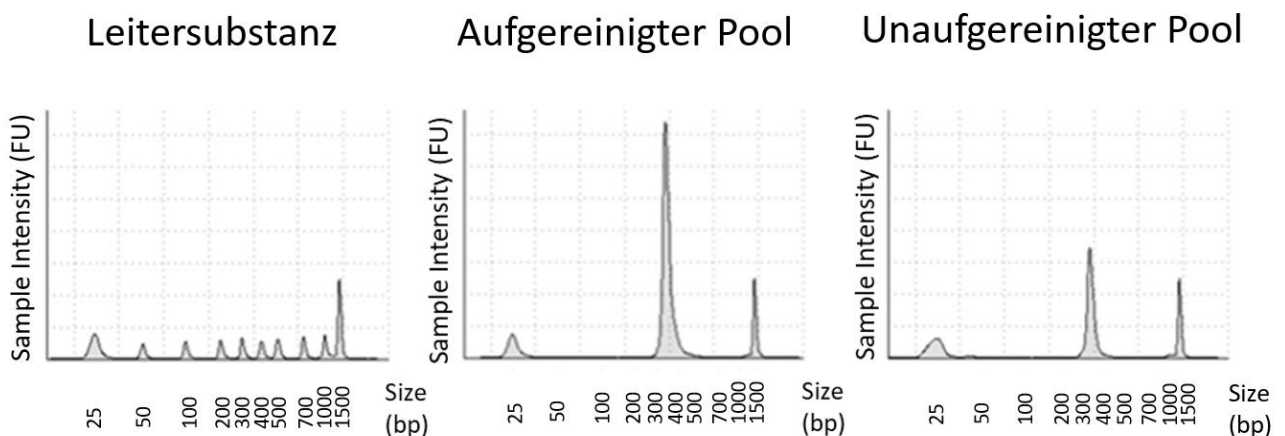


Abb. 6: Vergleich des aufgereinigten mit dem unaufgereinigten Pool.

Links: Leitersubstanz. Mitte: Aufgereinigter Pool. Rechts: Unaufgereinigter Pool.

Auf der x-Achse wird die genomische Größe der Fragmente in Basenpaaren (bp) angegeben, auf der y-Achse die Intensität des Produkts (*Sample Intensity*) in der Einheit *Normalized FU (Fluorescence Unit)*.

Der aufgereinigte Pool wurde nun durch Zugabe eines Verdünnungspuffers (*Elution Buffer*, EB, Qiagen, Hilden, Deutschland) auf eine Konzentration von 2 nmol/l verdünnt und auf einem Illumina MiSeq (San Diego, Kalifornien, Vereinigte Staaten) mittels eines "v2 Micro Kit 300 cycles" (Illumina, San Diego, Kalifornien, Vereinigte Staaten) sequenziert. Der auf 2 nmol/l verdünnte Pool sollte dazu zunächst denaturiert werden. Dafür wurden 10 µl von Natriumhydroxid (NaOH) 1 mol/l (VWR International, Radnor, Pennsylvania,

Vereinigte Staaten) mit 90 µl Wasser auf eine Konzentration von 0,1 mol/l verdünnt. Mit 10 µl dieses verdünnten NaOH wurde die DNA schließlich denaturiert und nach Zugabe von insgesamt 1530 µl HT1 *Hybridization Buffer* (Hybridisierungspuffer, Illumina, San Diego, Kalifornien, Vereinigte Staaten) wurde der Pool auf eine Konzentration von 9 pmol/l gebracht.

Forward-, *Reverse*- und *Index*-Primer wurden aufgetaut und jeweils 5 µl jedes Primers wurden mit jeweils 145 µl HT1 (Illumina, San Diego, Kalifornien, Vereinigte Staaten) in ein Eppendorf-Gefäß gemischt, gevortext und herunterzentrifugiert. Jeweils 600 µl Pool (9 pmol/l) sowie 150 µl jedes verdünnten Primers wurde an die im Illumina MiSeq vorgesehenen Positionen pipettiert und der Sequenzierlauf mit einer Leselänge von 2 x 150 bp gestartet. Um die sequenzierten *Reads* später den richtigen Proben zuordnen zu können, wurde der Software eine Probenliste bereitgestellt, die die Zuordnung zwischen dem Namen jeder Probe und dem dazugehörigen, verwendeten Barcode enthielt.

2.4 Anpassung der smMIP-Primer (*Balancing* und *Rebalancing*)

Nach der Sequenzierung wurde für jeden der 686 smMIP-Primer mittels bioinformatischer Auswertung die individuelle *Coverage*, also die Abdeckungsleistung in der Einheit *Reads* (Anzahl an Ablesungen) ermittelt. Die Ergebnisse zeigten für die einzelnen smMIPs über alle sechs Proben hinweg eine ähnliche *Coverage* (durchschnittlich ca. 1.000 *Reads*).

Je nach *Coverage* müssen smMIPs vor der Resequenzierung angepasst werden, um eine möglichst einheitliche *Coverage* aller smMIPs zu erhalten. Um das zu gewährleisten, wird bei smMIPs mit niedriger *Coverage* die Konzentration erhöht. smMIPs mit einer hohen *Coverage* werden attenuiert, indem nach erfolgter Phosphorylierung zusätzlich entsprechende unphosphorylierte smMIPs zum Gesamt-smMIP-Pool hinzugefügt werden. Da es nach Bindung der unphosphorylierten smMIPs an die DNA nicht zum Ringschluss kommt, entstehen letztendlich weniger *Reads*.

Je nach Anzahl der *Reads* wurden die smMIPs zur Anpassung in verschiedene Kategorien eingeteilt. Sieben smMIPs besaßen jeweils über 5.000 *Reads* und gehörten damit zur Kategorie "A". 69 smMIPs besaßen jeweils 11-100 *Reads* und wurden in die Kategorie "B" eingeteilt. In die Kategorie "C" gehörten 29 smMIPs mit jeweils 1-10 *Reads*, in die Kategorie "D" fielen 36 smMIPs mit jeweils 0 *Reads*. Besaß ein Gen über 20 % smMIPs

mit jeweils <100 *Reads*, wurde das ganze Gen aufgrund der schlechten Leistung ausgeschlossen. Dies traf für die vier Gene *ASB16*, *CDAN1*, *CILP2* und *HAPLN4* zu. Durch den Ausschluss der vier Gene mit unzureichender *Coverage* blieben noch 16 Gene, die durch insgesamt 579 smMIPs abgedeckt wurden, übrig.

Damit gab es noch 20 smMIPs aus der Kategorie "D", 10 aus der Kategorie "C", 37 aus der Kategorie "B" und 7 smMIPs der Kategorie "A", die nun für einen erneuten Testlauf, einen sogenannten *Balancing*-Lauf, angepasst wurden. Dafür wurde ein neuer smMIP-Pool erstellt, in dem die smMIPs der Kategorie "D" in ihrer Konzentration 21-fach erhöht wurden, die smMIPs der Kategorie "C" 11-fach und die smMIPs der Kategorie "B" 6-fach.

Die restlichen smMIPs wurden wie vorher in einfacher Konzentration dem smMIP-Pool hinzugefügt.

Dieser angepasste *Balancing*-Pool (aus 126,4 µl smMIPs bestehend) wurde gemäß Protokoll mit 5,1 µl T4 Polynukleotidkinase (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten) und 18 µl 10x T4 DNA Ligase Puffer mit 10 mmol/l ATP (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten) sowie 30,5 µl Wasser phosphoryliert und für ein praktikableres Volumen 1:300 verdünnt. Zur Attenuierung wurden unphosphorylierte smMIPs der Kategorie "A" (0,7 µl) nach der Phosphorylierung des neuen smMIP-Pools dazugegeben. Für den *Balancing*-Lauf wurden nun dieselben sechs Kontroll-Proben mit dem optimierten und phosphorylierten smMIP-Pool hybridisiert. Die restlichen Schritte (Exonuklease-Behandlung, PCR mit 24 Zyklen, Auftragen auf 2 % ethidiumbromidhaltiges Agarosegel, Pooling und Aufreinigen der PCR-Produkte, Verdünnen des aufgereinigten Pools auf 2 nmol/l) erfolgten analog zum ersten Testlauf. Nach der Sequenzierung auf dem Illumina MiSeq (San Diego, Kalifornien, Vereinigte Staaten) mit dem Sequenzierkit "v2 Micro Kit 300 cycles" (Illumina, San Diego, Kalifornien, Vereinigte Staaten) wurde die Leistung jedes einzelnen smMIP-Primers erneut evaluiert, die durchschnittliche *Coverage* der smMIPs zeigte nun 1.100 *Reads*. Etwa 90 % der 579 verbliebenen smMIP-Primer besaßen eine individuelle *Coverage* von >100 *Reads* (s. Abb. 7, S. 38). Die Attenuierung der sieben smMIPs der ursprünglichen Kategorie "A" war erfolgreich, die individuellen *Coverages* der sieben smMIPs lagen jetzt jeweils bei ca. 2.000-3.000 *Reads*. 13 smMIPs waren nun noch in den Kategorien "D" bzw. "C" mit <10 *Reads* vorhanden, die *Coverage* dieser smMIPs hatte sich im Vergleich zum ersten Testlauf trotz

Anpassung nicht gesteigert. Diese smMIPs wurden dann nicht erneut weiter angepasst, da eine weitere Erhöhung der Konzentration nicht zielführend schien. Von 22 smMIPs, die jetzt der Kategorie "B" angehörten, hatten sieben smMIPs im ersten Testlauf >100 *Reads*, diese wurden nun 6-fach konzentriert. Die 15 restlichen smMIPs der Kategorie "B" hatten sich im Vergleich zum Testlauf verbessert und wurden jetzt erneut höher konzentriert (von 6-facher zu 11-facher Konzentration und von 11-facher zu 16-facher Konzentration). smMIP-Primer aus der 21-fachen Konzentration wurden nicht mehr gesteigert. Drei smMIPs waren im Testlauf noch in Kategorie "D" oder "C" und erlangten im *Balancing* eine *Coverage* von >3.000 *Reads*. Diese smMIPs wurden eine Stufe niedriger konzentriert (von 21-facher zu 11-facher Konzentration bzw. von 11-facher zu 6-facher Konzentration). Restliche smMIPs, deren Anpassung erfolgreich war (d.h. smMIPs mit einer *Coverage* zwischen 100 und 5.000 *Reads*), wurden wieder in gleicher Weise wie für den *Balancing*-Lauf angepasst bzw. in einfacher Konzentration belassen. Diese erneute Anpassung der smMIPs war das sogenannte *Rebalancing*.

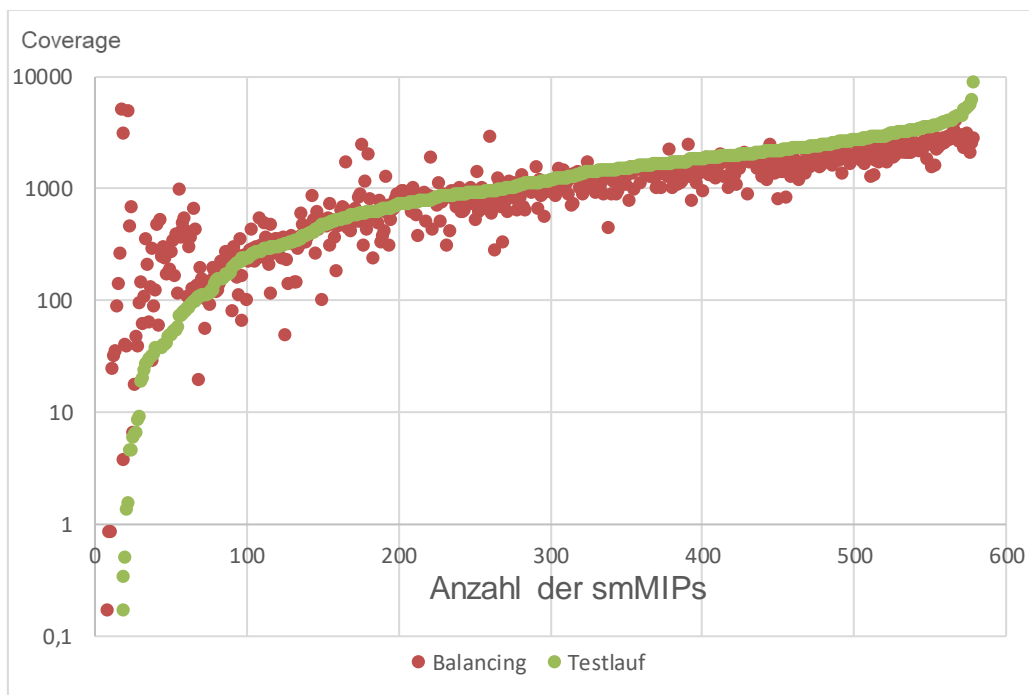


Abb. 7: Erzielte *Coverage* der 579 smMIPs im Test- und Balancinglauf.

Jeder rote Punkt zeigt die *Coverage* eines smMIPs im *Balancing*-Lauf, jeder grüne Punkt die *Coverage* eines smMIPs im Testlauf. Auf der x-Achse ist die Anzahl der smMIPs dargestellt, auf der y-Achse die *Coverage* in der Einheit *Reads* in logarithmischer Skalierung.

Die erneut angepassten smMIPs wurden gepoolt (106,4 µl), erneut mit 4,3 µl T4 Polynukleotidkinase (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten), 15 µl T4 DNA Ligase Puffer mit 10 mmol/l ATP (New England Biolabs, Ipswich, Massachusetts, Vereinigte Staaten) und 24,3 µl Wasser phosphoryliert. Anschließend wurden die sieben smMIPs der Kategorie "A" (0,7 µl) aus dem Testlauf erneut unphosphoryliert hinzugefügt, da die Attenuierung dort erfolgreich war. Dieser finale smMIP-Pool wurde 1:200 verdünnt und es wurden 30 Aliquots erstellt, um für jede zu hybridisierende 96er-Platte einen fertig phosphorylierten smMIP-Pool zu erhalten und so unnötige Frier-Tau-Zyklen des phosphorylierten Gesamtpools zu vermeiden.

Dieser final angepasste smMIP-Pool wurde nun für die Sequenzierung der 1.052 Patienten- und 1.247 Kontrollproben verwendet.

2.5 Vorbereitung der DNA-Proben

Die extrahierte DNA der Proben lag bereits in Konzentrationen von entweder 20 ng/µl oder 50 ng/µl vor. Da das Resequenzierungsprotokoll einen Einsatz von 100 ng DNA verlangte, wurde die Konzentration jeder DNA-Probe auf 8 ng/µl verdünnt und diese Konzentration mittels QuantIT PicoGreen dsDNA Assay (Thermo Fisher Scientific, Waltham, Massachusetts, Vereinigte Staaten) in Doppelbestimmung gemessen und ggf. angepasst. Alle 2.299 DNA-Proben wurden auf 24 96er-Platten aufgeteilt. Dabei wurden jeder Platte im Verhältnis von etwa 50:50 sowohl Patienten- als auch Kontrollproben zugeordnet, um Plattenartefakte zu vermeiden.

2.6 Sequenzierung

Für jede Hybridisierung wurde nun ein Aliquot des finalen phosphorylierten smMIP-Pools genutzt. Hierfür wurden jeweils 0,13 µl des verdünnten smMIP-Pools (Verhältnis DNA-Molekül/Phosphorylierte smMIPs 1/800) mit 100 ng (= 12,5 µl) DNA zu der Hybridisierungsmischung hinzugefügt. Hybridisierung, Exonukleasebehandlung und PCR wurden analog zum letzten *Balancing*-Lauf durchgeführt. Jeweils fünf Proben pro Platte wurden im Anschluss an die PCR mit 24 Zyklen auf ein 2 % ethidiumbromidhaltiges Agarosegel aufgetragen. Für alle Proben zeigte sich ein Bandenprodukt mit jeweils wenig unspezifischem Nebenprodukt. Damit war anzunehmen, dass die Hybridisierung bei nahezu allen Proben erfolgreich gewesen sein sollte (s. Abb. 8, S. 40).

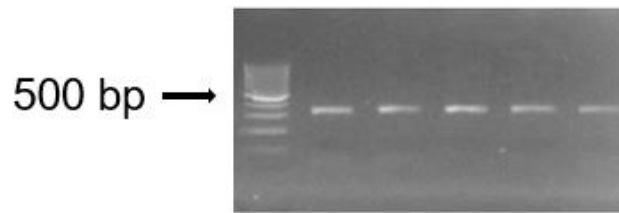


Abb. 8: Bandenprodukte von fünf zufällig ausgewählten Proben einer 96er-Platte nach 24 PCR-Zyklen auf Agarosegel.

Links: Leitersubstanz mit Pfeil auf der 500 bp-Markierung. Rechts: Fünf zufällig ausgewählte Proben. Zu erkennen ist eine deutliche Produktbande als Ausdruck des Hybridisierungserfolgs ohne unspezifisches Nebenprodukt.

Die PCR-Produkte von jeweils vier 96er-Platten ($4 \times 96 = 384$ DNA-Proben, jede mit ihrem jeweils einzigartigen PCR-Reverse-Primer-Barcode) wurden am Ende jeweils in einen Platten-Pool zusammengefügt und aufgereinigt. Die insgesamt sechs aufgereinigten Platten-Pools (4×384 PCR-Produkte, 1×382 PCR-Produkte, 1×381 PCR-Produkte = 2.299 PCR-Produkte) wurden analog zum Testlauf und Balancing mit EB auf eine Konzentration von 2 nmol/l verdünnt, um nun auf dem Illumina HiSeq2500 v4 (San Diego, Kalifornien, Vereinigte Staaten) resequenziert zu werden. Dafür war das "v4 PE 250 cycles and Cluster Kit" (Illumina, San Diego, Kalifornien, Vereinigte Staaten) notwendig. Erneut wurde zunächst die DNA denaturiert, indem 10 μ l des Pools (2 nmol/l) mit 10 μ l NaOH (0,1 mol/l) zusammenpipettiert wurden, um die DNA zu denaturieren. Nach Zugabe von 1230 μ l HT1 (Illumina, San Diego, Kalifornien, Vereinigte Staaten) wurden 750 μ l des Pools auf 15 pmol/l verdünnt. Zur Vorbereitung des *Forward*-Primers wurden pro smMIPs-Lane jeweils 1,65 μ l Primer mit 330 μ l HT1 (Illumina, San Diego, Kalifornien, Vereinigte Staaten) verdünnt, um eine Konzentration von 0,5 μ mol/l zu erreichen. 300 μ l des *Forward*-Primers wurden nach dem Vortexen und Herunterzentrifugieren anschließend in den dafür vorgesehenen Streifen pipettiert. 15 μ l des ebenfalls auf 0,5 μ mol/l verdünnten *Reverse*-Primers wurden ebenfalls in den dafür vorgesehenen Streifen gegeben. Vom *Index*-Primer wurden jeweils 12,5 μ l (0,5 μ mol/l) benötigt. Nach Eingabe der Probenliste wurde der HiSeq-Lauf gemäß der Software mit einer Leselänge von 2×125 bp gestartet und durchgeführt.

2.7 Bioinformatische Auswertung und Qualitätsanalyse

Die bioinformatische Auswertung (insbesondere die FDR (*False Discovery Rate*)-Korrektur und die genetische *Burdenanalyse*) erfolgte mit Unterstützung von Frau Anna Maaser-Hecker (naturwissenschaftliche Doktorandin am Institut für Humangenetik der Universität Bonn).

Nach dem Resequenzierungslauf auf dem HiSeq 2500 v4 (Illumina, San Diego, Kalifornien, Vereinigte Staaten) lagen die Daten zunächst in einem *Binary Base Call* (BCL)-Format vor. Zur bioinformatischen Auswertung wurden die Dateien zunächst mit der *bcl2fastq*-Software in FASTQ-Dateien umgewandelt. Durch den *Paired-End Read Merger* (PEAR) wurden die beiden *Reads* zusammengeführt (Zhang et al., 2014) und durch BWA-MEM (*Burrows-Wheeler Alignment Maximum Exact Matches*) wurde die Angleichung an das hg19 Referenzgenom durchgeführt (Li und Durbin, 2009). Mit Hilfe von verfügbaren MIPgen Skripten (Boyle et al., 2014) wurden die beiden smMIP-Primersequenzen von den *Reads* abgeschnitten und *Reads* mit demselben molekularen Marker (fünf degenerierte Basen bei der smMIPs-Option) in einem *Binary Alignment Map* (BAM)-Dokument zu einzelnen Konsensus-*Reads* vereint, da sie das Ergebnis der Amplifikation des gleichen PCR-Produkts darstellen. Dadurch wurden PCR-Artefakte reduziert und eine höhere Datenqualität gewährleistet. Mit Hilfe des GATK (*Genome Analysis ToolKit*) Unified Genotyper und Haplotype Caller (Van der Auwera et al., 2013) wurden die Varianten gecallt und es wurde anschließend ein nicht weiter annotiertes "rohes" *Variant Call Format* (VCF)-Dokument erstellt. Dieses VCF-Dokument wurde anschließend mit Hilfe des Programms ANNOVAR (*ANNOtate VARIation*, Wang et al., 2010) mit weiteren Informationen wie z.B. der funktionellen Konsequenz der Varianten und einem Phred-skalierten CADD-Wert (*Combined Annotation Dependent Depletion*, Version 1.3, Rentzsch et al., 2019) annotiert. Wurde eine Variante in >25 % der *Reads* identifiziert, wurde sie als heterozygot klassifiziert, wurde sie in >75 % der *Reads* identifiziert, als homozygot. Insgesamt 84 der 982 identifizierten Varianten wurden aufgrund der in der Bioinformatik-Pipeline vorgegebenen Qualitätskriterien (GATK-Filtergrenzen: Phred-skalierte Qualität <30.0, Intensitätsqualität <10.0) als Varianten mit geringer Qualität eingestuft und ausgeschlossen. Die weitere Datenbearbeitung wurde mit *vcftools* 0.1.15 (Danecek et al., 2011) durchgeführt. Insgesamt 17 Patientenproben und 38 Varianten mit einer durchschnittlichen *Coverage* von <30 *Reads* wurden ausgeschlossen. 99,5 % der Proben und 97,4 % der Varianten erreichten

danach eine mittlere *Coverage* von >15 *Reads*. Die mittlere *Coverage* über alle Varianten betrug 259 *Reads* für die Kontrollproben und 262 *Reads* für die Patientenproben. Alle identifizierten Varianten besaßen mindestens 10 informative *Reads* innerhalb einer Probe.

Varianten mit einer Allelfrequenz von <1 % in dem kombinierten Fall-Kontroll-Kollektiv wurden als selten deklariert, Varianten mit einem Phred-skalierten CADD-Wert von >10 wurden als potentiell funktionell relevant eingestuft. Um die statistische Signifikanz auf Einzelvariantenebene zu testen, führten wir einen automatisierten Fisher's Exact Test durch (Kim, 2017). Mittels FDR-Korrektur wurde für multiples Testen korrigiert (Aggarwal und Yadav, 2016).

2.8 Genetische *Burdenanalyse*

Um die statistische Signifikanz unserer Ergebnisse auf Genebene zu testen, führten wir eine genetische *Burdenanalyse* durch. In dieser bioinformatischen Testung wird die genetische Mutationslast eines Gens ermittelt, indem die seltenen Varianten innerhalb einer genetischen Region in ihrer Gesamtheit gegenüber dem Phänotypen assoziiert werden (Wu et al., 2011). Dieser klassische *Burdentest* ist aussagekräftig, wenn die meisten Varianten in einer Region kausal für die Erkrankungsentstehung sind und dieselbe Effektrichtung aufweisen (Lee et al., 2012b). Ist nur ein Teil der Varianten kausal für die Erkrankung oder sind die Effekte sowohl protektiv als auch schädlich, ist der *Sequence Kernel Association Test* (SKAT) aussagekräftiger (Ionita-Laza et al., 2013; Lee et al., 2012a; Lee et al., 2012b). SKAT ist ein Test auf Ebene der einzelnen Varianten, der die individuellen Varianten-Phänotyp-Assoziationen in einer Matrix vereint (Lee, et al., 2012b). Da die in dieser Studie identifizierten Varianten sowohl Risiko- als auch protektive Effekte vermitteln könnten, wurde ein optimierter SKAT (SKAT-O) benutzt, der eine Kombination aus klassischer *Burdenanalyse* und dem SKAT darstellt. SKAT-O wählt je nach Datensatz die bessere Strategie aus und maximiert damit die Testleistung und die individuelle Aussagekraft (Lee et al., 2012a; Lee et al., 2012b; Wu et al., 2011).

Zur Durchführung der SKAT-O Analyse wurden 20 multiallelische Varianten und InDels ausgeschlossen. Die nicht annotierte VCF-Datei mit den Varianten wurde für die Analyse zunächst in ein PLINK-Format überführt. Eine Auflistung der genomischen Zielregionen

(Exone der 16 Gene +/- sechs Basenpaare) wurde ebenfalls erstellt und in die Analyse eingepflegt, um später eine Statistik für die bestimmte Region zu erhalten. Ebenso benötigte die Analyse ein Informationsdokument der Proben inklusive Fall-/Kontroll-Status. Als Marker für funktionelle Relevanz benutzten wir den Phred-skalierten CADD-Wert, einen quantitativen Marker für die potentiellen funktionellen Konsequenzen einer Variante (Rentzsch et al., 2019).

Wir führten die kombinierte genetische *Burdenanalyse* zunächst für alle seltenen Varianten mit einem Phred-skalierten CADD-Wert >10 durch. In diese Analyse gingen 532 seltene Varianten aus den 16 Kandidatengenen ein. Weiterhin führten wir die Analyse für 273 seltene Varianten mit einem CADD-Wert >20 durch, um die 1 % der „schädlichsten“ Varianten in Relation zu allen möglichen Einzelbasenvarianten ebenfalls abzubilden (Rentzsch et al., 2019). Anschließend erfolgte eine Korrektur der 32 generierten P-Werte mittels FDR-Korrektur. P-Werte $<0,05$ nach Korrektur für multiples Testen wurden als statistisch signifikant angesehen.

Um die weitere klinische Signifikanz der Gene zu prüfen, bestimmten wir den RVIS-Wert (*Residual Variation Intolerance Score*) des CCDS Projekts (*Consensus Coding Sequence Project*, 9. Veröffentlichung, Petrovski et al., 2013) für die nominal signifikanten Gene bei einem CADD-Wert >10 und >20 . Je negativer der Wert, desto intoleranter ist das Gen gegenüber funktionellen Varianten. Zusätzlich zu dem linearen Wert wird ein Prozentsatz angezeigt, um die intolerantesten Gene einzustufen (Petrovski et al., 2013).

3. Ergebnisse

3.1 Identifizierte Varianten nach Qualitätsanalyse

Nach Anwendung der Qualitätsparameter wurden bei 1.035 Patienten mit Bipolarer Störung und 1.247 Kontrollprobanden insgesamt 860 Varianten in den 16 Genen identifiziert. Wir klassifizierten 790 dieser Varianten als selten (MAF <1 %, Tab. 4, S. 44).

Tab. 4: Übersicht der 860 identifizierten Varianten nach Qualitätsanalyse. 790 Varianten wurden als selten (Allelfrequenz <1 %), 543 als selten und potentiell funktionell relevant (CADD-Wert >10) eingestuft. Tabelle 4 zeigt die Aufteilung der Varianten auf die 16 untersuchten Gene.

Gen	Anzahl Varianten	Anzahl seltene Varianten	Anzahl seltene + potentiell funktionell relevante Varianten
<i>ADCY2</i>	52	46	34
<i>ADD3</i>	26	26	22
<i>ANK3</i>	165	150	102
<i>CACNA1C</i>	104	93	68
<i>FSTL5</i>	44	40	29
<i>GLYCTK</i>	25	22	14
<i>GNL3</i>	25	22	14
<i>GRIN2A</i>	50	48	22
<i>LMAN2L</i>	10	10	7
<i>NMB</i>	12	10	5
<i>PACS1</i>	31	30	22
<i>PLEKHO1</i>	26	25	14
<i>PPM1M</i>	24	24	22
<i>RIMS1</i>	75	70	51
<i>SHANK2</i>	103	98	69
<i>THSD7A</i>	88	76	48
Insgesamt	860	790	543

Für die 790 seltenen Varianten wurden statistische Analysen mittels Fisher's Exact Test durchgeführt. Insgesamt sieben Varianten in den drei Genen *ADD3*, *ANK3* und *GRIN2A* zeigten einen nominal signifikanten P-Wert ($<0,05$) (s. Tab. 5, S. 46). Nach einer FDR-Korrektur für multiples Testen waren die Assoziationen der sieben Varianten nicht mehr signifikant.

Für die weitere Priorisierung wurde der CADD-Wert der seltenen Varianten abgefragt, um deren mögliche funktionelle Relevanz zu bestimmen. Für sieben der 790 seltenen Varianten (vier InDels und drei intronische Varianten) konnte kein Phred-skaliertes CADD-Wert erhoben werden. Diese Varianten wurden jedoch jeweils nur in einer Person gefunden und waren in der statistischen Analyse (Fisher's Exact Test, s. o.) nicht nominal signifikant. In der vorliegenden Studie wurden insgesamt 543 Varianten als selten und potentiell funktionell relevant (MAF $<1\%$ und Phred-skaliertes CADD-Wert >10) eingestuft (s. Tab. 4, S. 44). Dazu zählten interessanterweise auch die sieben o. g. Varianten, die in der statistischen Einzelvariantenanalyse eine nominale Assoziation zeigten (Tab. 5, S. 46). Im Folgenden werden die seltenen, potentiell funktionellen Varianten, die in den Genen *ADD3*, *ANK3* und *GRIN2A* identifiziert wurden, detaillierter dargestellt, da diese drei Gene durch die sieben o. g., nominal signifikanten Varianten impliziert wurden.

Tab. 5: Nominal signifikante seltene Varianten gemäß Fisher's Exact Test.

SNV= *Single Nucleotide Variant* mit Basenaustausch. Aminosäureaustausch= Aminosäureaustausch mit Position, Transkript bei *GRIN2A* GRIN2A:NM_001134407, Transkript bei *ANK3* ANK3:NM_020987 bzw. ANK3:NM_001204404 (mit + markiert), Transkript bei *ADD3* ADD3:NM_001121 bzw. ADD3:NM_001320592 (mit + markiert). Funktion= Funktionelle Konsequenz. CADD-Wert= *Combined Annotation Dependent Depletion Score*. Het. Pat./Het. Kontr.= Heterozygote Patienten/Heterozygote Kontrollen= Anzahl an Patienten bzw. Kontrollpersonen, in der die Variante heterozygot identifiziert wurde. P_{nom} = nominal signifikanter P-Wert. P_{kor} = P-Wert nach FDR-Korrektur für multiples Testen. Die Varianten sind aufsteigend nach nominal signifikantem P-Wert sortiert.

Gen	SNV	Aminosäureaustausch	Funktion	CADD-Wert	Het. Pat./ Het. Kontr.	P_{nom}	P_{kor}
GRIN2A	rs61758995 G/T	N1076K	Nicht-synonym	23,0	8/31	0,002	>0,999
ANK3	rs41274676 G/A	P1489S	Nicht-synonym	17,7	0/8	0,010	>0,999
ADD3	rs140299264 A/G	V163V	Synonym	14,2	18/9	0,032	>0,999
ANK3	rs139092048 C/A	D1661Y+	Nicht-synonym	14,5	14/6	0,040	>0,999
ANK3	rs117138204 G/A	C349C	Synonym	13,2	16/8	0,040	>0,999
ADD3	rs149714005 C/A	Q602K+	Nicht-synonym	16,6	4/0	0,042	>0,999
ANK3	rs141939315 C/T	R4369Q	Nicht-synonym	35,0	13/31	0,047	>0,999

3.1.1 *ADD3*

Tab. 9 (s. Anhang, S. 59-60) zeigt eine Übersicht über die 22 identifizierten seltenen, potentiell funktionellen Varianten in *ADD3*. Vier synonyme Varianten zeigen einen vergleichsweise niedrigen CADD-Wert (von 10,3 bis 14,2), die 18 nichtsynonymen Varianten einen eher hohen Wert (von 13,6 bis 47,0). Die einzige identifizierte Stoppmutation bildet mit 47,0 den höchsten CADD-Wert ab, diese ließ sich in der vorliegenden Studie aber lediglich in einer gesunden Kontrollperson heterozygot identifizieren und ist dementsprechend nicht statistisch signifikant ($P_{\text{nom}} > 0,999$, $P_{\text{kor}} > 0,999$).

Im *ADD3*-Gen konnten zwei nominal signifikante Varianten identifiziert werden. Beide Varianten sind bei Patienten mit Bipolarer Störung gegenüber gesunden Kontrollpersonen überrepräsentiert. Die Variante rs140299264 ($P_{\text{nom}} = 0,032$, $P_{\text{kor}} > 0,999$) stellt eine synonyme Variante mit einem Phred-skalierten CADD-Wert von 14,2 dar, rs149714005 ($P_{\text{nom}} = 0,042$, $P_{\text{kor}} > 0,999$) eine nichtsynonyme Variante, die zu einem Aminosäureaustausch von Glutamin (Q) zu Lysin (K) an Position 602 führt und einen CADD-Wert von 16,6 aufweist.

3.1.2 *ANK3*

Insgesamt fanden sich in *ANK3* 102 seltene und potentiell funktionelle Varianten (s. Tab. 10 im Anhang, S. 61-67). Den Großteil (82 Varianten) bildeten die nichtsynonymen Varianten ab. Die Spannweite des CADD-Wertes lag hier bei 10,1 bis 35,0.

In *ANK3* konnten in dieser Studie die meisten nominal signifikanten Varianten ($n = 4$) identifiziert werden. Diese zeigten unterschiedliche Effektrichtungen: Zwei der Varianten zeigten eine Überrepräsentation bei Patienten mit Bipolarer Störung gegenüber gesunden Kontrollprobanden, darunter die nichtsynonyme Variante rs139092048 ($P_{\text{nom}} = 0,040$, $P_{\text{kor}} > 0,999$) mit einem CADD-Wert von 14,5 und einem Aminosäureaustausch von Asparaginsäure (D) zu Tyrosin (Y) an Position 1661 und rs117138204 ($P_{\text{nom}} = 0,040$, $P_{\text{kor}} > 0,999$) als synonyme Variante mit einem CADD-Wert von 13,2. Die anderen beiden, nichtsynonymen Varianten traten in gesunden Kontrollpersonen häufiger als bei Patienten mit Bipolarer Störung auf. rs41274676 ($P_{\text{nom}} = 0,010$, $P_{\text{kor}} > 0,999$) besaß einen CADD-Wert von 17,7 und führt zu einem Aminosäureaustausch von Prolin (P) zu Serin (S) an

Position 1489, rs141939315 ($P_{\text{nom}}= 0,047$, $P_{\text{kor}}>0,999$) besaß mit 35,0 den höchsten CADD-Wert (Aminosäureaustausch Arginin (R) zu Glutamin (Q) an Position 4369).

3.1.3 *GRIN2A*

Im *GRIN2A* Gen wurden insgesamt 22 seltene und potentiell funktionelle Varianten identifiziert (s. Tab. 11 im Anhang, S. 68-69). Sechs synonyme Varianten zeigten einen tendenziell niedrigen CADD-Wert (Spannweite 10,4 bis 15,2), wohingegen die restlichen 16 nichtsynonymen Varianten CADD-Werte von 11,8 bis 34,0 aufwiesen.

Die in *GRIN2A* nachgewiesene, nominal signifikante seltene Variante rs61758995 ($P_{\text{nom}}= 0,002$, $P_{\text{kor}}>0,999$) ist eine nichtsynonyme Variante, die an der Aminosäureposition 1076 zu einem Austausch von Asparagin (N) zu Lysin (K) führt. Die Variante besaß einen CADD-Wert von 23,0 und war in der vorliegenden Studie bei Kontrollpersonen nominal signifikant überrepräsentiert.

3.2 Genetische *Burdenanalyse*

Die genetische *Burdenanalyse* wurde für die 16 untersuchten Gene mittels SKAT-O (Lee et al., 2012b) für seltene (MAF <1 %), möglicherweise funktionell relevante Varianten mit den CADD-Werten >10 (insgesamt 543 Varianten abzüglich 11 multiallelischer Varianten und InDels= 532 Varianten) und >20 (278 Varianten abzüglich 5 multiallelischer Varianten und InDels= 273 Varianten) durchgeführt. Im Anschluss erfolgte eine FDR-Korrektur der 32 generierten P-Werte (s. Tab. 6, S. 49). In der *Burdenanalyse* zeigten sich für seltene Varianten mit einem CADD-Wert von >10 nominal signifikante Assoziationen für die Gene *ADD3*, *ANK3* und *GRIN2A*, die nach FDR-Korrektur nicht korrekturstabil waren. In der Analyse von Varianten mit einem CADD-Wert von >20 konnte eine nominal signifikante Assoziation für vier Gene (*ADD3*, *ANK3*, *GRIN2A* und *LMAN2L*) nachgewiesen werden. Nach FDR-Korrektur für multiples Testen zeigte sich jedoch für keines dieser Gene ein korrekturstabil signifikantes Ergebnis (s. Tab. 6, S. 49).

Tab. 6: Genbasierte P-Werte der *Burdenanalyse*.

$P_{CADD>10}$ = P-Wert aller seltenen Varianten mit einem CADD-Wert >10. $P_{korrr10}$ = FDR-korrigierte P-Werte (für 32 Gene/Tests) aller seltenen Varianten mit einem CADD-Wert >10. $P_{CADD>20}$ = P-Wert aller seltenen Varianten mit einem CADD-Wert >20. $P_{korrr20}$ = FDR-korrigierte P-Werte (für 32 Gene/Tests) aller seltenen Varianten mit einem CADD-Wert >20. Nominal signifikante P-Werte sind fett hervorgehoben.

Gen	$P_{CADD>10}$	$P_{korrr10}$	$P_{CADD>20}$	$P_{korrr20}$
<i>ADCY2</i>	0,872	0,946	0,286	0,703
<i>ADD3</i>	0,006	0,059	0,044	0,200
<i>ANK3</i>	0,009	0,070	0,042	0,200
<i>CACNA1C</i>	0,420	0,840	0,756	0,946
<i>FSTL5</i>	0,479	0,901	0,811	0,946
<i>GLYCTK</i>	0,328	0,750	0,717	0,946
<i>GNL3</i>	0,519	0,923	0,172	0,689
<i>GRIN2A</i>	0,003	0,054	0,002	0,051
<i>LMAN2L</i>	0,214	0,695	0,043	0,200
<i>NMB</i>	0,852	0,946	0,797	0,946
<i>PACS1</i>	0,726	0,946	0,239	0,695
<i>PLEKHO1</i>	0,395	0,840	0,887	0,946
<i>PPM1M</i>	0,230	0,695	0,789	0,946
<i>RIMS1</i>	>0,999	>0,999	0,283	0,703
<i>SHANK2</i>	0,630	0,946	0,567	0,946
<i>THSD7A</i>	0,798	0,946	>0,999	>0,999

3.3 Residual Variation Intolerance Score (RVIS)

In der *Burdenanalyse* von seltenen Varianten mit einem CADD-Wert >10 und >20 zeigten insgesamt drei Gene eine nominal signifikante Assoziation in beiden Untersuchungen: *ADD3*, *ANK3* und *GRIN2A*. Von diesen besitzen *GRIN2A* und *ANK3* einen negativen RVIS-Wert und gehören demzufolge zu den 3,89 % bzw. 0,34 % am meisten gegenüber genetischer Variation intoleranten Genen. *ADD3* weist einen positiven Wert und einen Prozentrang von 64,82 % auf, dieses Gen toleriert also eher genetische Variation (s. Tab. 7, S. 50).

Tab. 7: RVIS Wert der Gene *ADD3*, *ANK3* und *GRIN2A*.
Pro Gen ist jeweils ein nominaler Wert und ein Prozentrang angegeben.

Gen	RVIS Wert	Prozentrang
<i>ADD3</i>	0,16	64,82 %
<i>ANK3</i>	-3,50	0,34 %
<i>GRIN2A</i>	-1,46	3,89 %

Quelle: <https://genic-intolerance.org/>, Zugriffsdatum: 14.08.2023

4. Diskussion

4.1 Vielversprechende Kandidatengene und -varianten

In der vorliegenden Studie wurden proteinkodierende Gene, die an den 30 genomweit signifikanten Loci für Bipolare Störung aus der GWAS von Stahl et al. (2019) lokalisiert waren, mittels Resequenzierung auf das Vorliegen von seltenen, potentiell funktionell relevanten Varianten untersucht. Nach Beginn der vorliegenden Studie wurde eine neuere GWAS-Metaanalyse veröffentlicht, die die Proben der Stahl et al. (2019) GWAS weitgehend beinhaltete und insgesamt 41.917 Patienten mit Bipolarer Störung sowie 371.549 Kontrollen umfasste (Mullins et al., 2021). In dieser neueren GWAS-Metaanalyse wurden 64 genomweit signifikante Genloci für die Bipolare Störung identifiziert (Mullins et al., 2021). Die 64 genomweit signifikanten Loci haben die chromosomalen Regionen um *ADD3*, *ANK3* und *GRIN2A* erneut eingeschlossen, insgesamt konnten in der Mullins et al. (2021) GWAS 28 der 30 Loci aus Stahl et al. (2019) bestätigt werden. Trotz der Vergrößerung der GWAS-Datensätze und der Anwendung bioinformatischer Analysen ist jedoch weiterhin weitestgehend unklar, welches Gen bzw. welche Gene an den assoziierten Loci zur Entstehung der Bipolaren Störung beitragen. Daher ist die Durchführung weiterführender Analysen, wie die in der vorliegenden Studie erfolgte Resequenzierung, sinnvoll, um die Bedeutung der implizierten Gene weiter abzuklären.

In der vorliegenden Studie konnten wir nach der Resequenzierung von 16 Kandidatengen in über 1.000 Patienten mit Bipolarer Störung und über 1.000 gesunden Kontrollprobanden sieben seltene Varianten in den drei Genen *ADD3*, *ANK3* und *GRIN2A* identifizieren, die nominal mit der Bipolaren Störung assoziiert waren. Diese Varianten zeigten sowohl risikobezogene als auch protektive Effekte, die nachgewiesenen Assoziationen hielten jedoch einer Korrektur für multiples Testen nicht stand. Von den sieben nominal signifikanten Varianten waren die vier Varianten rs149714005 (nichtsynonyme Variante in *ADD3*), rs140299264 (synonyme Variante in *ADD3*), rs117138204 (synonyme Variante in *ANK3*) und rs139092048 (nichtsynonyme Variante in *ANK3*) in Patienten mit Bipolarer Störung im Vergleich zu gesunden Kontrollprobanden überrepräsentiert. Dem gegenüber stehen die seltenen nichtsynonymen Varianten rs141939315 (in *ANK3*), rs41274676 (in *ANK3*) und rs61758995 (in *GRIN2A*), die in der vorliegenden Studie in Kontrollprobanden

überrepräsentiert waren. Die Variante mit dem niedrigsten, nominalen P-Wert in dieser Studie war rs61758995 im *GRIN2A*-Gen. Sie stellt einen Einzelbasenaustausch von Guanin zu Thymin dar, der auf Proteinebene zum Austausch von Asparagin zu Lysin an Position 1076 führt (p.N1076K), und konnte in heterozygoter Form in acht Patienten und 31 Kontrollprobanden identifiziert werden ($P_{\text{nom}} = 0,002$, $P_{\text{kor}} > 0,999$).

Die Ergebnisse der o. g. sieben Varianten wurden mit den Resultaten des *Bipolar-Exome*-Sequenzierprojekts (BipEx, Palmer et al., 2022) verglichen (13.933 Patienten mit Bipolarer Störung und 14.422 Kontrollen). Im BipEx-Projekt konnten Ergebnisse für sechs unserer sieben nominal signifikanten, seltenen Varianten erhoben werden (s. Tab. 8, S. 53). Für alle sechs Varianten bestätigten die Allelfrequenzdaten der BipEx-Studie, dass es sich um seltene Varianten handelt ($\text{MAF} < 0,01$). Fünf der Varianten zeigten in der BipEx-Studie eine nahezu vergleichbare Frequenz zwischen Patienten mit Bipolarer Störung und Kontrollpersonen ($P \geq 0,194$; Palmer et al., 2022). Als einzige der sieben Varianten zeigte rs117138204 in *ANK3* in der BipEx-Studie eine nominal signifikante Assoziation mit Bipolarer Störung ($P = 0,012$; Palmer et al., 2022), interessanterweise war diese Variante wie in unserer Studie in Patienten mit Bipolarer Störung im Vergleich zu Kontrollpersonen leicht überrepräsentiert. Diese gleiche Effektrichtung könnte darauf hinweisen, dass die seltene Variante rs117138204 in *ANK3* zur Entstehung der Bipolaren Störung beiträgt. Da die nachgewiesenen Assoziationen jedoch nur nominal signifikant waren, sollten die in dieser Studie identifizierten Einzelvarianten in zukünftigen Studien mit noch größeren Fall- und Kontrollzahlen untersucht werden, um ihre Relevanz für die Ätiologie der Bipolaren Störung abschließend zu klären.

Tab. 8: Vergleich der Ergebnisse der BipEx-Daten mit unserer Studie für die sieben seltenen Varianten mit nominal signifikanter Assoziation.

Die Ergebnisse der BipEx-Studie (einschließlich der P-Werte) wurden dem BipEx: Bipolar Exomes Browser (<https://bipex.broadinstitute.org/>) entnommen. Variante= Gen, SNP ID und Basenaustausch der zugrundeliegenden Variante. AF Pat.= Allelfrequenz in Patienten. AF Kontr.= Allelfrequenz in Kontrollen. P-Wert= nominal signifikanter P-Wert. k.A.= keine Angabe. Nominal signifikante Ergebnisse sind fett markiert.

Variante	Unsere Studie			BipEx		
	AF Pat.	AF Kontr.	P-Wert	AF Pat.	AF Kontr.	P-Wert
GRIN2A rs61758995 G/T	0,004	0,012	0,002	0,009	0,009	0,853
ANK3 rs41274676 G/A	0,000	0,003	0,010	0,003	0,003	0,418
ADD3 rs140299264 A/G	0,009	0,004	0,032	0,008	0,009	0,194
ANK3 rs139092048 C/A	0,007	0,002	0,040	k.A.	k.A.	k.A.
ANK3 rs117138204 G/A	0,008	0,003	0,040	0,010	0,008	0,012
ADD3 rs149714005 C/A	0,002	0,000	0,042	0,000	0,001	0,465
ANK3 rs141939315 C/T	0,006	0,012	0,047	0,006	0,006	0,583

Um die Signifikanz der Varianten auf Genebene zu testen, führten wir eine kombinierte genbasierte *Burdenanalyse* durch und benutzen den optimierten SKAT-O, der effizienter ist, wenn die Varianten verschiedene Effektrichtungen aufweisen (Ionita-Laza et al., 2013; Lee et al., 2012a; Lee et al., 2012b). In dieser Untersuchung (s. Tab. 6, S. 49) zeigte sich sowohl bei der Analyse seltener Varianten mit CADD >10 als auch bei der Untersuchung von seltenen Varianten mit CADD >20 eine nominal signifikante Assoziation für die drei Gene *ADD3*, *ANK3* und *GRIN2A*. Ein korrekturstabiler Befund nach FDR-Korrektur konnte jedoch für keines dieser Gene erhoben werden. Da die drei Gene *ADD3*, *ANK3* und *GRIN2A* sowohl in der Einzelvariantenanalyse als auch in der genbasierten *Burdenanalyse* mit nominaler Signifikanz impliziert wurden, werden die Funktion der Gene sowie ihre mögliche Beteiligung an der Bipolaren Störung und anderen psychischen Erkrankungen im Folgenden diskutiert.

4.1.1 *GRIN2A*

Das *GRIN2A*-Gen kodiert für das Protein GluN2A, das eine Untereinheit des ionotropen Glutamaterezeptors vom NMDA-Typ darstellt (Strehlow et al., 2019). NMDA-Rezeptoren sind an der Vermittlung der exzitatorischen Neurotransmission beteiligt und besitzen dementsprechend eine wichtige Rolle bei Lern- und Gedächtnisprozessen (Strehlow et al., 2019). Interessanterweise ist *GRIN2A* nicht nur ein für die Entstehung verschiedener entwicklungsneurologischer Erkrankungen bekanntes Gen (Strehlow et al., 2019), sondern zeigt auch Assoziationen zur Schizophrenie (Poltavskaya et al., 2021): *GRIN2A* liegt an einem genomweit signifikanten Locus für die Schizophrenie (Trubetsky et al., 2022), außerdem sind sehr seltene, funktionell relevante Varianten in *GRIN2A* signifikant mit einem erhöhten Erkrankungsrisiko für die Schizophrenie assoziiert (Singh et al., 2022). In der vorliegenden Studie konnten wir für die Variante rs61758995 eine nominal signifikante Überrepräsentation in Kontrollprobanden im Vergleich zu Patienten mit Bipolarer Störung nachweisen. In den genbasierten Analysen zeigte *GRIN2A* ebenfalls eine nominal signifikante Assoziation. Da diese Ergebnisse jedoch nicht korrekturstabil waren und die Variante rs61758995 in der größeren BipEx-Studie eine nahezu vergleichbare Frequenz zwischen Patienten mit Bipolarer Störung und Kontrollpersonen zeigte ($P=0,853$; Palmer et al., 2022), liefert unsere Studie insgesamt keine starken Hinweise darauf, dass seltene Sequenzvarianten in *GRIN2A* zur Entstehung der Bipolaren Störung beitragen. Unsere

Ergebnisse schließen eine entsprechende Beteiligung jedoch auch nicht gänzlich aus, sodass zukünftige Studien in größeren Fall-Kontroll-Stichproben benötigt werden, um den Beitrag seltener, funktionell relevanter Varianten im *GRIN2A*-Gen zur Entwicklung der Bipolaren Störung zu klären.

4.1.2 *ANK3*

Ankyrin G, das Genprodukt von *ANK3*, ist für die Integration spannungsabhängiger Natriumkanäle an den Ranvier'schen Schnürringen zuständig und spielt infolgedessen eine Rolle bei der synaptischen Übertragung (Iqbal et al., 2013). *ANK3* wurde bereits in mehreren vorangegangenen Studien als Kandidatengen für die Bipolare Störung impliziert (Fiorentino et al., 2014; Georgi et al., 2014; Mühleisen et al., 2014; Stahl et al., 2019), auch sind Mutationen in *ANK3* für Autismus-Spektrum-Störungen und Entwicklungsverzögerungen beschrieben (Kloth et al., 2017). In einer kombinierten WES- und Kopplungsanalyse von 15 australischen, mehrfach betroffenen Familien mit Bipolarer Störung konnten Toma und Kollegen eine signifikante Segregation von seltenen Varianten in *ANK3* nachweisen (Toma et al., 2021). In einer anderen Studie von Hughes et al. (2018) konnte bestätigt werden, dass eine seltene Spleißvariante (rs41283526), die eine Isoform von *ANK3* betrifft, protektive Effekte gegenüber Bipolarer Störung und Schizophrenie aufweist (Hughes et al., 2018). Die Variante rs41283526 wurde auch in der vorliegenden Studie etwas häufiger bei Kontrollpersonen im Vergleich zu Patienten mit Bipolarer Störung nachgewiesen ($P_{\text{nom}} = 0,084$; s. Tab. 10 im Anhang, S. 61-67). Bei der Interpretation dieses Befundes ist jedoch zu beachten, dass zwischen der Hughes et al. (2018) und der vorliegenden Studie eine Überlappung bei den untersuchten Patienten und Kontrollen besteht.

In der vorliegenden Studie wurden seltene, nominal signifikante Varianten in *ANK3* nachgewiesen, die unterschiedliche Effektrichtungen aufwiesen. In Zusammenschau mit den o. g. Befunden könnten diese Ergebnisse darauf hinweisen, dass seltene Varianten im *ANK3*-Gen sowohl risikoe erhöhende, als auch protektive Effekte im Hinblick auf die Entstehung einer Bipolaren Störung vermitteln könnten, was in zukünftigen Studien weiter abgeklärt werden sollte.

4.1.3 *ADD3*

ADD3 kodiert für Adducin 3, das zu einer Familie von zytoskelettalen Proteinen gehört, die am Aufbau des Spektrin-Aktin-Netzwerks beteiligt sind, indem sie für die Kappung des wachsenden Endes der Aktinfilamente verantwortlich sind und die Bindung von Spektrin an Aktin fördern (Charney et al., 2017). Homozygote Varianten im *ADD3*-Gen wurden als ursächlich für die Spastische Zerebralparese beschrieben (Kruer et al., 2013; Sanchez Marco et al., 2022). Allerdings fehlt es bisher an Evidenz, dass seltene Varianten in diesem Gen an der Entstehung psychischer Erkrankungen wie der Bipolaren Störung oder der Schizophrenie beteiligt sind (Palmer et al., 2022; Singh et al., 2022). Auch in der vorliegenden Studie konnten keine seltenen, funktionell relevanten Varianten in *ADD3* identifiziert werden, die korrekturstabil mit der Entstehung der Bipolaren Störung assoziiert waren.

4.2 Limitationen

Die durchgeführte Studie hatte zahlreiche Limitationen. Bei einem Teil der untersuchten Kontrollpersonen wurde das Vorliegen einer psychischen Störung nicht ausgeschlossen, sodass wir nicht sicher sagen können, ob einige dieser Kontrollpersonen von einer Bipolaren Störung betroffen waren. Aufgrund der insgesamt eher geringen Lebenszeitprävalenz der Bipolaren Störung (ca. 1-2 % laut Merikangas et al., 2011) ist aber nicht davon auszugehen, dass die statistische Power der vorliegenden Studie durch die Verwendung von nicht gescreenten Kontrollen wesentlich verringert wurde (Moskvina et al., 2005). Die Ergebnisse unserer smMIP-Sequenzierung wurden nicht mittels einer anderen Methode (bspw. Sanger-Sequenzierung, s. Abb. 2, S. 14) überprüft. Da die bei unserer Resequenzierung identifizierten Varianten nach Qualitätsanalyse aber mindestens 10 informative *Reads* innerhalb einer Probe hatten und insgesamt eine hohe mittlere *Coverage* der smMIP-Primer vorlag, wurde in der vorliegenden Studie auf eine Bestätigung der identifizierten Varianten mittels einer anderen, unabhängigen Methode verzichtet. Da wir unsere Kandidatengene teilweise basierend auf der physischen Nähe zu den GWAS Loci gewählt haben, können wir nicht mit Sicherheit sagen, dass diese die kausalen Gene für das Erkrankungsrisiko darstellen (Westra und Franke, 2014). Zukünftige Studien sollten verschiedene bioinformatische Ansätze kombinieren, um die an den Loci gelegenen, relevanten Gene zu identifizieren (z. B. Mullins et al., 2021). Bezüglich der CADD-Werte sollte

die biologische Relevanz der identifizierten Varianten in funktionellen Analysen weiter untersucht und repliziert werden.

Obwohl in der vorliegenden Studie mehr als 1.000 Patienten mit Bipolarer Störung und 1.000 Kontrollen untersucht wurden, war die statistische Power zur Detektion von seltenen, funktionell relevanten Varianten limitiert (Bansal et al., 2010). Die Kandidatengene der vorliegenden Studie sollten daher in zukünftigen Forschungsvorhaben mit größeren Fall- und Kontroll-Stichproben weiter untersucht werden.

4.3 Ausblick

Zukünftige Sequenzierungsstudien in größeren Datensätzen werden den Beitrag seltener Varianten in den an den GWAS Loci für die Bipolare Störung gelegenen Genen weiter klären. Diese Studien sollten insbesondere Sequenziermethoden anwenden, die das ganze Genom untersuchen, um die Rolle aller (proteinkodierenden und nicht-kodierenden) Gene an den GWAS Loci systematisch zu untersuchen. Durch diese Analysen wird es auch möglich sein, die Relevanz der in der vorliegenden Studie identifizierten Sequenzvarianten hinsichtlich der Entstehung der Bipolaren Störung weiter zu charakterisieren. Langfristig könnte die Identifikation der zugrundeliegenden Gene zur Entwicklung neuer diagnostischer und therapeutischer Ansätze für die Bipolare Störung führen (Harrison et al., 2016).

5. Zusammenfassung

Die Bipolare Störung ist eine häufige psychiatrische Erkrankung mit einer geschätzten Heritabilität von bis zu 85 %. Da die in genomweiten Assoziationsanalysen untersuchten häufigen Varianten nur etwa 20-30 % der genetischen Varianz der Bipolaren Störung erklären, könnten seltene Varianten mit höherer Penetranz ebenfalls die Suszeptibilität der Bipolaren Störung beeinflussen. In dieser Arbeit wurde der Beitrag von seltenen Varianten zur Entstehung der Bipolaren Störung durch Resequenzierung von proteinkodierenden Genen an den 30 genomweit signifikanten Genloci der GWAS von Stahl et al. (2019) untersucht. Die Resequenzierung von 20 priorisierten Kandidatengenen in jeweils über 1.000 Patienten und Kontrollprobanden deutscher Herkunft wurde mittels innovativer Next-Generation-Sequencing-Methoden durchgeführt. Es konnten 860 Varianten in 16 Genen einschließlich 543 seltener (Allelfrequenz <1 %) und potentiell funktionell relevanter (CADD-Wert >10) Varianten identifiziert werden. Sieben dieser Varianten in drei Genen zeigten eine nominal signifikante Assoziation für die Bipolare Störung, nach stringenter FDR-Korrektur für multiples Testen blieb die Assoziation jedoch für keine der identifizierten Varianten signifikant. In einer genbasierten *Burdenanalyse* von seltenen Varianten mit einem CADD-Wert >10 und >20 zeigten insgesamt drei Gene (*ADD3*, *ANK3* und *GRIN2A*) eine nominal signifikante Assoziation, die einer FDR-Korrektur nicht Stand hielt. Interessanterweise kodieren die drei implizierten Gene *ADD3*, *ANK3* und *GRIN2A* für Proteine, die durch ihre Funktion eine mögliche Beteiligung an der Entstehung der Bipolaren Störung plausibel erscheinen lassen. Zusammenfassend liefern die in dieser Studie erbrachten Ergebnisse keine starken Hinweise auf einen Beitrag von seltenen Varianten in den an den GWAS Loci gelegenen Genen zur Entstehung der Bipolaren Störung. Weitere Resequenzierungsstudien in größeren Kollektiven sowie funktionelle Analysen sind notwendig, um die Relevanz der implizierten Varianten und Gene für die Pathogenese der Bipolaren Störung weiter einzuschätzen.

6. Anhang

Tab. 9: Übersicht über die 22 identifizierten seltenen (MAF <1 %) und potentiell funktionellen (CADD >10) Varianten in *ADD3*.

Position= Chromosomale Position (hg19). SNP ID= *Single Nucleotide Polymorphism Identification*. Basenaustausch= Referenz- und Alternatives Allel. Aminosäureaustausch= Aminosäureaustausch mit Position, Transkript *ADD3*:NM_001121 bzw. *ADD3*:NM_001320592 (mit + markiert). Funktion= Funktionelle Konsequenz. CADD-Wert= *Combined Annotation Dependent Depletion Score*. Het. Pat./Het. Kontr.= Heterozygote Patienten/Heterozygote Kontrollen= Anzahl an Patienten bzw. Kontrollpersonen, in der die Variante heterozygot identifiziert wurde. k.A.= keine Angabe. Die Varianten sind aufsteigend nach chromosomaler Position sortiert. Zwei nominal signifikante Varianten sind fett markiert.

Position	SNP ID	Basenaustausch	Aminosäureaustausch	Funktion	CADD-Wert	Het. Pat./ Het. Kontr.
chr10: 111860435	rs367830044	C/T	G8G	Synonym	12,5	0/1
chr10: 111860531	k.A.	C/T	N40N	Synonym	10,3	0/1
chr10: 111872595	k.A.	A/G	T86A	Nicht-synonym	15,5	1/0
chr10: 111872650	rs139343347	C/T	S104L	Nicht-synonym	23,7	2/1
chr10: 111876041	k.A.	A/G	N120S	Nicht-synonym	20,8	1/0
chr10: 111877102	rs140299264	A/G	V163V	Synonym	14,2	18/9
chr10: 111878412	k.A.	G/T	G212V	Nicht-synonym	27,6	1/0
chr10: 111879027	rs55763936	A/G	D259G	Nicht-synonym	27,2	6/2
chr10: 111881902	rs41291894	A/G	N332S	Nicht-synonym	24,6	0/1
chr10: 111881933	k.A.	A/C	K342N	Nicht-synonym	25,1	1/0

Position	SNP ID	Basen- aus- tausch	Amino- säure- austausch	Funktion	CADD- Wert	Het. Pat./ Het. Kontr.
chr10: 111883815	k.A.	G/A	R395Q	Nicht- synonym	26,8	0/1
chr10: 111884003	k.A.	G/A	G458R	Nicht- synonym	32,0	1/0
chr10: 111886249	rs143966199	T/A	D532E	Nicht- synonym	15,5	2/1
chr10: 111890122	k.A.	C/T	T537I	Nicht- synonym	17,3	1/0
chr10: 111890151	rs148633520	C/T	P547S	Nicht- synonym	21,9	3/2
chr10: 111890229	k.A.	C/T	Q573X	Stopp- mutation	47,0	0/1
chr10: 111892134	rs149714005	C/A	Q602K+	Nicht- synonym	16,6	4/0
chr10: 111893113	k.A.	A/T	I588F	Nicht- synonym	15,9	0/1
chr10: 111893133	rs145028289	C/T	V594V	Synonym	12,1	0/5
chr10: 111893160	rs73352950	G/C	M603I	Nicht- synonym	17,4	1/0
chr10: 111893161	k.A.	C/T	H604Y	Nicht- synonym	24,3	2/0
chr10: 111893198	k.A.	G/A	R616K	Nicht- synonym	13,6	0/1

Tab. 10: Übersicht über die 102 identifizierten seltenen (MAF <1 %) und potentiell funktionellen (CADD >10) Varianten in *ANK3*.

Position= Chromosomale Position (hg19). SNP ID= *Single Nucleotide Polymorphism Identification*. Basenaustausch= Referenz- und Alternatives Allel. Aminosäureaustausch= Aminosäureaustausch mit Position, Transkript *ANK3:NM_020987* bzw. *ANK3:NM_001204404* (mit + markiert) bzw. *ANK3:NM_001204403* (mit ++ markiert). Funktion= Funktionelle Konsequenz. CADD-Wert= *Combined Annotation Dependent Depletion Score*. Het. Pat./Het. Kontr.= Heterozygote Patienten/Heterozygote Kontrollen= Anzahl an Patienten bzw. Kontrollpersonen, in der die Variante heterozygot identifiziert wurde. k.A.= keine Angabe. Del= Deletion. Die Varianten sind aufsteigend nach chromosomaler Position sortiert. Vier nominal signifikante Varianten sind fett markiert.

Position	SNP ID	Basenaustausch	Aminosäureaustausch	Funktion	CADD-Wert	Het. Pat./ Het. Kontr.
chr10: 61802477	rs141939315	C/T	R4369Q	Nicht-synonym	35,0	13/31
chr10: 61802492	k.A.	G/A	T4364M	Nichtsynonym	15,9	1/0
chr10: 61815451	k.A.	G/A	H4344Y	Nichtsynonym	23,1	1/0
chr10: 61815643	k.A.	T/C	K4280E	Nichtsynonym	23,2	1/0
chr10: 61815712	rs12261793	T/C	I4257V	Nichtsynonym	11,6	1/1
chr10: 61819543	rs139092048	C/A	D1661Y+	Nicht-synonym	14,5	14/6
chr10: 61819571	k.A.	G/A	V1651V+	Synonym	12,9	1/0
chr10: 61819599	rs201194312	A/G	I1642T+	Nichtsynonym	11,2	1/0
chr10: 61822964	k.A.	A/G	V4167A	Nichtsynonym	27,5	0/1
chr10: 61827772	rs200190212	G/A	k.A.	k.A.	16,0	0/2
chr10: 61828425	k.A.	C/T	G4072S	Nichtsynonym	19,0	1/0

Position	SNP ID	Basen- aus- tausch	Amino- säure- austausch	Funktion	CADD- Wert	Het. Pat./ Het. Kontr.
chr10: 61828521	rs201115226	C/T	E4040K	Nichtsynonym	33,0	1/0
chr10: 61829273	rs144123544	T/C	N3789S	Nichtsynonym	19,6	1/2
chr10: 61829384	k.A.	C/T	C3752Y	Nichtsynonym	12,6	1/0
chr10: 61829549	rs148109897	G/C	S3697C	Nichtsynonym	16,1	2/0
chr10: 61829571	k.A.	C/T	G3690R	Nichtsynonym	23,0	0/2
chr10: 61829713	k.A.	C/G	K3642N	Nichtsynonym	27,0	1/0
chr10: 61829856	k.A.	T/C	T3595A	Nichtsynonym	24,7	0/1
chr10: 61829878	rs138986388	C/T	T3587T	Synonym	12,9	1/0
chr10: 61829951	rs147527383	T/C	E3563G	Nichtsynonym	22,9	3/3
chr10: 61830176	k.A.	T/C	E3488G	Nichtsynonym	23,3	1/0
chr10: 61830242	rs201398340	A/T	V3466D	Nichtsynonym	25,9	0/2
chr10: 61830270	k.A.	G/A	R3457C	Nichtsynonym	33,0	0/1
chr10: 61830419	k.A.	G/A	T3407M	Nichtsynonym	25,9	1/0
chr10: 61830421	k.A.	G/T	D3406E	Nichtsynonym	21,9	0/1
chr10: 61830584	rs61845768	T/C	E3352G	Nichtsynonym	14,2	16/10

Position	SNP ID	Basen- aus- tausch	Amino- säure- austausch	Funktion	CADD- Wert	Het. Pat./ Het. Kontr.
chr10: 61830642	rs140183285	T/A	T3333S	Nichtsynonym	15,9	3/3
chr10: 61830704	rs201625904	G/A	A3312V	Nichtsynonym	16,0	5/6
chr10: 61831150	k.A.	C/T	G3163G	Synonym	10,6	0/1
chr10: 61831298	rs181031970	A/G	V3114A	Nichtsynonym	17,2	0/1
chr10: 61831566	k.A.	C/A	V3025F	Nichtsynonym	17,7	1/0
chr10: 61832066	rs78194762	G/A	S2858L	Nichtsynonym	23,7	1/0
chr10: 61832105	rs139022925	C/T	G2845E	Nichtsynonym	25,5	2/3
chr10: 61832247	k.A.	C/T	E2798K	Nichtsynonym	22,8	0/2
chr10: 61832552	k.A.	G/A	S2696L	Nichtsynonym	22,1	0/1
chr10: 61832903	rs61732397	C/T	R2579K	Nichtsynonym	26,2	1/0
chr10: 61833170	rs140741466	G/A	P2490L	Nichtsynonym	11,1	0/1
chr10: 61833414	rs148904927	A/G	S2409P	Nichtsynonym	25,4	15/16
chr10: 61833564	k.A.	A/G	Y2359H	Nichtsynonym	22,4	1/0
chr10: 61833649	k.A.	T/C	I2330M	Nichtsynonym	23,0	0/1
chr10: 61833684	rs140463162	C/T	D2319N	Nichtsynonym	28,2	2/5

Position	SNP ID	Basen- aus- tausch	Amino- säure- austausch	Funktion	CADD- Wert	Het. Pat./ Het. Kontr.
chr10: 61833723	rs144270555	T/C	K2306E	Nichtsynonym	21,1	1/0
chr10: 61833749	k.A.	T/C	K2297R	Nichtsynonym	23,1	1/0
chr10: 61833851	k.A.	G/A	A2263V	Nichtsynonym	17,6	1/0
chr10: 61834084	rs144465058	C/A	Q2185H	Nichtsynonym	15,3	5/3
chr10: 61834098	rs368258557	C/A	G2181W	Nichtsynonym	28,9	2/0
chr10: 61834658	rs141251507	G/A	S1994L	Nichtsynonym	22,5	1/4
chr10: 61834700	k.A.	T/C	D1980G	Nichtsynonym	23,1	1/0
chr10: 61834861- 61834863	k.A.	CTC/-	E1926Del	Deletion ohne Rasterschub	21,7	0/1
chr10: 61835027	k.A.	C/T	R1871Q	Nichtsynonym	29,9	0/1
chr10: 61835057	rs117475706	G/A	T1861M	Nichtsynonym	24,5	6/12
chr10: 61835221	rs200232886	C/T	S1806S	Synonym	12,4	0/1
chr10: 61835429	k.A.	G/A	A1737V	Nichtsynonym	17,0	1/0
chr10: 61835462	k.A.	G/A	P1726L	Nichtsynonym	26,2	1/0
chr10: 61835518	rs202183196	C/T	M1707I	Nichtsynonym	11,8	0/1
chr10: 61835633	rs145794254	G/A	P1669L	Nichtsynonym	25,6	3/1

Position	SNP ID	Basen- aus- tausch	Amino- säure- austausch	Funktion	CADD- Wert	Het. Pat./ Het. Kontr.
chr10: 61835821	k.A.	C/T	T1606T	Synonym	11,5	1/1
chr10: 61835822	k.A.	G/A	T1606M	Nichtsynonym	22,6	0/1
chr10: 61835870	k.A.	T/C	Y1590C	Nichtsynonym	24,1	0/1
chr10: 61835921	rs201134561	G/A	S1573F	Nichtsynonym	23,3	0/1
chr10: 61835934	rs375050420	A/C	S1569A	Nichtsynonym	24,5	0/1
chr10: 61836109	k.A.	C/T	P1510P	Synonym	10,3	0/1
chr10: 61836110	k.A.	G/A	P1510L	Nichtsynonym	28,0	0/1
chr10: 61836174	rs41274676	G/A	P1489S	Nicht- synonym	17,7	0/8
chr10: 61836201	k.A.	G/A	R1480W	Nichtsynonym	31,0	1/0
chr10: 61840303	rs200130698	G/A	P1475L	Nichtsynonym	24,5	0/1
chr10: 61840327	rs139982397	T/C	K1467R	Nichtsynonym	23,6	1/2
chr10: 61840330	k.A.	C/T	R1466H	Nichtsynonym	35,0	0/1
chr10: 61840347	rs75052415	G/A	F1460F	Synonym	13,8	0/1
chr10: 61842437	rs201213520	G/C	T1420R	Nichtsynonym	25,2	1/0
chr10: 61842449	k.A.	T/C	K1416R	Nichtsynonym	23,1	0/1

Position	SNP ID	Basen- aus- tausch	Amino- säure- austausch	Funktion	CADD- Wert	Het. Pat./ Het. Kontr.
chr10: 61843365	rs200556767	C/G	G1362A	Nichtsynonym	28,4	1/1
chr10: 61844437	k.A.	T/C	M1333V	Nichtsynonym	22,2	2/1
chr10: 61844465	rs140104995	G/A	P1323P	Synonym	12,6	1/0
chr10: 61844537	rs150367434	C/T	T1299T	Synonym	15,1	3/3
chr10: 61848064	rs187349220	C/T	T1127T	Synonym	15,5	1/0
chr10: 61848104	k.A.	G/A	P1114L	Nichtsynonym	21,4	0/1
chr10: 61865807	k.A.	C/T	V1062M	Nichtsynonym	29,6	0/1
chr10: 61865821	k.A.	-/A	k.A.	k.A.	11,7	0/1
chr10: 61874016	k.A.	T/C	N972S	Nichtsynonym	25,3	1/0
chr10: 61894068	k.A.	G/T	S934R	Nichtsynonym	26,2	0/1
chr10: 61894126	k.A.	C/T	R915H	Nichtsynonym	27,5	1/0
chr10: 61905727	rs41283526	T/C	K872K+	Synonym	20,6	13/28
chr10: 61905734	k.A.	G/A	T870M+	Nichtsynonym	16,6	0/2
chr10: 61926345	k.A.	A/G	k.A.	k.A.	13,1	1/0
chr10: 61932131	k.A.	G/A	R805C	Nichtsynonym	27,2	0/1

Position	SNP ID	Basen- aus- tausch	Amino- säure- austausch	Funktion	CADD- Wert	Het. Pat./ Het. Kontr.
chr10: 61932749	k.A.	T/C	Y764C	Nichtsynonym	24,9	1/0
chr10: 61932932	k.A.	G/C	P733A	Nichtsynonym	24,6	0/1
chr10: 61941098	rs144051466	C/T	A725T	Nichtsynonym	23,5	4/7
chr10: 61956328	k.A.	T/C	N582S	Nichtsynonym	10,1	2/0
chr10: 61958119	rs16914613	C/T	A556A	Synonym	19,6	7/2
chr10: 61958253	k.A.	C/A	A512S	Nichtsynonym	25,0	0/1
chr10: 61962809	k.A.	A/G	I448T	Nichtsynonym	28,0	1/0
chr10: 61967941	rs117138204	G/A	C349C	Synonym	13,2	16/8
chr10: 62021659	rs117266082	C/T	T252T	Synonym	12,9	7/9
chr10: 62023621	rs200837964	T/C	N224S	Nichtsynonym	25,6	0/1
chr10: 62023637	k.A.	C/T	A219T	Nichtsynonym	32,0	0/1
chr10: 62023640	k.A.	C/A	A218S	Nichtsynonym	33,0	0/1
chr10: 62023647	rs145034480	C/T	T215T	Synonym	14,9	0/1
chr10: 62038547	rs77316697	G/A	V133V	Synonym	12,7	2/3
chr10: 62149205	rs141511303	C/T	R31Q	Nichtsynonym	22,8	1/0
chr10: 62374954	k.A.	C/T	S29N++	Nichtsynonym	15,3	1/3

Tab. 11: Übersicht über die 22 identifizierten seltenen (MAF <1 %) und potentiell funktionellen (CADD >10) Varianten in *GRIN2A*.

Position= Chromosomale Position (hg19). SNP ID= *Single Nucleotide Polymorphism Identification*. Basenaustausch= Referenz- und Alternatives Allel. Aminosäureaustausch= Aminosäureaustausch mit Position, Transkript GRIN2A:NM_001134407. Funktion= Funktionelle Konsequenz. CADD-Wert= *Combined Annotation Dependent Depletion Score*. Het. Pat./Het. Kontr.= Heterozygote Patienten/Heterozygote Kontrollen= Anzahl an Patienten bzw. Kontrollpersonen, in der die Variante heterozygot identifiziert wurde. k.A.= keine Angabe. Die Varianten sind aufsteigend nach chromosomaler Position sortiert. Die nominal signifikante Variante ist fett markiert.

Position	SNP ID	Basenaustausch	Aminosäureaustausch	Funktion	CADD-Wert	Het. Pat./ Het. Kontr.
chr16: 9857216	k.A.	C/T	R1281Q	Nichtsynonym	12,3	0/1
chr16: 9857517	k.A.	A/G	I1295T	Nichtsynonym	25,4	1/0
chr16: 9857574	rs145063086	G/C	A1276G	Nichtsynonym	11,8	3/1
chr16: 9858173	rs61758995	G/T	N1076K	Nichtsynonym	23,0	8/31
chr16: 9858396	k.A.	G/A	T1002I	Nichtsynonym	17,5	0/1
chr16: 9858493	k.A.	G/A	R970W	Nichtsynonym	26,6	1/0
chr16: 9858549	rs139795367	C/A	G951V	Nichtsynonym	19,8	2/1
chr16: 9858652	k.A.	A/C	S917A	Nichtsynonym	23,0	0/1
chr16: 9916206	k.A.	G/A	R695W	Nichtsynonym	34,0	0/1
chr16: 9923439	rs372324916	G/A	S616S	Synonym	13,3	0/1
chr16: 9927973	k.A.	G/C	A589G	Nichtsynonym	29,8	0/1

Position	SNP ID	Basen- aus- tausch	Amino- säure- aus- tausch	Funktion	CADD- Wert	Het. Pat./ Het. Kontr.
chr16: 9934916	rs368221866	C/T	G458G	Synonym	15,2	0/1
chr16: 9934950	k.A.	T/C	N447S	Nichtsynonym	12,4	1/1
chr16: 9943780	rs367543146	G/A	H387H	Synonym	11,5	1/0
chr16: 10032235	k.A.	C/T	V196V	Synonym	10,4	0/1
chr16: 10032269	k.A.	C/G	S185T	Nichtsynonym	14,9	0/1
chr16: 10032276	k.A.	A/T	F183I	Nichtsynonym	27,1	1/0
chr16: 10032376	rs144464666	C/T	A149A	Synonym	14,7	0/1
chr16: 10032377	k.A.	G/A	A149V	Nichtsynonym	25,3	1/0
chr16: 10032378	k.A.	C/G	A149P	Nichtsynonym	23,6	1/0
chr16: 10032400	k.A.	C/T	T141T	Synonym	13,7	0/1
chr16: 10032401	rs78631453	G/A	T141M	Nichtsynonym	12,7	4/6

7. Abbildungsverzeichnis

Abbildung 1: Symptome der affektiven Störungen.	10
Abbildung 2: Sequenziermethode nach Sanger.	14
Abbildung 3: Manhattan Plot der GWAS aus 2019, adaptiert von Stahl et al., 2018*, mit Darstellung der in der vorliegenden Studie ausgewählten Kandidatengene.	22
Abbildung 4: Aufbau eines smMIP-Primers.	29
Abbildung 5: Agarosegelbilder der fünf Test-PCR-Zyklen.	34
Abbildung 6: Vergleich des aufgereinigten mit dem unaufgereinigten Pool.	35
Abbildung 7: Erzielte <i>Coverage</i> der 579 smMIPs im Test- und Balancinglauf.	38
Abbildung 8: Bandenprodukte von fünf zufällig ausgewählten Proben einer 96er-Platte nach 24 PCR-Zyklen auf Agarosegel.	40

8. Tabellenverzeichnis

Tabelle 1: Identifizierte Kandidatengene.	23
Tabelle 2: Kandidatengene inklusive Genprodukt und Funktion.	27
Tabelle 3: Anzahl der smMIP-Primer pro Gen und Erstellung von Gen- und Gesamt-smMIP-Pool.	31
Tabelle 4: Übersicht der 860 identifizierten Varianten nach Qualitätsanalyse.	44
Tabelle 5: Nominal signifikante seltene Varianten gemäß Fisher's Exact Test.	46
Tabelle 6: Genbasierte P-Werte der <i>Burdenanalyse</i> .	49
Tabelle 7: RVIS Wert der Gene <i>ADD3</i> , <i>ANK3</i> und <i>GRIN2A</i> .	50
Tabelle 8: Vergleich der Ergebnisse der BipEx-Daten mit unserer Studie für die sieben seltenen Varianten mit nominal signifikanter Assoziation.	53
Tabelle 9: Übersicht über die 22 identifizierten seltenen (MAF <1 %) und potentiell funktionellen (CADD >10) Varianten in <i>ADD3</i> .	59
Tabelle 10: Übersicht über die 102 identifizierten seltenen (MAF <1 %) und potentiell funktionellen (CADD >10) Varianten in <i>ANK3</i> .	61
Tabelle 11: Übersicht über die 22 identifizierten seltenen (MAF <1 %) und potentiell funktionellen (CADD >10) Varianten in <i>GRIN2A</i> .	68

9. Literaturverzeichnis

Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* 2013; 7: 7.20

Aggarwal S, Yadav AK. False Discovery Rate Estimation in Proteomics. *Methods Mol Biol* 2016; 1362: 119-128

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5® (5th ed.). Am Psychiatr Publ, 2013

Bansal V, Libiger O, Torkamani A, Schork NJ. Statistical analysis strategies for association studies involving rare variants. *Nat Rev Genet* 2010; 11: 773-785

Behjati S, Tarpey PS. What is next generation sequencing?. *Arch Dis Child Educ Pract Ed* 2013; 98: 236–238

Biennu OJ, Davydov DS, Kendler KS. Psychiatric ‘diseases’ versus behavioral disorders and degree of genetic influence. *Psychol Med* 2011; 41: 33-40

Boyle EA, O’Roak BJ, Martin BK, Kumar A, Shendure J. MIPgen: optimized modeling and design of molecular inversion probes for targeted resequencing, *Bioinformatics* 2014; 30: 2670–2672

Brainstorm Consortium, Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, Duncan L, Escott-Price V, Falcone GJ, Gormley P, Malik R, Patsopoulos NA, Ripke S, Wei Z, Yu D, Lee PH, Turley P, Grenier-Boley B, Chouraki V, Kamatani Y, Berr C, Letenneur L, Hannequin D, Amouyel P, Boland A, Deleuze JF, Duron E, Vardarajan BN, Reitz C, Goate AM, Huentelman MJ, Kamboh MI, Larson EB, Rogaeva E, St George-Hyslop P, Hakonarson H, Kukull WA, Farrer LA, Barnes LL, Beach TG, Demirci FY, Head E, Hulette CM, Jicha GA, Kauwe JSK, Kaye JA, Leverenz JB, Levey AI, Lieberman AP, Pankratz VS, Poon WW, Quinn JF, Saykin AJ, Schneider LS, Smith AG, Sonnen JA, Stern RA, Van Deerlin VM, Van Eldik LJ, Harold D, Russo G, Rubinsztein DC, Bayer A, Tsolaki M, Proitsi P, Fox NC, Hampel H, Owen MJ, Mead S, Passmore P, Morgan K, Nöthen MM,

Rossor M, Lupton MK, Hoffmann P, Kornhuber J, Lawlor B, McQuillin A, Al-Chalabi A, Bis JC, Ruiz A, Boada M, Seshadri S, Beiser A, Rice K, van der Lee SJ, De Jager PL, Geschwind DH, Riemenschneider M, Riedel-Heller S, Rotter JI, Ransmayr G, Hyman BT, Cruchaga C, Alegret M, Winsvold B, Palta P, Farh KH, Cuenca-Leon E, Furlotte N, Kurth T, Ligthart L, Terwindt GM, Freilinger T, Ran C, Gordon SD, Borck G, Adams HHH, Lehtimäki T, Wedenoja J, Buring JE, Schürks M, Hrafnisdottir M, Hottenga JJ, Penninx B, Artto V, Kaunisto M, Vepsäläinen S, Martin NG, Montgomery GW, Kurki MI, Hämäläinen E, Huang H, Huang J, Sandor C, Webber C, Muller-Myhsok B, Schreiber S, Salomaa V, Loehrer E, Göbel H, Macaya A, Pozo-Rosich P, Hansen T, Werge T, Kaprio J, Metspalu A, Kubisch C, Ferrari MD, Belin AC, van den Maagdenberg AMJM, Zwart JA, Boomsma D, Eriksson N, Olesen J, Chasman DI, Nyholt DR, Avbersek A, Baum L, Berkovic S, Bradfield J, Buono RJ, Catarino CB, Cossette P, De Jonghe P, Depondt C, Dlugos D, Ferraro TN, French J, Hjalgrim H, Jamnadas-Khoda J, Kälviäinen R, Kunz WS, Lerche H, Leu C, Lindhout D, Lo W, Lowenstein D, McCormack M, Møller RS, Molloy A, Ng PW, Oliver K, Privitera M, Radtke R, Ruppert AK, Sander T, Schachter S, Schankin C, Scheffer I, Schoch S, Sisodiya SM, Smith P, Sperling M, Striano P, Surges R, Thomas GN, Visscher F, Whelan CD, Zara F, Heinzen EL, Marson A, Becker F, Stroink H, Zimprich F, Gasser T, Gibbs R, Heutink P, Martinez M, Morris HR, Sharma M, Ryten M, Mok KY, Pulit S, Bevan S, Holliday E, Attia J, Battey T, Boncoraglio G, Thijs V, Chen WM, Mitchell B, Rothwell P, Sharma P, Sudlow C, Vicente A, Markus H, Kourkoulis C, Pera J, Raffeld M, Silliman S, Boraska Perica V, Thornton LM, Huckins LM, William Rayner N, Lewis CM, Gratacos M, Rybakowski F, Keski-Rahkonen A, Raevuori A, Hudson JI, Reichborn-Kjennerud T, Monteleone P, Karwautz A, Mannik K, Baker JH, O'Toole JK, Trace SE, Davis OSP, Helder SG, Ehrlich S, Herpertz-Dahlmann B, Danner UN, van Elburg AA, Clementi M, Forzan M, Docampo E, Lissowska J, Hauser J, Tortorella A, Maj M, Gonidakis F, Tziouvas K, Papezova H, Yilmaz Z, Wagner G, Cohen-Woods S, Herms S, Julià A, Rabionet R, Dick DM, Ripatti S, Andreassen OA, Espeseth T, Lundervold AJ, Steen VM, Pinto D, Scherer SW, Aschauer H, Schosser A, Alfredsson L, Padyukov L, Halmi KA, Mitchell J, Strober M, Bergen AW, Kaye W, Szatkiewicz JP, Cormand B, Ramos-Quiroga JA, Sánchez-Mora C, Ribasés M, Casas M, Hervas A, Arranz MJ, Haavik J, Zayats T, Johansson S, Williams N, Dempfle A, Rothenberger A, Kuntsi J, Oades RD, Banaschewski T, Franke B, Buitelaar JK, Arias Vasquez A, Doyle AE, Reif A, Lesch KP,

Freitag C, Rivero O, Palmason H, Romanos M, Langley K, Rietschel M, Witt SH, Dalsgaard S, Børglum AD, Waldman I, Wilmot B, Molly N, Bau CHD, Crosbie J, Schachar R, Loo SK, McGough JJ, Grevet EH, Medland SE, Robinson E, Weiss LA, Bacchelli E, Bailey A, Bal V, Battaglia A, Betancur C, Bolton P, Cantor R, Celestino-Soper P, Dawson G, De Rubeis S, Duque F, Green A, Klauck SM, Leboyer M, Levitt P, Maestrini E, Mane S, De-Luca DM, Parr J, Regan R, Reichenberg A, Sandin S, Vorstman J, Wassink T, Wijsman E, Cook E, Santangelo S, Delorme R, Rogé B, Magalhaes T, Arking D, Schulze TG, Thompson RC, Strohmaier J, Matthews K, Melle I, Morris D, Blackwood D, McIntosh A, Bergen SE, Schalling M, Jamain S, Maaser A, Fischer SB, Reinbold CS, Fullerton JM, Guzman-Parra J, Mayoral F, Schofield PR, Cichon S, Mühleisen TW, Degenhardt F, Schumacher J, Bauer M, Mitchell PB, Gershon ES, Rice J, Potash JB, Zandi PP, Craddock N, Ferrier IN, Alda M, Rouleau GA, Turecki G, Ophoff R, Pato C, Anjorin A, Stahl E, Leber M, Czerski PM, Cruceanu C, Jones IR, Posthuma D, Andlauer TFM, Forstner AJ, Streit F, Baune BT, Air T, Sinnamon G, Wray NR, MacIntyre DJ, Porteous D, Homuth G, Rivera M, Grove J, Middeldorp CM, Hickie I, Pergadia M, Mehta D, Smit JH, Jansen R, de Geus E, Dunn E, Li QS, Nauck M, Schoevers RA, Beekman AT, Knowles JA, Viktorin A, Arnold P, Barr CL, Bedoya-Berrio G, Bienvenu OJ, Brentani H, Burton C, Camarena B, Cappi C, Cath D, Cavallini M, Cusi D, Darrow S, Denys D, Derks EM, Dietrich A, Fernandez T, Figeo M, Freimer N, Gerber G, Grados M, Greenberg E, Hanna GL, Hartmann A, Hirschtritt ME, Hoekstra PJ, Huang A, Huyser C, Illmann C, Jenike M, Kuperman S, Leventhal B, Lochner C, Lyon GJ, Macciardi F, Madruga-Garrido M, Malaty IA, Maras A, McGrath L, Miguel EC, Mir P, Nestadt G, Nicolini H, Okun MS, Pakstis A, Paschou P, Piacentini J, Pittenger C, Plessen K, Ramensky V, Ramos EM, Reus V, Richter MA, Riddle MA, Robertson MM, Roessner V, Rosário M, Samuels JF, Sandor P, Stein DJ, Tsetsos F, Van Nieuwerburgh F, Weatherall S, Wendland JR, Wolanczyk T, Worbe Y, Zai G, Goes FS, McLaughlin N, Nestadt PS, Grabe HJ, Depienne C, Konkashbaev A, Lanzagorta N, Valencia-Duarte A, Bramon E, Buccola N, Cahn W, Cairns M, Chong SA, Cohen D, Crespo-Facorro B, Crowley J, Davidson M, DeLisi L, Dinan T, Donohoe G, Drapeau E, Duan J, Haan L, Hougaard D, Karachanak-Yankova S, Khrunin A, Klovins J, Kučinskas V, Lee Chee Keong J, Limborska S, Loughland C, Lönnqvist J, Maher B, Mattheisen M, McDonald C, Murphy KC, Nenadic I, van Os J, Pantelis C, Pato M, Petryshen T, Quedsted D, Roussos P, Sanders AR, Schall U, Schwab SG, Sim K, So

HC, Stögmann E, Subramaniam M, Toncheva D, Waddington J, Walters J, Weiser M, Cheng W, Cloninger R, Curtis D, Gejman PV, Henskens F, Mattingsdal M, Oh SY, Scott R, Webb B, Breen G, Churchhouse C, Bulik CM, Daly M, Dichgans M, Faraone SV, Guerreiro R, Holmans P, Kendler KS, Koeleman B, Mathews CA, Price A, Scharf J, Sklar P, Williams J, Wood NW, Cotsapas C, Palotie A, Smoller JW, Sullivan P, Rosand J, Corvin A, Neale BM, Schott JM, Anney R, Elia J, Grigoriou-Serbanescu M, Edenberg HJ, Murray R. Analysis of shared heritability in common disorders of the brain. *Science* 2018; 360: eaap8757

Budde M, Forstner AJ, Adorjan K, Schaupp SK, Nöthen MM, Schulze TG. Genetische Grundlagen der bipolaren Störung [Genetics of bipolar disorder]. *Nervenarzt* 2017; 88: 755-759

Charney AW, Ruderfer DM, Stahl EA, Moran JL, Chambert K, Belliveau RA, Forty L, Gordon-Smith K, Di Florio A, Lee PH, Bromet EJ, Buckley PF, Escamilla MA, Fanous AH, Fochtmann LJ, Lehrer DS, Malaspina D, Marder SR, Morley CP, Nicolini H, Perkins DO, Rakofsky JJ, Rapaport MH, Medeiros H, Sobell JL, Green EK, Backlund L, Bergen SE, Juréus A, Schalling M, Lichtenstein P, Roussos P, Knowles JA, Jones I, Jones LA, Hultman CM, Perlis RH, Purcell SM, McCarroll SA, Pato CN, Pato MT, Craddock N, Landén M, Smoller JW, Sklar P. Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder. *Transl Psychiatry* 2017; 7: e993

Cichon S, Mühleisen TW, Degenhardt FA, Mattheisen M, Miró X, Strohmaier J, Steffens M, Meesters C, Herms S, Weingarten M, Priebe L, Haenisch B, Alexander M, Vollmer J, Breuer R, Schmääl C, Tessmann P, Moebus S, Wichmann HE, Schreiber S, Müller-Myhsok B, Lucae S, Jamain S, Leboyer M, Bellivier F, Etain B, Henry C, Kahn JP, Heath S; Bipolar Disorder Genome Study (BiGS) Consortium; Hamshere M, O'Donovan MC, Owen MJ, Craddock N, Schwarz M, Vedder H, Kammerer-Ciernioch J, Reif A, Sasse J, Bauer M, Hautzinger M, Wright A, Mitchell PB, Schofield PR, Montgomery GW, Medland SE, Gordon SD, Martin NG, Gustafsson O, Andreassen O, Djurovic S, Sigurdsson E, Steinberg S, Stefansson H, Stefansson K, Kapur-Pojkic L, Oruc L, Rivas F, Mayoral F, Chuchalin A, Babadjanova G, Tiganov AS, Pantelejeva G, Abramova LI, Grigoriou-Serbanescu M, Diaconu CC, Czerski PM, Hauser J, Zimmer A, Lathrop M, Schulze TG,

Wienker TF, Schumacher J, Maier W, Propping P, Rietschel M, Nöthen MM. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am J Med Genet* 2011; 88: 372-381

Craddock N, Sklar P. Genetics of bipolar disorder. *Lancet* 2013; 381:1654-1662

Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* 2019; 179: 1469-1482.e11

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R; 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics* 2011; 27: 2156-2158

Fiorentino A, O'Brien NL, Locke DP, McQuillin A, Jarram A, Anjorin A, Kandaswamy R, Curtis D, Blizard RA, Gurling HM. Analysis of ANK3 and CACNA1C variants identified in bipolar disorder whole genome sequence data. *Bipolar Disord* 2014; 16: 583-591

Forstner AJ, Hofmann A, Maaser A, Sumer S, Khudayberdiev S, Mühleisen TW, Leber M, Schulze TG, Strohmaier J, Degenhardt F, Treutlein J, Mattheisen M, Schumacher J, Breuer R, Meier S, Herms S, Hoffmann P, Lacour A, Witt SH, Reif A, Müller-Myhsok B, Lucae S, Maier W, Schwarz M, Vedder H, Kammerer-Ciernioch J, Pfennig A, Bauer M, Hautzinger M, Moebus S, Priebe L, Sivalingam S, Verhaert A, Schulz H, Czerski PM, Hauser J, Lissowska J, Szeszenia-Dabrowska N, Brennan P, McKay JD, Wright A, Mitchell PB, Fullerton JM, Schofield PR, Montgomery GW, Medland SE, Gordon SD, Martin NG, Krasnov V, Chuchalin A, Babadjanova G, Pantelejeva G, Abramova LI, Tiganov AS, Polonikov A, Khusnutdinova E, Alda M, Cruceanu C, Rouleau GA, Turecki G, Laprise C, Rivas F, Mayoral F, Kogevinas M, Grigoriu-Serbanescu M, Propping P, Becker T, Rietschel M, Cichon S, Schratt G, Nöthen MM. Genome-wide analysis implicates microRNAs and their target genes in the development of bipolar disorder. *Transl Psychiatry* 2015; 5: e678

Forstner AJ, Fischer SB, Schenk LM, Strohmaier J, Maaser-Hecker A, Reinbold CS, Sivalingam S, Hecker J, Streit F, Degenhardt F, Witt SH, Schumacher J, Thiele H, Nürnberg P, Guzman-Parra J, Orozco Diaz G, Auburger G, Albus M, Borrmann-

Hassenbach M, González MJ, Gil Flores S, Cabaleiro Fabeiro FJ, Del Río Noriega F, Perez Perez F, Haro González J, Rivas F, Mayoral F, Bauer M, Pfennig A, Reif A, Herms S, Hoffmann P, Pirooznia M, Goes FS, Rietschel M, Nöthen MM, Cichon S. Whole-exome sequencing of 81 individuals from 27 multiply affected bipolar disorder families. *Transl Psychiatry* 2020; 10: 57

Forstner AJ, Hoffmann P, Nöthen MM, Cichon S. Insights into the genomics of affective disorders. *Med Genet* 2020; 32: 9-18

Frías Á, Baltasar I, Birmaher B. Comorbidity between bipolar disorder and borderline personality disorder: Prevalence, explanatory theories, and clinical impact. *J Affect Disord* 2016; 202: 210–219

Giudicessi JR, Ackerman MJ. Determinants of incomplete penetrance and variable expressivity in heritable cardiac arrhythmia syndromes. *Transl Res.* 2013; 161: 1-14

GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016; 388: 1545-1602

Georgi B, Craig D, Kember RL, Liu W, Lindquist I, Nasser S, Brown C, Egeland JA, Paul SM, Bućan M. Genomic view of bipolar disorder revealed by whole genome sequencing in a genetic isolate. *PLoS Genet* 2014; 10: e1004229

Goes FS. Genetics of Bipolar Disorder: Recent Update and Future Directions. *Psychiatr Clin North Am* 2016; 39: 139-155

Green EK, Rees E, Walters JT, Smith KG, Forty L, Grozeva D, Moran JL, Sklar P, Ripke S, Chambert KD, Genovese G, McCarroll SA, Jones I, Jones L, Owen MJ, O'Donovan MC, Craddock N, Kirov G. Copy number variation in bipolar disorder. *Mol Psychiatry* 2016; 21: 89-93

GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013; 45: 580-585

Harrison PJ, Cipriani A, Harmer CJ, Nobre AC, Saunders K, Goodwin GM, Geddes JR. Innovative approaches to bipolar disorder and its treatment. *Ann N Y Acad Sci* 2016; 1366: 76-89

Hiatt JB, Pritchard CC, Salipante SJ, O'Roak BJ, Shendure J. Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. *Genome Res* 2013; 23: 843-854

Hoffmeister, H, 2022a: Freies Lehrbuch: Einführung in die Biologie (für die Sekundarstufe 1 und 2). Kapitel 08.01: Zelluläre Grundlagen der Vererbung. https://hoffmeister.it/biologie/08.01zellulaere_grundlagen_der_vererbung.pdf (Zugriffsdatum: 29.03.2023)

Hoffmeister, H, 2022b: Freies Lehrbuch: Einführung in die Biologie (für die Sekundarstufe 1 und 2). Kapitel 08.03: Mendel und die klassische Genetik I. https://hoffmeister.it/biologie/08.03mendel_und_die_klassische_genetik_i.pdf (Zugriffsdatum: 02.04.2023)

Hoffmeister, H, 2022c: Freies Lehrbuch: Einführung in die Biologie (für die Sekundarstufe 1 und 2). Kapitel 08.12: Molekulare Grundlagen der Vererbung (Molekulargenetik). [https://hoffmeister.it/biologie/08.12molekulare_grundlagen_der_vererbung_\(molekulargenetik\).pdf](https://hoffmeister.it/biologie/08.12molekulare_grundlagen_der_vererbung_(molekulargenetik).pdf) (Zugriffsdatum: 29.03.2023)

Hoffmeister, H, 2022d: Freies Lehrbuch: Einführung in die Biologie (für die Sekundarstufe 1 und 2). Kapitel 08.14: Variabilität des Erbguts (Mutationen, Krebs, DNA-Reparatur). https://hoffmeister.it/biologie/08.14variabilitaet_des_erbguts.pdf (Zugriffsdatum: 29.03.2023)

Hughes T, Sønderby IE, Polushina T, Hansson L, Holmgren A, Athanasiu L, Melbø-Jørgensen C, Hassani S, Hoeffding LK, Herms S, Bergen SE, Karlsson R, Song J, Rietschel M, Nöthen MM, Forstner AJ, Hoffmann P, Hultman CM, Landén M, Cichon S, Werge T, Andreassen OA, Le Hellard S, Djurovic S. Elevated expression of a minor isoform of ANK3 is a risk factor for bipolar disorder. *Transl Psychiatry* 2018; 8: 210

Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. *Am J Hum Genet* 2013; 92 :841-853

Iqbal Z, Vandeweyer G, van der Voet M, Waryah AM, Zahoor MY, Besseling JA, Roca LT, Vulto-van Silfhout AT, Nijhof B, Kramer JM, Van der Aa N, Ansar M, Peeters H, Helsmoortel C, Gilissen C, Vissers LE, Veltman JA, de Brouwer AP, Frank Kooy R, Riazuddin S, Schenck A, van Bokhoven H, Rooms L. Homozygous and heterozygous disruptions of ANK3: at the crossroads of neurodevelopmental and psychiatric disorders. *Hum Mol Genet* 2013; 22: 1960-1970

Joslyn C, Hawes DJ, Hunt C, Mitchell PB. Is age of onset associated with severity, prognosis, and clinical features in bipolar disorder? A meta-analytic review. *Bipolar Disord* 2016; 18: 389-403

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res* 2002; 12: 996-1006

Kim HY. Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restor Dent Endod* 2017; 42: 152-155

Kircher T, Wöhr M, Nenadic I, Schwarting R, Schrott G, Alferink J, Culmsee C, Garn H, Hahn T, Müller-Myhsok B, Dempfle A, Hahmann M, Jansen A, Pfefferle P, Renz H, Rietschel M, Witt SH, Nöthen M, Krug A, Dannlowski U. Neurobiology of the major psychoses: a translational perspective on brain structure and function-the FOR2107 consortium. *Eur Arch Psychiatry Clin Neurosci* 2019; 269: 949-962

Kloth K, Denecke J, Hempel M, Johannsen J, Strom TM, Kubisch C, Lessel D. First de novo ANK3 nonsense mutation in a boy with intellectual disability, speech impairment and autistic features. *Eur J Med Genet* 2017; 60: 494-498

Kruer MC, Jepperson T, Dutta S, Steiner RD, Cottenie E, Sanford L, Merkens M, Russman BS, Blasco PA, Fan G, Pollock J, Green S, Woltjer RL, Mooney C, Kretzschmar D, Paisán-Ruiz C, Houlden H. Mutations in γ adducin are associated with inherited cerebral palsy. *Ann Neurol* 2013; 74: 805-814

Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA; NHLBI GO Exome Sequencing Project—ESP Lung Project Team; Christiani DC, Wurfel MM, Lin X. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet* 2012; 91: 224-237

Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. *Biostatistics* 2012; 13: 762-775

Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 2011; 88: 294-305

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatic* 2009; 25: 1754-1760

Lykke-Andersen S, Jensen TH. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol* 2015; 16: 665-677

Maaser A, Forstner AJ, Strohmaier J, Hecker J, Ludwig KU, Sivalingam S, Streit F, Degenhardt F, Witt SH, Reinbold CS, Koller AC, Raff R, Heilmann-Heimbach S, Fischer SB; Bipolar Disorder Working Group of the Psychiatric Genomics Consortium; Herms S, Hoffmann P, Thiele H, Nürnberg P, Löhlein Fier H, Orozco-Díaz G, Carmenate-Naranjo D, Proenza-Barzaga N, Auburger GWJ, Andlauer TFM, Cichon S, Marcheco-Teruel B, Mors O, Rietschel M, Nöthen MM. Exome sequencing in large, multiplex bipolar disorder families from Cuba. *PLoS One* 2018; 13: e0205895

Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, Antaki D, Shetty A, Holmans PA, Pinto D, Gujral M, Brandler WM, Malhotra D, Wang Z, Fajardo KVF, Maile MS, Ripke S, Agartz I, Albus M, Alexander M, Amin F, Atkins J, Bacanu SA, Belliveau RA Jr, Bergen SE, Bertalan M, Bevilacqua E, Bigdeli TB, Black DW, Bruggeman R, Buccola NG, Buckner RL, Bulik-Sullivan B, Byerley W, Cahn W, Cai G, Cairns MJ, Champion D, Cantor RM, Carr VJ, Carrera N, Catts SV, Chambert KD, Cheng W, Cloninger CR, Cohen D, Cormican P, Craddock N, Crespo-Facorro B, Crowley JJ, Curtis D, Davidson M, Davis KL, Degenhardt F, Del Favero J, DeLisi LE, Dikeos D, Dinan T, Djurovic S, Donohoe G, Drapeau E, Duan J, Dudbridge F, Eichhammer P, Eriksson J,

Escott-Price V, Essioux L, Fanous AH, Farh KH, Farrell MS, Frank J, Franke L, Freedman R, Freimer NB, Friedman JI, Forstner AJ, Fromer M, Genovese G, Georgieva L, Gershon ES, Giegling I, Giusti-Rodríguez P, Godard S, Goldstein JI, Gratten J, de Haan L, Hamshere ML, Hansen M, Hansen T, Haroutunian V, Hartmann AM, Henskens FA, Herms S, Hirschhorn JN, Hoffmann P, Hofman A, Huang H, Ikeda M, Joa I, Kähler AK, Kahn RS, Kalaydjieva L, Karjalainen J, Kavanagh D, Keller MC, Kelly BJ, Kennedy JL, Kim Y, Knowles JA, Konte B, Laurent C, Lee P, Lee SH, Legge SE, Lerer B, Levy DL, Liang KY, Lieberman J, Lönngqvist J, Loughland CM, Magnusson PKE, Maher BS, Maier W, Mallet J, Mattheisen M, Mattingsdal M, McCarley RW, McDonald C, McIntosh AM, Meier S, Meijer CJ, Melle I, Meshulam-Gately RI, Metspalu A, Michie PT, Milani L, Milanova V, Mokrab Y, Morris DW, Müller-Myhsok B, Murphy KC, Murray RM, Myin-Germeys I, Nenadic I, Nertney DA, Nestadt G, Nicodemus KK, Nisenbaum L, Nordin A, O'Callaghan E, O'Dushlaine C, Oh SY, Olincy A, Olsen L, O'Neill FA, Van Os J, Pantelis C, Papadimitriou GN, Parkhomenko E, Pato MT, Paunio T; Psychosis Endophenotypes International Consortium; Perkins DO, Pers TH, Pietiläinen O, Pimm J, Pocklington AJ, Powell J, Price A, Pulver AE, Purcell SM, Quested D, Rasmussen HB, Reichenberg A, Reimers MA, Richards AL, Roffman JL, Roussos P, Ruderfer DM, Salomaa V, Sanders AR, Savitz A, Schall U, Schulze TG, Schwab SG, Scolnick EM, Scott RJ, Seidman LJ, Shi J, Silverman JM, Smoller JW, Söderman E, Spencer CCA, Stahl EA, Strengman E, Strohmaier J, Stroup TS, Suvisaari J, Svrakic DM, Szatkiewicz JP, Thirumalai S, Tooney PA, Veijola J, Visscher PM, Waddington J, Walsh D, Webb BT, Weiser M, Wildenauer DB, Williams NM, Williams S, Witt SH, Wolen AR, Wormley BK, Wray NR, Wu JQ, Zai CC, Adolfsson R, Andreassen OA, Blackwood DHR, Bramon E, Buxbaum JD, Cichon S, Collier DA, Corvin A, Daly MJ, Darvasi A, Domenici E, Esko T, Gejman PV, Gill M, Gurling H, Hultman CM, Iwata N, Jablensky AV, Jönsson EG, Kendler KS, Kirov G, Knight J, Levinson DF, Li QS, McCarroll SA, McQuillin A, Moran JL, Mowry BJ, Nöthen MM, Ophoff RA, Owen MJ, Palotie A, Pato CN, Petryshen TL, Posthuma D, Rietschel M, Riley BP, Rujescu D, Sklar P, St Clair D, Walters JTR, Werge T, Sullivan PF, O'Donovan MC, Scherer SW, Neale BM, Sebat J. CNV and Schizophrenia Working Groups of the Psychiatric Genomics Consortium. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet* 2017; 49: 27-35

Merikangas KR, Jin R, He JP, Kessler RC, Lee S, Sampson NA, Viana MC, Andrade LH, Hu C, Karam EG, Ladea M, Medina-Mora ME, Ono Y, Posada-Villa J, Sagar R, Wells JE, Zarkov Z. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gen Psychiatry* 2011; 68: 241-251

Moskvina V, Holmans P, Schmidt KM, Craddock N. Design of case-controls studies with unscreened controls. *Ann Hum Genet.* 2005; 69: 566-576

Mühleisen TW, Leber M, Schulze TG, Strohmaier J, Degenhardt F, Treutlein J, Mattheisen M, Forstner AJ, Schumacher J, Breuer R, Meier S, Herms S, Hoffmann P, Lacour A, Witt SH, Reif A, Müller-Myhsok B, Lucae S, Maier W, Schwarz M, Vedder H, Kammerer-Ciernioch J, Pfennig A, Bauer M, Hautzinger M, Moebus S, Priebe L, Czerski PM, Hauser J, Lissowska J, Szeszenia-Dabrowska N, Brennan P, McKay JD, Wright A, Mitchell PB, Fullerton JM, Schofield PR, Montgomery GW, Medland SE, Gordon SD, Martin NG, Krasnow V, Chuchalin A, Babadjanova G, Pantelejeva G, Abramova LI, Tiganov AS, Polonikov A, Khusnutdinova E, Alda M, Grof P, Rouleau GA, Turecki G, Laprise C, Rivas F, Mayoral F, Kogevinas M, Grigoriu-Serbanescu M, Propping P, Becker T, Rietschel M, Nöthen MM, Cichon S. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun* 2014; 5: 3339

Mühleisen TW, Reinbold CS, Forstner AJ, Abramova LI, Alda M, Babadjanova G, Bauer M, Brennan P, Chuchalin A, Cruceanu C, Czerski PM, Degenhardt F, Fischer SB, Fullerton JM, Gordon SD, Grigoriu-Serbanescu M, Grof P, Hauser J, Hautzinger M, Herms S, Hoffmann P, Kammerer-Ciernioch J, Khusnutdinova E, Kogevinas M, Krasnow V, Lacour A, Laprise C, Leber M, Lissowska J, Lucae S, Maaser A, Maier W, Martin NG, Mattheisen M, Mayoral F, McKay JD, Medland SE, Mitchell PB, Moebus S, Montgomery GW, Müller-Myhsok B, Oruc L, Pantelejeva G, Pfennig A, Pojskic L, Polonikov A, Reif A, Rivas F, Rouleau GA, Schenk LM, Schofield PR, Schwarz M, Streit F, Strohmaier J, Szeszenia-Dabrowska N, Tiganov AS, Treutlein J, Turecki G, Vedder H, Witt SH, Schulze TG, Rietschel M, Nöthen MM, Cichon S. Gene set enrichment analysis and expression pattern exploration implicate an involvement of neurodevelopmental processes in bipolar disorder. *J Affect Disord* 2018; 228: 20-25

Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, Qiao Z, Als TD, Bigdeli TB, Børte S, Bryois J, Charney AW, Drange OK, Gandal MJ, Hagenaars SP, Ikeda M, Kamitaki N, Kim M, Krebs K, Panagiotaropoulou G, Schilder BM, Sloofman LG, Steinberg S, Trubetskoy V, Winsvold BS, Won HH, Abramova L, Adorjan K, Agerbo E, Al Eissa M, Albani D, Alliey-Rodriguez N, Anjorin A, Antilla V, Antoniou A, Awasthi S, Baek JH, Bækvad-Hansen M, Bass N, Bauer M, Beins EC, Bergen SE, Birner A, Bøcker Pedersen C, Bøen E, Boks MP, Bosch R, Brum M, Brumpton BM, Brunkhorst-Kanaan N, Budde M, Bybjerg-Grauholm J, Byerley W, Cairns M, Casas M, Cervantes P, Clarke TK, Cruceanu C, Cuellar-Barboza A, Cunningham J, Curtis D, Czerski PM, Dale AM, Dalkner N, David FS, Degenhardt F, Djurovic S, Dobbyn AL, Douzenis A, Elvsåshagen T, Escott-Price V, Ferrier IN, Fiorentino A, Foroud TM, Forty L, Frank J, Frei O, Freimer NB, Frisé L, Gade K, Garnham J, Gelernter J, Giørtz Pedersen M, Gizer IR, Gordon SD, Gordon-Smith K, Greenwood TA, Grove J, Guzman-Parra J, Ha K, Haraldsson M, Hautzinger M, Heilbronner U, Hellgren D, Herms S, Hoffmann P, Holmans PA, Huckins L, Jamain S, Johnson JS, Kalman JL, Kamatani Y, Kennedy JL, Kittel-Schneider S, Knowles JA, Kogevinas M, Koromina M, Kranz TM, Kranzler HR, Kubo M, Kupka R, Kushner SA, Lavebratt C, Lawrence J, Leber M, Lee HJ, Lee PH, Levy SE, Lewis C, Liao C, Lucae S, Lundberg M, MacIntyre DJ, Magnusson SH, Maier W, Maihofer A, Malaspina D, Maratou E, Martinsson L, Mattheisen M, McCarroll SA, McGregor NW, McGuffin P, McKay JD, Medeiros H, Medland SE, Millischer V, Montgomery GW, Moran JL, Morris DW, Mühleisen TW, O'Brien N, O'Donovan C, Olde Loohuis LM, Oruc L, Papiol S, Pardiñas AF, Perry A, Pfennig A, Porichi E, Potash JB, Quedsted D, Raj T, Rapaport MH, DePaulo JR, Regeer EJ, Rice JP, Rivas F, Rivera M, Roth J, Roussos P, Ruderfer DM, Sánchez-Mora C, Schulte EC, Senner F, Sharp S, Shilling PD, Sigurdsson E, Sirignano L, Slaney C, Smeland OB, Smith DJ, Sobell JL, Sørholm Hansen C, Soler Artigas M, Spijker AT, Stein DJ, Strauss JS, Świątkowska B, Terao C, Thorgeirsson TE, Toma C, Tooney P, Tsermpini EE, Vawter MP, Vedder H, Walters JTR, Witt SH, Xi S, Xu W, Yang JMK, Young AH, Young H, Zandi PP, Zhou H, Zillich L; HUNT All-In Psychiatry; Adolfsson R, Agartz I, Alda M, Alfredsson L, Babadjanova G, Backlund L, Baune BT, Bellivier F, Bengesser S, Berrettini WH, Blackwood DHR, Boehnke M, Børghlum AD, Breen G, Carr VJ, Catts S, Corvin A, Craddock N, Dannlowski U, Dikeos D, Esko T, Etain B, Ferentinos P, Frye M, Fullerton JM, Gawlik M, Gershon ES, Goes FS, Green MJ, Grigoriou-Serbanescu M,

Hauser J, Henskens F, Hillert J, Hong KS, Hougaard DM, Hultman CM, Hveem K, Iwata N, Jablensky AV, Jones I, Jones LA, Kahn RS, Kelsoe JR, Kirov G, Landén M, Leboyer M, Lewis CM, Li QS, Lissowska J, Lochner C, Loughland C, Martin NG, Mathews CA, Mayoral F, McElroy SL, McIntosh AM, McMahon FJ, Melle I, Michie P, Milani L, Mitchell PB, Morken G, Mors O, Mortensen PB, Mowry B, Müller-Myhsok B, Myers RM, Neale BM, Nievergelt CM, Nordentoft M, Nöthen MM, O'Donovan MC, Oedegaard KJ, Olsson T, Owen MJ, Paciga SA, Pantelis C, Pato C, Pato MT, Patrinos GP, Perlis RH, Posthuma D, Ramos-Quiroga JA, Reif A, Reininghaus EZ, Ribasés M, Rietschel M, Ripke S, Rouleau GA, Saito T, Schall U, Schalling M, Schofield PR, Schulze TG, Scott LJ, Scott RJ, Serretti A, Shannon Weickert C, Smoller JW, Stefansson H, Stefansson K, Stordal E, Streit F, Sullivan PF, Turecki G, Vaaler AE, Vieta E, Vincent JB, Waldman ID, Weickert TW, Werge T, Wray NR, Zwart JA, Biernacka JM, Nurnberger JI, Cichon S, Edenberg HJ, Stahl EA, McQuillin A, Di Florio A, Ophoff RA, Andreassen OA. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet* 2021; 53: 817-829

Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci* 2015; 18: 199-209

Nurnberger JI Jr, Koller DL, Jung J, Edenberg HJ, Foroud T, Guella I, Vawter MP, Kelsoe JR; Psychiatric Genomics Consortium Bipolar Group. Identification of pathways for bipolar disorder: a meta-analysis. *JAMA Psychiatry* 2014; 71: 657-664

O'Connell KS, Coombes BJ. Genetic contributions to bipolar disorder: current status and future directions. *Psychol Med* 2021; 51: 2156-2167

Palmer DS, Howrigan DP, Chapman SB, Adolfsson R, Bass N, Blackwood D, Boks MPM, Chen CY, Churchhouse C, Corvin AP, Craddock N, Curtis D, Di Florio A, Dickerson F, Freimer NB, Goes FS, Jia X, Jones I, Jones L, Jonsson L, Kahn RS, Landén M, Locke AE, McIntosh AM, McQuillin A, Morris DW, O'Donovan MC, Ophoff RA, Owen MJ, Pedersen NL, Posthuma D, Reif A, Risch N, Schaefer C, Scott L, Singh T, Smoller JW, Solomonson M, Clair DS, Stahl EA, Vreeker A, Walters JTR, Wang W, Watts NA, Yolken

R, Zandi PP, Neale BM. Exome sequencing in bipolar disorder identifies AKAP11 as a risk gene shared with schizophrenia. *Nat Genet* 2022; 54: 541-547

Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet* 2013; 9: e1003709

Poltavskaya EG, Fedorenko OY, Kornetova EG, Loonen AJM, Kornetov AN, Bokhan NA, Ivanova SA. Study of Early Onset Schizophrenia: Associations of *GRIN2A* and *GRIN2B* Polymorphisms. *Life (Basel)* 2021; 11: 997

Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 2011; 43: 977-985

Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* 2019; 47: D886-D894

Salloum IM, Brown ES. Management of comorbid bipolar disorder and substance use disorders. *Am J Drug Alcohol Abuse* 2017; 43: 366-376

Sanchez Marco SB, Buhl E, Firth R, Zhu B, Gainsborough M, Beleza-Meireles A, Moore S, Caswell R, Stals K, Ellard S, Kennedy C, Hodge JLL, Majumdar A. Hereditary spastic paraparesis presenting as cerebral palsy due to *ADD3* variant with mechanistic insight provided by a *Drosophila* γ -adducin model. *Clin Genet* 2022; 102: 494-502

Schmermund A, Möhlenkamp S, Stang A, Grönemeyer D, Seibel R, Hirche H, Mann K, Siffert W, Lauterbach K, Siegrist J, Jöckel KH, Erbel R. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. *Risk Factors, Evaluation of Coronary Calcium and Lifestyle. Am Heart J* 2002; 144: 212-218

Schulz H, Ruppert AK, Herms S, Wolf C, Mirza-Schreiber N, Stegle O, Czamara D, Forstner AJ, Sivalingam S, Schoch S, Moebus S, Pütz B, Hillmer A, Fricker N, Vatter H,

Müller-Myhsok B, Nöthen MM, Becker AJ, Hoffmann P, Sander T, Cichon S. Genome-wide mapping of genetic determinants influencing DNA methylation and gene expression in human hippocampus. *Nat Commun* 2017; 8: 1511

Singh T, Poterba T, Curtis D, Akil H, Al Eissa M, Barchas JD, Bass N, Bigdeli TB, Breen G, Bromet EJ, Buckley PF, Bunney WE, Bybjerg-Grauholm J, Byerley WF, Chapman SB, Chen WJ, Churchhouse C, Craddock N, Cusick CM, DeLisi L, Dodge S, Escamilla MA, Eskelinen S, Fanous AH, Faraone SV, Fiorentino A, Francioli L, Gabriel SB, Gage D, Gagliano Taliun SA, Ganna A, Genovese G, Glahn DC, Grove J, Hall MH, Hämäläinen E, Heyne HO, Holi M, Hougaard DM, Howrigan DP, Huang H, Hwu HG, Kahn RS, Kang HM, Karczewski KJ, Kirov G, Knowles JA, Lee FS, Lehrer DS, Lescai F, Malaspina D, Marder SR, McCarroll SA, McIntosh AM, Medeiros H, Milani L, Morley CP, Morris DW, Mortensen PB, Myers RM, Nordentoft M, O'Brien NL, Olivares AM, Ongur D, Ouwehand WH, Palmer DS, Paunio T, Quedsted D, Rapaport MH, Rees E, Rollins B, Satterstrom FK, Schatzberg A, Scolnick E, Scott LJ, Sharp SI, Sklar P, Smoller JW, Sobell JL, Solomonson M, Stahl EA, Stevens CR, Suvisaari J, Tiao G, Watson SJ, Watts NA, Blackwood DH, Børglum AD, Cohen BM, Corvin AP, Esko T, Freimer NB, Glatt SJ, Hultman CM, McQuillin A, Palotie A, Pato CN, Pato MT, Pulver AE, St Clair D, Tsuang MT, Vawter MP, Walters JT, Werge TM, Ophoff RA, Sullivan PF, Owen MJ, Boehnke M, O'Donovan MC, Neale BM, Daly MJ. Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* 2022; 604: 509-516

Spektrum Akademischer Verlag, 1999: Genetische Heterogenität.
<https://www.spektrum.de/lexikon/biologie/genetische-heterogenitaet/27276>
 (Zugriffsdatum: 23.11.2023)

Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, Mattheisen M, Wang Y, Coleman JRI, Gaspar HA, de Leeuw CA, Steinberg S, Pavlides JMW, Trzaskowski M, Byrne EM, Pers TH, Holmans PA, Richards AL, Abbott L, Agerbo E, Akil H, Albani D, Alliey-Rodriguez N, Als TD, Anjorin A, Antilla V, Awasthi S, Badner JA, Bækvad-Hansen M, Barchas JD, Bass N, Bauer M, Belliveau R, Bergen SE, Pedersen CB, Bøen E, Boks MP, Boocock J, Budde M, Bunney W, Burmeister M, Bybjerg-Grauholm J, Byerley W, Casas M, Cerrato F, Cervantes P, Chambert K, Charney AW, Chen D, Churchhouse C,

Clarke TK, Coryell W, Craig DW, Cruceanu C, Curtis D, Czerski PM, Dale AM, de Jong S, Degenhardt F, Del-Favero J, DePaulo JR, Djurovic S, Dobbyn AL, Dumont A, Elvsåshagen T, Escott-Price V, Fan CC, Fischer SB, Flickinger M, Foroud TM, Forty L, Frank J, Fraser C, Freimer NB, Frisé L, Gade K, Gage D, Garnham J, Giambartolomei C, Pedersen MG, Goldstein J, Gordon SD, Gordon-Smith K, Green EK, Green MJ, Greenwood TA, Grove J, Guan W, Guzman-Parra J, Hamshere ML, Hautzinger M, Heilbronner U, Herms S, Hipolito M, Hoffmann P, Holland D, Huckins L, Jamain S, Johnson JS, Juréus A, Kandaswamy R, Karlsson R, Kennedy JL, Kittel-Schneider S, Knowles JA, Kogevinas M, Koller AC, Kupka R, Lavebratt C, Lawrence J, Lawson WB, Leber M, Lee PH, Levy SE, Li JZ, Liu C, Lucae S, Maaser A, MacIntyre DJ, Mahon PB, Maier W, Martinsson L, McCarroll S, McGuffin P, McInnis MG, McKay JD, Medeiros H, Medland SE, Meng F, Milani L, Montgomery GW, Morris DW, Mühleisen TW, Mullins N, Nguyen H, Nievergelt CM, Adolfsson AN, Nwulia EA, O'Donovan C, Loohuis LMO, Ori APS, Oruc L, Ösby U, Perlis RH, Perry A, Pfennig A, Potash JB, Purcell SM, Regeer EJ, Reif A, Reinbold CS, Rice JP, Rivas F, Rivera M, Roussos P, Ruderfer DM, Ryu E, Sánchez-Mora C, Schatzberg AF, Scheftner WA, Schork NJ, Shannon Weickert C, Shekhtman T, Shilling PD, Sigurdsson E, Slaney C, Smeland OB, Sobell JL, Sørholm Hansen C, Spijker AT, St Clair D, Steffens M, Strauss JS, Streit F, Strohmaier J, Szlinger S, Thompson RC, Thorgeirsson TE, Treutlein J, Vedder H, Wang W, Watson SJ, Weickert TW, Witt SH, Xi S, Xu W, Young AH, Zandi P, Zhang P, Zöllner S; eQTLGen Consortium; BIOS Consortium; Adolfsson R, Agartz I, Alda M, Backlund L, Baune BT, Bellivier F, Berrettini WH, Biernacka JM, Blackwood DHR, Boehnke M, Børghlum AD, Corvin A, Craddock N, Daly MJ, Dannlowski U, Esko T, Etain B, Frye M, Fullerton JM, Gershon ES, Gill M, Goes F, Grigoriu-Serbanescu M, Hauser J, Hougaard DM, Hultman CM, Jones I, Jones LA, Kahn RS, Kirov G, Landén M, Leboyer M, Lewis CM, Li QS, Lissowska J, Martin NG, Mayoral F, McElroy SL, McIntosh AM, McMahon FJ, Melle I, Metspalu A, Mitchell PB, Morken G, Mors O, Mortensen PB, Müller-Myhsok B, Myers RM, Neale BM, Nimgaonkar V, Nordentoft M, Nöthen MM, O'Donovan MC, Oedegaard KJ, Owen MJ, Paciga SA, Pato C, Pato MT, Posthuma D, Ramos-Quiroga JA, Ribasés M, Rietschel M, Rouleau GA, Schalling M, Schofield PR, Schulze TG, Serretti A, Smoller JW, Stefansson H, Stefansson K, Stordal E, Sullivan PF, Turecki G, Vaaler AE, Vieta E, Vincent JB, Werge T, Nurnberger JI, Wray NR, Di Florio A, Edenberg HJ, Cichon S, Ophoff RA, Scott LJ, Andreassen OA,

Kelsoe J, Sklar P. Bipolar Disorder Working Group of the Psychiatric Genomics Consortium. Genome-wide association study identifies 30 loci associated with bipolar disorder. *BioRxiv* 2018, 173062. <https://doi.org/10.1101/173062>

Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetsky V, Mattheisen M, Wang Y, Coleman JRI, Gaspar HA, de Leeuw CA, Steinberg S, Pavlides JMW, Trzaskowski M, Byrne EM, Pers TH, Holmans PA, Richards AL, Abbott L, Agerbo E, Akil H, Albani D, Alliey-Rodriguez N, Als TD, Anjorin A, Antilla V, Awasthi S, Badner JA, Bækvad-Hansen M, Barchas JD, Bass N, Bauer M, Belliveau R, Bergen SE, Pedersen CB, Bøen E, Boks MP, Boocock J, Budde M, Bunney W, Burmeister M, Bybjerg-Grauholm J, Byerley W, Casas M, Cerrato F, Cervantes P, Chambert K, Charney AW, Chen D, Churchhouse C, Clarke TK, Coryell W, Craig DW, Cruceanu C, Curtis D, Czerski PM, Dale AM, de Jong S, Degenhardt F, Del-Favero J, DePaulo JR, Djurovic S, Dobbyn AL, Dumont A, Elvsåshagen T, Escott-Price V, Fan CC, Fischer SB, Flickinger M, Foroud TM, Forty L, Frank J, Fraser C, Freimer NB, Frisén L, Gade K, Gage D, Garnham J, Giambartolomei C, Pedersen MG, Goldstein J, Gordon SD, Gordon-Smith K, Green EK, Green MJ, Greenwood TA, Grove J, Guan W, Guzman-Parra J, Hamshere ML, Hautzinger M, Heilbronner U, Herms S, Hipolito M, Hoffmann P, Holland D, Huckins L, Jamain S, Johnson JS, Juréus A, Kandaswamy R, Karlsson R, Kennedy JL, Kittel-Schneider S, Knowles JA, Kogevinas M, Koller AC, Kupka R, Lavebratt C, Lawrence J, Lawson WB, Leber M, Lee PH, Levy SE, Li JZ, Liu C, Lucae S, Maaser A, MacIntyre DJ, Mahon PB, Maier W, Martinsson L, McCarroll S, McGuffin P, McInnis MG, McKay JD, Medeiros H, Medland SE, Meng F, Milani L, Montgomery GW, Morris DW, Mühleisen TW, Mullins N, Nguyen H, Nievergelt CM, Adolfsson AN, Nwulia EA, O'Donovan C, Loohuis LMO, Ori APS, Oruc L, Ösby U, Perlis RH, Perry A, Pfennig A, Potash JB, Purcell SM, Regeer EJ, Reif A, Reinbold CS, Rice JP, Rivas F, Rivera M, Roussos P, Ruderfer DM, Ryu E, Sánchez-Mora C, Schatzberg AF, Scheftner WA, Schork NJ, Shannon Weickert C, Shekhtman T, Shilling PD, Sigurdsson E, Slaney C, Smeland OB, Sobell JL, Sørholm Hansen C, Spijker AT, St Clair D, Steffens M, Strauss JS, Streit F, Strohmaier J, Szlinger S, Thompson RC, Thorgeirsson TE, Treutlein J, Vedder H, Wang W, Watson SJ, Weickert TW, Witt SH, Xi S, Xu W, Young AH, Zandi P, Zhang P, Zöllner S; eQTLGen Consortium; BIOS Consortium; Adolfsson R, Agartz I, Alda M, Backlund L, Baune BT, Bellivier F,

Berrettini WH, Biernacka JM, Blackwood DHR, Boehnke M, Børglum AD, Corvin A, Craddock N, Daly MJ, Dannlowski U, Esko T, Etain B, Frye M, Fullerton JM, Gershon ES, Gill M, Goes F, Grigoriu-Serbanescu M, Hauser J, Hougaard DM, Hultman CM, Jones I, Jones LA, Kahn RS, Kirov G, Landén M, Leboyer M, Lewis CM, Li QS, Lissowska J, Martin NG, Mayoral F, McElroy SL, McIntosh AM, McMahon FJ, Melle I, Metspalu A, Mitchell PB, Morken G, Mors O, Mortensen PB, Müller-Myhsok B, Myers RM, Neale BM, Nimgaonkar V, Nordentoft M, Nöthen MM, O'Donovan MC, Oedegaard KJ, Owen MJ, Paciga SA, Pato C, Pato MT, Posthuma D, Ramos-Quiroga JA, Ribasés M, Rietschel M, Rouleau GA, Schalling M, Schofield PR, Schulze TG, Serretti A, Smoller JW, Stefansson H, Stefansson K, Stordal E, Sullivan PF, Turecki G, Vaaler AE, Vieta E, Vincent JB, Werge T, Nurnberger JI, Wray NR, Di Florio A, Edenberg HJ, Cichon S, Ophoff RA, Scott LJ, Andreassen OA, Kelsoe J, Sklar P. Bipolar Disorder Working Group of the Psychiatric Genomics Consortium. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* 2019; 51: 793-803

Strehlow V, Heyne HO, Vlaskamp DRM, Marwick KFM, Rudolf G, de Bellecize J, Biskup S, Brilstra EH, Brouwer OF, Callenbach PMC, Hentschel J, Hirsch E, Kind PC, Mignot C, Platzer K, Rump P, Skehel PA, Wyllie DJA, Hardingham GE, van Ravenswaaij-Arts CMA, Lesca G, Lemke JR. GRIN2A study group. GRIN2A-related disorders: genotype and functional consequence predict phenotype. *Brain* 2019; 142: 80-92

Toma C, Shaw AD, Allcock RJN, Heath A, Pierce KD, Mitchell PB, Schofield PR, Fullerton JM. An examination of multiple classes of rare variants in extended families with bipolar disorder. *Transl Psychiatry* 2018; 8: 65

Toma C, Shaw AD, Heath A, Pierce KD, Mitchell PB, Schofield PR, Fullerton JM. A linkage and exome study of multiplex families with bipolar disorder implicates rare coding variants of ANK3 and additional rare alleles at 10q11-q21. *J Psychiatry Neurosci* 2021; 46: E247-E257

Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, Bryois J, Chen CY, Dennison CA, Hall LS, Lam M, Watanabe K, Frei O, Ge T, Harwood JC, Koopmans F, Magnusson S, Richards AL, Sidorenko J, Wu Y, Zeng J, Grove J, Kim M, Li Z, Voloudakis G, Zhang W, Adams M, Agartz I, Atkinson EG, Agerbo E, Al Eissa M, Albus

M, Alexander M, Alizadeh BZ, Alptekin K, Als TD, Amin F, Arolt V, Arrojo M, Athanasiu L, Azevedo MH, Bacanu SA, Bass NJ, Begemann M, Belliveau RA, Bene J, Benyamin B, Bergen SE, Blasi G, Bobes J, Bonassi S, Braun A, Bressan RA, Bromet EJ, Bruggeman R, Buckley PF, Buckner RL, Bybjerg-Grauholm J, Cahn W, Cairns MJ, Calkins ME, Carr VJ, Castle D, Catts SV, Chambert KD, Chan RCK, Chaumette B, Cheng W, Cheung EFC, Chong SA, Cohen D, Consoli A, Cordeiro Q, Costas J, Curtis C, Davidson M, Davis KL, de Haan L, Degenhardt F, DeLisi LE, Demontis D, Dickerson F, Dikeos D, Dinan T, Djurovic S, Duan J, Ducci G, Dudbridge F, Eriksson JG, Fañanás L, Faraone SV, Fiorentino A, Forstner A, Frank J, Freimer NB, Fromer M, Frustaci A, Gadelha A, Genovese G, Gershon ES, Giannitelli M, Giegling I, Giusti-Rodríguez P, Godard S, Goldstein JI, González Peñas J, González-Pinto A, Gopal S, Gratten J, Green MF, Greenwood TA, Guillin O, Gülöksüz S, Gur RE, Gur RC, Gutiérrez B, Hahn E, Hakonarson H, Haroutunian V, Hartmann AM, Harvey C, Hayward C, Henskens FA, Herms S, Hoffmann P, Howrigan DP, Ikeda M, Iyegbe C, Joa I, Julià A, Kähler AK, Kam-Thong T, Kamatani Y, Karachanak-Yankova S, Kebir O, Keller MC, Kelly BJ, Khrunin A, Kim SW, Klovins J, Kondratiev N, Konte B, Kraft J, Kubo M, Kučinskis V, Kučinskiene ZA, Kusumawardhani A, Kuzelova-Ptackova H, Landi S, Lazzeroni LC, Lee PH, Legge SE, Lehrer DS, Lencer R, Lerer B, Li M, Lieberman J, Light GA, Limborska S, Liu CM, Lönnqvist J, Loughland CM, Lubinski J, Luykx JJ, Lynham A, Macek M Jr, Mackinnon A, Magnusson PKE, Maher BS, Maier W, Malaspina D, Mallet J, Marder SR, Marsal S, Martin AR, Martorell L, Mattheisen M, McCarley RW, McDonald C, McGrath JJ, Medeiros H, Meier S, Melegh B, Melle I, Meshulam-Gately RI, Metspalu A, Michie PT, Milani L, Milanova V, Mitjans M, Molden E, Molina E, Molto MD, Mondelli V, Moreno C, Morley CP, Muntané G, Murphy KC, Myin-Germeys I, Nenadić I, Nestadt G, Nikitina-Zake L, Noto C, Nuechterlein KH, O'Brien NL, O'Neill FA, Oh SY, Olincy A, Ota VK, Pantelis C, Papadimitriou GN, Parellada M, Paunio T, Pellegrino R, Periyasamy S, Perkins DO, Pfulmann B, Pietiläinen O, Pimm J, Porteous D, Powell J, Quattrone D, Quested D, Radant AD, Rampino A, Rapaport MH, Rautanen A, Reichenberg A, Roe C, Roffman JL, Roth J, Rothermundt M, Rutten BPF, Saker-Delye S, Salomaa V, Sanjuan J, Santoro ML, Savitz A, Schall U, Scott RJ, Seidman LJ, Sharp SI, Shi J, Siever LJ, Sigurdsson E, Sim K, Skarabis N, Slominsky P, So HC, Sobell JL, Söderman E, Stain HJ, Steen NE, Steixner-Kumar AA, Stögmann E, Stone WS, Straub RE, Streit F, Strengman E, Stroup TS,

Subramaniam M, Sugar CA, Suvisaari J, Svrakic DM, Swerdlow NR, Szatkiewicz JP, Ta TMT, Takahashi A, Terao C, Thibaut F, Toncheva D, Tooney PA, Torretta S, Tosato S, Tura GB, Turetsky BI, Üçok A, Vaaler A, van Amelsvoort T, van Winkel R, Veijola J, Waddington J, Walter H, Waterreus A, Webb BT, Weiser M, Williams NM, Witt SH, Wormley BK, Wu JQ, Xu Z, Yolken R, Zai CC, Zhou W, Zhu F, Zimprich F, Atbaşoğlu EC, Ayub M, Benner C, Bertolino A, Black DW, Bray NJ, Breen G, Buccola NG, Byerley WF, Chen WJ, Cloninger CR, Crespo-Facorro B, Donohoe G, Freedman R, Galletly C, Gandal MJ, Gennarelli M, Hougaard DM, Hwu HG, Jablensky AV, McCarroll SA, Moran JL, Mors O, Mortensen PB, Müller-Myhsok B, Neil AL, Nordentoft M, Pato MT, Petryshen TL, Pirinen M, Pulver AE, Schulze TG, Silverman JM, Smoller JW, Stahl EA, Tsuang DW, Vilella E, Wang SH, Xu S; Indonesia Schizophrenia Consortium; PsychENCODE; Psychosis Endophenotypes International Consortium; SynGO Consortium; Adolfsson R, Arango C, Baune BT, Belangero SI, Børglum AD, Braff D, Bramon E, Buxbaum JD, Champion D, Cervilla JA, Cichon S, Collier DA, Corvin A, Curtis D, Forti MD, Domenici E, Ehrenreich H, Escott-Price V, Esko T, Fanous AH, Gareeva A, Gawlik M, Gejman PV, Gill M, Glatt SJ, Golimbet V, Hong KS, Hultman CM, Hyman SE, Iwata N, Jönsson EG, Kahn RS, Kennedy JL, Khusnutdinova E, Kirov G, Knowles JA, Krebs MO, Laurent-Levinson C, Lee J, Lencz T, Levinson DF, Li QS, Liu J, Malhotra AK, Malhotra D, McIntosh A, McQuillin A, Menezes PR, Morgan VA, Morris DW, Mowry BJ, Murray RM, Nimgaonkar V, Nöthen MM, Ophoff RA, Paciga SA, Palotie A, Pato CN, Qin S, Rietschel M, Riley BP, Rivera M, Rujescu D, Saka MC, Sanders AR, Schwab SG, Serretti A, Sham PC, Shi Y, St Clair D, Stefánsson H, Stefansson K, Tsuang MT, van Os J, Vawter MP, Weinberger DR, Werge T, Wildenauer DB, Yu X, Yue W, Holmans PA, Pocklington AJ, Roussos P, Vassos E, Verhage M, Visscher PM, Yang J, Posthuma D, Andreassen OA, Kendler KS, Owen MJ, Wray NR, Daly MJ, Huang H, Neale BM, Sullivan PF, Ripke S, Walters JTR, O'Donovan MC. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* 2022; 604: 502-508

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S,

DePristo MA. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013; 43: 11.10.1-11.10.33

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010; 38: e164

Ward AJ, Cooper TA. The pathobiology of splicing. *J Pathol* 2010; 220: 152-163

Westra HJ, Franke L. From genome to function by studying eQTLs. *Biochim Biophys Acta* 2014; 1842: 1896-1902

World Medical Association, 2023: Declaration of Helsinki 1983. <https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki/doh-oct1983/> (Zugriffsdatum: 23.11.2023)

Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet* 2011; 89: 82-93

Yapici Eser H, Kacar AS, Kilciksiz CM, Yalçınay-Inan M, Ongur D. Prevalence and Associated Features of Anxiety Disorder Comorbidity in Bipolar Disorder: A Meta-Analysis and Meta-Regression Study. *Front Psychiatry* 2018; 9: 229

Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 2014; 30: 614-620

10. Danksagung

Ich möchte mich an dieser Stelle ganz herzlich bei meinem Doktorvater Herrn Prof. Dr. med. Markus M. Nöthen sowie bei der ganzen Abteilung für Humangenetik für die gute Betreuung und die aufregende Zeit bedanken.

Außerdem gilt mein Dank meinen lieben Eltern, Geschwistern, Freunden und meinem Ehemann für die stetige Unterstützung während der Fertigstellung dieser Arbeit.