

Aus dem Zentralinstitut für Seelische Gesundheit Mannheim
der Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität zu Heidelberg

Direktor: Prof. Dr. med. Andreas Meyer-Lindenberg

Nutzung genomweiter Daten zum besseren Verständnis der molekulargenetischen Basis von Schizophrenie und Alkoholabhängigkeit

Inauguraldissertation

zur Erlangung des

Doctor scientiarum humanarum (Dr. sc. hum.)

der

Medizinischen Fakultät Mannheim

der Ruprecht-Karls-Universität

zu

Heidelberg

vorgelegt von

Josef Frank

aus

Riedenburg

2018

Dekan: Herr Prof. Dr. med. Sergij Goerd

Betreuerin: Frau Prof. Dr. med. Marcella Rietschel

2	EINLEITUNG	3
2.1	Suche nach genetischen Faktoren bei psychiatrischen Erkrankungen durch genomweite Assoziationsuntersuchungen.	3
2.1.1	Ausgangslage	3
2.1.2	Statistische Power von GWAS	6
2.1.3	GWAS Befunde	8
2.2	Ziele der Arbeit	10
2.2.1	Schizophrenie	10
2.2.2	Alkoholabhängigkeit	11
2.3	Beitrag des Doktoranden	14
3	MATERIAL UND METHODEN.....	16
3.1	Stichproben.....	16
3.1.1	Schizophrenie-Studie	16
3.1.2	Studien zur Alkoholabhängigkeit	17
3.2	Probengewinnung und Genetische Analyse	18
3.3	Benutzte Programme/Softwarepakete	19
3.4	Statistische Verfahren.....	19
3.4.1	Qualitätskontrolle (QC) der Genetischen Daten	19
3.4.2	Testverfahren	20
4	ERGEBNISSE	25
4.1	Schizophrenie-Studie.....	25
4.1.1	Ergebnisse der Publikation: Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients	25
4.2	Studien zur Alkoholabhängigkeit.....	28
4.2.1	Ergebnisse der Publikation: Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster	
	28	
4.2.2	Ergebnisse der Publikation: Genetic contribution to alcohol dependence: investigation of a heterogeneous German sample of	

individuals with alcohol dependence, chronic alcoholic pancreatitis, and alcohol-related cirrhosis	32
4.2.3 Ergebnisse der Publikation: Shared genetic etiology between alcohol dependence and major depressive disorder	34
5 DISKUSSION	37
5.1 Schizophrenie	37
5.1.1 Publikation: Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients	37
5.2 Alkoholabhängigkeit.....	38
5.2.1 Publikation: Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster	38
5.2.2 Publikation: Genetic contribution to alcohol dependence: investigation of a heterogeneous German sample of individuals with alcohol dependence, chronic alcoholic pancreatitis, and alcohol- related cirrhosis.....	40
5.2.3 Publikation: Shared genetic etiology between alcohol dependence and major depressive disorder	42
6 ZUSAMMENFASSUNG.....	46
7 LITERATURVERZEICHNIS.....	48
8 ENTHALTENE EIGENE VERÖFFENTLICHUNGEN	62
9 LEBENSLAUF	63
10 DANKSAGUNG	64

ABKÜRZUNGSVERZEICHNIS

- ACP = alkoholinduzierte chronische Pankreatitis
- ALZ = alkoholinduzierte Leberzirrhose
- GWAS = Genomweite Assoziationsstudie
- PCA = Principal Component Analysis
(= Hauptkomponentenanalyse)
- SNP = Single Nucleotide Polymorphism
(= Einzelbasenaustausch)
- TRS = treatment resistant schizophrenia;
Behandlungsresistente Schizophrenie:
Erfordernis der Behandlung mit Clozapin
- ETRS = extreme treatment resistant schizophrenia
extrem behandlungsresistente Schizophrenie:
Nicht-Ansprechen auf Clozapin
- ETRS+ = ETRS und zusätzlich:
früher und schleichender Krankheitsbeginn sowie
schlechte prämorbid soziale Anpassung
- SCZ = Schizophrenie

2 EINLEITUNG

2.1 Suche nach genetischen Faktoren bei psychiatrischen Erkrankungen durch genomweite Assoziationsuntersuchungen.

2.1.1 Ausgangslage

Psychiatrische Störungen sind häufig. Sie beginnen in der Regel im frühen Erwachsenenalter und zeigen oftmals einen rezidivierenden Verlauf. Sie gehören weltweit zu den Hauptverursachern der Globalen Krankheitslast (Vigo, Thornicroft, & Atun, 2016). Das Wissen um die genauen Ursachen ist allerdings noch unzureichend und somit auch die Prävention und Therapie der Störungen.

Wie alle Volkskrankheiten sind psychiatrische Störungen komplex bedingt, d.h. dass sowohl genetische als auch Umweltfaktoren an der Krankheitsentstehung beteiligt sind. Im Gegensatz zu anderen Volkskrankheiten, wie Krebs, Diabetes oder Bluthochdruck, gibt es jedoch bislang keine objektiven biologischen Marker, welche die Diagnosestellung oder die Vorhersage des Verlaufes unterstützen könnten (Sullivan, Daly, & O'Donovan, 2012). So stützt sich die Diagnose bislang ausschließlich auf klinisch beobachtbare Symptome. Eine Diagnose wird dann gestellt, wenn eine definierte Mindestanzahl von Symptomen – aus einem Katalog von spezifischen Symptomen – über einen definierten Zeitraum aufgetreten sind. Daraus ergibt sich, dass Patienten mit den gleichen Diagnosen unterschiedliche Symptome und Verlaufsformen aufweisen können.

Darüber hinaus besteht eine hohe Komorbidität zwischen den psychischen Störungen untereinander. So zeigen insbesondere affektive Störungen eine hohe Komorbidität mit schizophrenen Störungen (Buckley, Miller, Lehrer, & Castle, 2009; Laursen, Agerbo, & Pedersen, 2009) und Alkoholabhängigkeit (Lynskey, 1998). Es ist davon auszugehen, dass diese klinische Heterogenität auch mit einer genetischen Heterogenität einhergeht. Andererseits zeigen Familienstudien, wie z.B. bei Schizophrenie, dass Angehörige ein erhöhtes Risiko für Spektrumsstörungen haben, also, dass eine erhöhte genetische Belastung für Schizophrenie mit einer Vielzahl von klinischen Symptombildern einhergehen kann, die noch nicht einmal die Diagnose einer Schizophrenie erfüllen müssen.

Formalgenetische Untersuchungen belegen weiterhin den substantiellen Beitrag genetischer Faktoren bei der Entstehung dieser Störungen. Bei der Schizophrenie, den bipolaren Störungen oder den Autismus Spektrum Störungen ist die Erbllichkeit, also der Anteil, der durch genetische Faktoren erklärt wird, sehr hoch und liegt bis über 80% (Cardno et al., 1999; Sullivan et al., 2012; Hilker et al., 2017; McGuffin et al., 2003). Bei den sehr häufig in der Population vorkommenden Störungen wie Alkoholabhängigkeit liegt sie niedriger aber immerhin noch bei ca. 40-50% (Sullivan et al., 2012)

Deshalb erwartet man von der Identifizierung der genetischen Faktoren auf der molekularen Ebene, ganz neue Einblicke in die biologischen Krankheitsprozesse zu gewinnen. Hierbei wird untersucht, ob sich Patienten hinsichtlich des Vorkommens von genetischen Varianten von nicht betroffenen Personen unterscheiden.

Die genetischen Varianten betreffen Veränderungen in der Desoxyribonucleinsäure (DNA) und zwar in Abfolge der insgesamt $3,2 \times 10^9$ Basenpaare (A, T, C, G; zur Nomenklatur siehe <http://varnomen.hgvs.org/recommendations/DNA/>). Die größte Quelle genetischer Variabilität beruht auf Austausch einzelner solcher Basen, sogenannter Single Nucleotide Polymorphisms (SNPs, Abbildung 1) auf mindestens einem der beiden Chromosomen gegenüber dem Referenzgenom

```
ACTGAATATGGGCATCGCAGATACCCGATTTAGGAAACTGTTTTTGGGCTCCCCGGGAATATCGCGCGTT
ACTGAATATGGGCATTGAAGATACCCGATTTAGCAAAC T GTTTTTGGGCTA CCCCCGGGAATATCGCGCGTT
ACTGAATATGGGCATCGAAGATACCCGATTTAGCAAAC T GTTTTTGGGCT CCCCCGGGAATATCGCGCGTT
```

Abbildung 1: Darstellung von Einzelbasenaustauschen in einer genomischen Sequenz. Oben in blau der Referenzstrang, unten in schwarzer Schrift die Sequenzen für beide Chromosomen eines Individuums an demselben Genlocus. Es können entweder beide Allele ersetzt sein (homozygot), oder auch nur eines davon (heterozygot).

Schätzungsweise enthält das menschliche Genom 11 Millionen solcher SNPs, hiervon weisen ca. 7 Millionen eine Frequenz von über 5% und der Rest eine Frequenz von 1-5% in der europäischen Normalbevölkerung auf (Frazer, Murray, Schork, & Topol, 2009).

Bislang ist noch ungeklärt, in wieweit bei einzelnen Störungen und bei einzelnen Patienten spezifische seltene Varianten mit größeren oder großen Effekten oder

aber Kombinationen mehrerer häufiger/seltener genetischer Varianten eine Rolle spielen, und ob eine Vielzahl von ganz unterschiedlichen Varianten zu dem gleichen klinischen Erscheinungsbild führen kann.

Anfangs wurde angenommen, dass – ähnlich wie bei monogenen Erbkrankheiten – seltene Varianten mit großen Effekten ursächlich für die psychiatrischen Störungen seien. Deshalb konzentrierten sich die genetischen Studien auf Familien mit mehrfach Betroffenen. Hier sollten mittels Kopplungsuntersuchungen Genregionen gefunden werden, in denen sich diese seltenen Varianten befinden. Entgegen den Erwartungen konnten hierbei allerdings keine eindeutigen Loci mit starken Risikoeffekten identifiziert werden. Eine Ausnahme bildete bei Alkoholabhängigkeit eine Region auf dem langen Arm des Chromosom 4 in der Gene lokalisiert sind, die für Isoformen der Alkoholdehydrogenase kodieren, die mit unterschiedlichen Geschwindigkeit Alkohol verstoffwechseln (Tawa, Hall, & Lohoff, 2016). Auch wenn sich somit keine neuen Erkenntnisse auftaten, konnte zumindest gezeigt werden, dass mit dieser Methode prinzipiell auch bei komplexen Erkrankungen Risikogene gefunden werden können – falls diese ausreichend große Effekte aufweisen.

Theoretische Überlegungen zeigen, dass es unwahrscheinlich ist, dass häufige Erkrankungen, die mit Selektionsnachteilen einhergehen, wie dies durch die empirisch belegten verminderten Fertilitätsraten bei schizophrenen Patienten nahe liegt, durch hochfrequente Risikovarianten mit großen Effekten bedingt sind. Insbesondere wenn Verwandte erhöhte Fertilitätsraten zeigen, wie dies bei den meisten psychiatrischen Störungen der Fall ist, weist dies auf die Beteiligung von häufigen Risikovarianten mit kleineren positiven Effekten hin (Power et al., 2013).

Demgemäß verlagerten sich die Anstrengungen auf Untersuchungen von häufig vorkommenden Varianten mit kleineren Effekten bei unverwandten Patienten und Kontrollen. Diese Untersuchungen konzentrierten sich in der Vergangenheit, als es noch nicht möglich war, das gesamte Genom zu untersuchen, auf sogenannte Kandidatengene, das heißt Gene von denen man auf Grund pharmakologischer oder anderer Beobachtungen annahm, dass sie bei der Erkrankung eine Rolle spielen. Da beim Kandidatengenansatz allerdings auf bereits vorliegende Hypothesen zurückgegriffen wird, ist der Erkenntnisgewinn über die weiteren molekulargenetischen Ursachen der Störungen gering. Zudem waren die meisten der Ergebnisse inkonsistent

(Sullivan et al., 2012), wobei die Befunde bei Alkoholabhängigkeit mit Varianten in den Alkohol metabolisierenden Genen Stoffwechselgenen *ADH* und *ALDH* eine Ausnahme bilden (Tawa et al., 2016).

Durch sinkende Genotypisierungskosten wurde es ab 2005 faktisch möglich, in zunehmend größerer Zahl genomweite Assoziationsstudien (GWAS) durchzuführen (Hirschhorn & Daly, 2005). Dies bedeutet, dass zeitnah genomweit eine Vielzahl von SNPs, heute in der Regel ein bis zwei Millionen, bei einer großen Zahl von Personen untersucht werden, um so Unterschiede in der Häufung von genetischen Varianten zwischen Gruppen festzustellen (Burton et al., 2007). Mit Hilfe dieses Untersuchungsansatzes wurden in den letzten zehn Jahren bei einer Vielzahl von komplexen Erkrankungen eine zunehmende Anzahl an Varianten, Genen, Pathways und bislang unbekannte biologische Prozesse identifiziert (Abbildung 2)

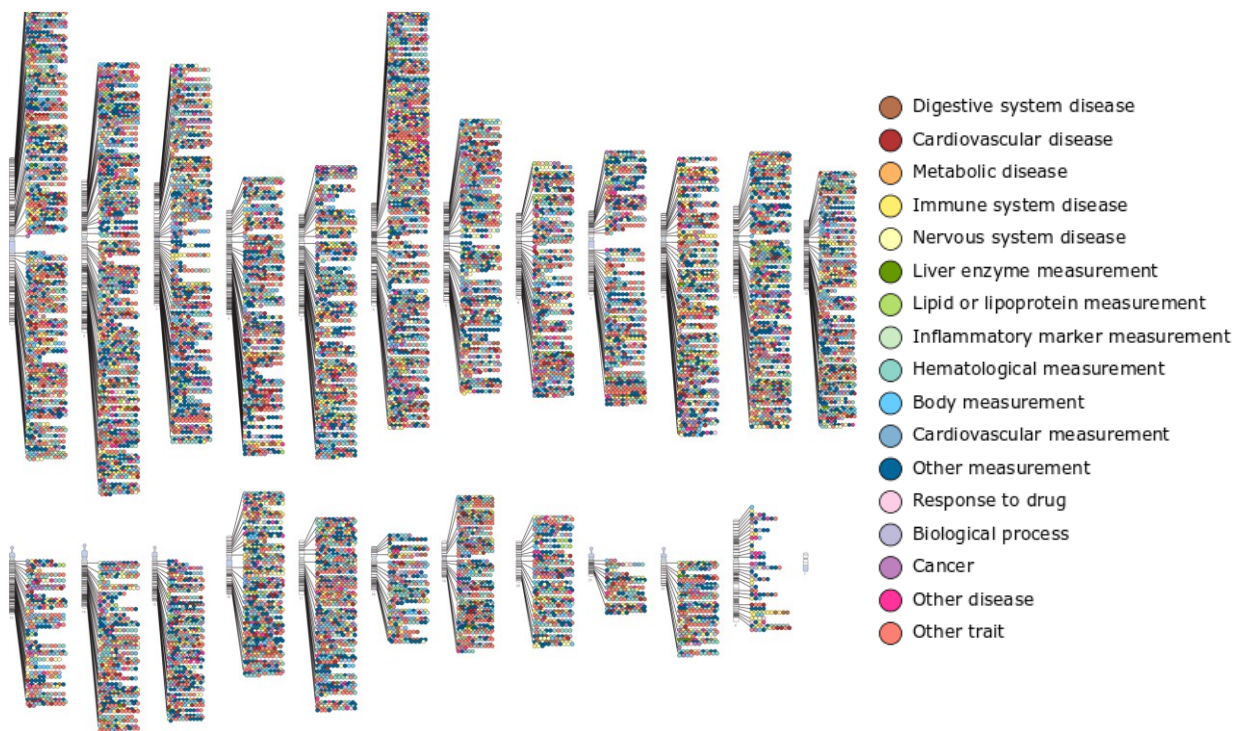


Abbildung 2: Überblick über bereits publizierte genomweit signifikante Assoziationsbefunde (abgerufen von <https://www.ebi.ac.uk/gwas/diagram> am 15.11.2018)

2.1.2 Statistische Power von GWAS

Ein Nachteil dieses systematischen genomweiten Screenings ist es, dass für diese große Anzahl von durchgeführten Assoziationstests statistisch korrigiert wer-

den muss (Yang et al., 2010). Dies führt dazu, dass die meisten echten Assoziationsbefunde nicht als solche erkannt werden, weil sie die Signifikanzgrenze von 5×10^{-8} (Pe'er, Yelensky, Altshuler, & Daly, 2008) nicht erreichen (Zondervan & Cardon, 2007). Durch Vergrößerung der untersuchten Stichprobe kann eine erhöhte statistische Aussagekraft erreicht werden (Spencer, Su, Donnelly, & Marchini, 2009). Hierbei wird geschätzt, dass bei einer bei einer Störung wie Schizophrenie, die eine Erblichkeit von ca. 80% und eine Prävalenz von ca. 1% aufweist, eine Stichprobengröße von etwa 50 000 Patienten und ebenso vielen Kontrollen benötigt wird, um einen Effekt von etwa 10% Aufklärung der phänotypischen Varianz durch die typisierten SNPs mit einer ausreichenden Wahrscheinlichkeit zu entdecken (Wray, Purcell, & Visscher, 2011), wie es ähnlich für Körpergröße berichtet wurde (Lango Allen et al., 2010). Solche großen Stichprobenzahlen können aber nur durch die Zusammenarbeit vieler Zentren erreicht werden.

Während bei GWAS bei Tumorerkrankungen objektive Parameter zur Phänotypisierung herangezogen werden können, beschränkte sich der Krankheitsphänotyp bei psychiatrischen GWAS bislang üblicherweise auf die Diagnosen. Dies führt wie oben aufgeführt dazu, dass Patienten mit unterschiedlicher Symptomatik und Krankheitsverläufen in einer Patientengruppe zusammengefasst werden. Prinzipiell wäre es natürlich sinnvoll, große Patientenstichproben mit vergleichbarer klinischer Symptomatik zu untersuchen. Dies wäre allerdings mit einer verminderten Stichprobengröße verbunden und würde vor allem eine Zentren-übergreifende systematische Erfassung der Symptome und Verläufe notwendig machen. Angesichts des dafür erforderlichen großen Aufwandes wurde in der Praxis bislang darauf weitgehend verzichtet. Kritiker argumentieren, dass der Zuwachs an der Samplegröße mit einer erhöhten klinischen und genetischen Heterogenität erkauft wird. Das führt dazu, dass die statistische Power mit Anwachsen der Samplegröße nicht automatisch zunimmt, und im ungünstigsten Falle sogar abnimmt (Bigdeli et al., 2017). Die Befürworter des Ansatzes argumentieren hingegen, dass sich ein solcher eventueller Verlust der Power durch das Anwachsen der Samplegröße ausgeglichen werden könne (Sullivan, 2012)

2.1.3 GWAS Befunde

Durch solche Zentren-übergreifenden Studien gelang 2008 erstmals die Identifikation zweier signifikanter Befunde durch diesen genomweiten Ansatz und zwar für eine Variante im Zinc finger protein 804A (*ZNF804A*) Gen in einem Sample von 479 schizophrenen Patienten und 2 937 Kontrollen für Schizophrenie (O'Donovan et al., 2008) und bei 1 233 bipolaren Patienten und 1 439 Kontrollen im Diacylglycerol Kinase Eta (*DGKH*) Gen für bipolare Störungen (Baum et al., 2008), 2009 folgte der erste genomweite Befund für Alkoholabhängigkeit bei 487 Patienten und 1358 Kontrollen mit einer intergenischen Variante in nächster Nähe zum Peroxisomal trans-2-enoyl-CoA reductase (*PECR*) Gen. Diese, durch unsere Arbeitsgruppe durchgeführte Studie unterschied sich von den üblichen GWAS-Analysen dadurch, dass wir uns auf klinisch und genetisch homogenere Sample konzentrierten, indem wir für Ethnizität, Geschlecht, Krankheitsschwere und Krankheitsbeginn stratifizierten. Eine weitere Untersuchung durch unsere Arbeitsgruppe – die Bestandteil der hier vorliegenden Dissertation ist – an einem vergrößerten Sample von 1333 Patienten und 2168 Kontrollen, zeigte einen weiteren genomweit signifikanten Befund mit einem Marker im *ADH*-Gencluster (Frank et al., 2012). Nachfolgende GWAS unterstützen diesen Befund mit nominell als auch genomweit signifikanten Assoziationsbefunden mit Varianten in den *ADH*- und *ALDH*-Clustern (Park et al., 2013; Gelernter et al., 2014).

In den nachfolgenden Jahren wurden durch kontinuierliche Vergrößerung der Stichproben weitere genomweit signifikante Varianten bei psychiatrischen Störungen identifiziert. Diese Befunde bildeten die Grundlage für eine Vielzahl von Studien, die die differentiellen Effekte dieser Varianten – insbesondere auf Hirnfunktion und Hirnstruktur – ermöglichten. Allerdings wuchs, im Gegensatz zu anderen komplexen Erkrankungen, wie z.B. Makuladegeneration (Black & Clark, 2016) oder Cholelithiasis (The NIDDK IBD Genetics Consortium et al., 2010), die Anzahl der genomweit signifikanten Varianten bei psychiatrischen Störungen nur relativ langsam und lag beispielsweise im Jahre 2012 bei insgesamt 15 Varianten für Schizophrenie und Bipolare Störung und bei 2 für Alkoholabhängigkeit. Eine zunehmende Zahl von Kritikern gelangten zu der Meinung, dass psychiatrische Störungen zu komplex für den GWAS Ansatz wären (Goldstein, 2009), während die Befürworter darauf verwie-

sen, dass die hierfür notwendigen Samplegrößen noch nicht erreicht waren (Sullivan et al., 2012).

Der Durchbruch gelang 2014, als eine insgesamt 36 989 Patienten und 113 075 Kontrollprobanden umfassende multizentrische Studie – an der unsere Arbeitsgruppe beteiligt war – 108 unabhängige Loci für Schizophrenie identifizieren konnte (Ripke et al., 2014). 2017 lag diese Zahl bereits bei 145 Loci (Pardiñas et al., 2018). Bei bipolaren Störungen liegt die berichtete Anzahl genomweit signifikanter Loci zwischenzeitlich bei 30 (Stahl et al., 2017), und auch bei den hochheritablen Autismusspektrumsstörungen und dem Aufmerksamkeitsdefizit-Hyperaktivitätssyndrom (Demontis et al., 2018) werden eine zunehmende Anzahl an genomweit signifikanten Varianten identifiziert.

Bei Depression, die wie oben erwähnt, eine weitaus geringere Heritabilität aufweist, gelang die Identifikation von 44 genomweit signifikanten Varianten erst nach einer substantiellen Vergrößerung der Stichprobe auf über 100 000 Betroffene (Wray et al., 2018). Für die in einem ähnlich niedrigen Heritabilitätsbereich liegende Alkoholabhängigkeit wurden solche Befunde noch nicht erzielt. Neben den Loci im Bereich der beiden Stoffwechsel-Gene *ADH* und *ADLH* wurden bislang nur drei weitere Regionen berichtet (Gelernter et al., 2014; Treutlein et al., 2009; Wang et al., 2012), die allerdings in der Metaanalyse des Substance Abuse PGC Konsortium (Walters et al., 2018) nicht mehr genomweite Signifikanz erreichen. Es steht abzuwarten, ob die Zahl der Loci mit zunehmender Samplegröße im gleichen Ausmaß wie bei der Depression ansteigen wird. Diese Frage kann zum jetzigen Zeitpunkt noch nicht beantwortet werden.

Die genomweit signifikanten Befunde stellen allerdings nur einen Teil der Information dar, die in den Daten der GWAS enthalten sind. Durch neu entwickelte statistische Auswertungsverfahren kann zunehmend auch die Information der Befunde genutzt werden, die unter der genomweiten Signifikanzschwelle liegen. Hierzu zählen Gen- und Genset-basierte Auswertemethoden sowie die Anwendung polygener Risikoscores auf die im Methodenteil eingegangen wird.

2.2 Ziele der Arbeit

2.2.1 Schizophrenie

2.2.1.1 Spezifischer Hintergrund zur Fragestellung zur Schizophrenie

Schizophrenie ist eine schwere psychiatrische Störung, die in Deutschland eine Lebenszeitprävalenz von ca. 0,7%-1,4% aufweist (Gaebel & Wölwer, 2010), und zu den Hauptverursachern des Global Burden of Disease zählt (Hay, 2017). Die Krankheit, die mit tiefgreifenden Störungen des Denkens, Fühlens und Handelns einhergeht, zeigt individuell Verlaufsformen. Im günstigsten Fall erfahren manche Patienten nur eine einzelne Episode mit vollständiger Remission, im ungünstigsten Fall hingegen findet eine kontinuierliche Verschlechterung statt.

Bislang ist es allerdings weder auf der Basis klinischer noch biologischer Parameter möglich, individuelle Vorhersagen über den Verlauf der Erkrankung zu machen. Allerdings ist bekannt, dass eine schlechte prämorbid soziale Anpassung (Schennach, Riedel, Musil, & Möller, 2012) und ein früher Krankheitsbeginn die Erfolgsaussichten der Behandlung verschlechtern, wohingegen eine frühzeitige Behandlung diese verbessern (Marshall et al., 2005). Zudem deutet der Zusammenhang von familiärer Vorbelastung mit dem Behandlungserfolg (Crespo-Facorro et al., 2013) auf den Einfluss genetischer Faktoren hin..

Ca. 30% aller Patienten weisen eine Therapieresistenz auf. Diese ist dadurch definiert, dass eine medikamentöse Therapie trotz ausreichender Behandlungsdauer mit verschiedenen Substanzen in ausreichend hohen Dosen zu unzureichenden oder zu gar keiner Besserung der Symptome führt. In diesem Fall ist eine Behandlung mit Clozapin angezeigt, welches bislang das einzige Medikament ist, für welches eine Wirksamkeit bei dieser schweren Verlaufsform belegt ist (Essali, Al-Haj Haasan, Li, & Rathbone, 2009).

Allerdings können bei Behandlung mit Clozapin ernste, z.T. lebensbedrohliche Nebenwirkungen wie die sog. Agranulozytose auftreten, die zu einem Zusammenbruch des Immunsystems führt. Clozapin wird daher nur dann verordnet, wenn zuvor mindestens zwei Behandlungsversuche mit anderen Medikamenten gescheitert sind (Nielsen, Nielsen, & Correll, 2012). Dieser Prozess kann sich jedoch über Jahre hin-

ziehen (Taylor, Young, & Paton, 2003). Gleichzeitig sind aber ein längeres Fortbestehen der Krankheit ohne adäquate Behandlung sowie fehlende Wirksamkeit der Erstbehandlung mit schlechteren Erfolgsaussichten verbunden (Schennach et al., 2012). Es erscheint daher angezeigt, frühzeitig die Patienten zu identifizieren, die nicht auf die klassischen antipsychotischen Medikamente ansprechen, und bei denen nach einem längeren erfolglosen Therapieverlauf schließlich doch eine Behandlung mit Clozapin erforderlich sein wird, um durch einen frühzeitigen Beginn die Erfolgsaussichten verbessern zu können.

2.2.1.2 Fragestellung und Ziele der Untersuchung bei Schizophrenie

Publikation: Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients

Frage 1: Ist eine erhöhte genetische Belastung für Schizophrenie mit dem Therapieansprechen bzw. dem Krankheitsverlauf schizophrener Patienten assoziiert?

Für die Beantwortung dieser Frage, wird untersucht, ob der Krankheitsverlauf und das Therapieansprechen von schizophrenen Patienten eine Assoziation mit einem erhöhten polygenen Risikoscore aufweist. Zudem wird der Beitrag bekannter klinischer Faktoren in Kombination mit genetischen Faktoren untersucht.

2.2.2 Alkoholabhängigkeit

2.2.2.1 Spezifischer Hintergrund für die Fragestellung zur Alkoholabhängigkeit

Alkoholabhängigkeit ist eine weltweit sehr verbreitete Störung, die in Deutschland eine Prävalenz von ca. 4% (Meyer, Rumpf, Hapke, Dilling, & John, 2000) bis 6% (Beesdo-Baum et al., 2015) aufweist. Die Störung, die mit einer Reihe von Komorbiditäten, wie z.B. Depression und Angststörungen (Kessler et al., 1996; Regier, 1990) als auch mit alkoholinduzierten Folgestörungen wie Leberzirrhose, Pankreatitis etc. (Schoepf & Heun, 2015) einhergeht, zählt zu den zehn führenden Ursachen des sog. Global Burden of Disease (Connor, Haber, & Hall, 2016; World Health Organization, 2009).

Trotz ausgiebiger Forschung ist immer noch unklar, warum ein Teil der Personen, die Alkohol konsumieren, eine Abhängigkeit entwickeln, der Rest aber nicht. Zwillingsstudien belegen eine genetische Komponente von ca. 40-60% (Enoch & Goldman, 2002).

Epidemiologische Untersuchungen belegen darüber hinaus die Komorbidität zwischen Alkoholabhängigkeit und Depression und zeigen, dass das Vorliegen einer Alkoholabhängigkeit das Risiko für das Auftreten einer Depression erhöht und umgekehrt (Crum et al., 2008; Foulds, Adamson, Boden, Williman, & Mulder, 2015; Kessler et al., 1996; Swendsen & Merikangas, 2000; Woodruff, Guze, Clayton, & Carr, 1973). Es wurde vermutet, dass diese Komorbidität eine gemeinsame Ursache haben könnte die teilweise auf gemeinsame genetische Faktoren zurückzuführen sein könnte (Kendler et al., 1995; Lyons et al., 2006; Maier, Lichtermann, & Mingos, 1994; Nurnberger, Foroud, Flury, Meyer, & Wiegand, 2002; Prescott, Aggen, & Kendler, 2000; Winokur & Coryell, 1991). In einer kürzlich erschienenen Studie konnten Andersen et al (2017) zeigen, dass ein polygener Risikoscore für Depression, der auf den MDD Working Group of the Psychiatric Genomics Consortium (PGC) (PGC-MDD1: $n_{\text{Patienten}}=9\ 240$; $n_{\text{Kontrollen}}=9\ 519$) GWAS (Ripke et al., 2013) beruhte, mit einem erhöhten Risiko für Alkoholabhängigkeit in vier unabhängigen GWAS Datensätzen (mit 317 bis 2 135 Patienten) assoziiert war.

Die Aufklärung dieser genetischen Faktoren könnte zu einem verbesserten Verständnis der Ätiologie und hierdurch zur Entwicklung erfolgversprechender Ansätze zu Prävention und Behandlung beitragen.

Wie in der Einleitung aufgeführt sind die einzigen Gene, die durch systematische Genomuntersuchungen repliziert nachgewiesen werden konnten solche, die bei der Verstoffwechslung von Alkohol eine Rolle spielen. Diese Befunde sind allerdings konsistent und weisen in manchen Populationen sehr starke Effekte auf. Während also – anders als bei den anderen psychiatrischen Erkrankungen – bereits einige Varianten mit starken Effekten identifiziert werden konnten, ist die Gesamtzahl genomweit signifikanter Varianten gering. Dies wird sich wahrscheinlich mit wachsender statistischer Power durch Vergrößerung der Samples ändern. Andererseits ist bislang nur wenig über die genetische Architektur der Alkoholabhängigkeit bekannt.

2.2.2.2 Fragestellung und Ziele der Untersuchungen bei Alkoholabhängigkeit

2.2.2.2.1 Publikation: Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster

Frage 2: Handelt es sich bei der Alkoholabhängigkeit ebenfalls um eine komplexe genetische Störung mit einem polygenen Anteil und tragen Genvarianten, die häufig in der Bevölkerung vorkommen, zur Entstehung von Alkoholabhängigkeit bei?

Durchführung einer Alkohol-GWAS bei 1 333 Patienten und 2 168 Kontrollen und Suche nach signifikant mit Alkohol assoziierten Genvarianten, Genen und Pathways. Basierend auf den Individual-Genotypdaten wird in unabhängigen Fall-Kontrollstichproben überprüft, ob Alkoholpatienten mehr Risikovarianten als Kontrollen tragen, d.h. einen höheren polygenen Risikoscore aufweisen.

2.2.2.2.2 Publikation: Genetic contribution to alcohol dependence: Investigation of a heterogeneous German sample of individuals with alcohol dependence, chronic alcoholic pancreatitis, and alcohol-related cirrhosis

Frage 3: *Ist eine Vergrößerung der Stichprobe auch auf Kosten der klinischen Homogenität zielführend, um weitere Gene und Pathways zu finden?*

Die oben genannte Stichprobe von 1 333 Patienten und 2 168 Kontrollen wird um weitere 1 510 Patienten und 1 750 Kontrollpersonen erweitert, wobei es sich bei den alkoholabhängigen Patienten um solche mit komorbiden Störungen, d.h. mit Leberzirrhose und chronischer Pankreatitis handelt. An Hand des gesicherten ADH Assoziationsbefundes wird untersucht, ob der Befund durch die Stichprobenvergrößerung robuster wird.

2.2.2.2.3 Publikation: Shared genetic etiology between alcohol dependence and major depressive disorder

Frage 4: Zwischen Alkoholabhängigkeit und Depression ist auf klinischer Ebene eine Komorbidität belegt, wobei sich die Krankheiten gegenseitig bedingen könnten und/oder beide Krankheiten einen gemeinsamen genetischen Hintergrund haben. Hier soll die Frage beantwortet werden, ob Risikovarianten für Depression auch das Risiko für Alkoholabhängigkeit erhöhen.

Basierend auf den Daten der neuesten Depressions-GWAS (PGC-MDD2), wird in der Kohorte von 1 333 alkoholabhängigen Patienten und 2 168 Kontrollen untersucht, ob alkoholabhängige Patienten einen erhöhten polygenen Risikoscore für Depression aufweisen.

2.3 Beitrag des Doktoranden

Die Stichproben wurden im Rahmen größerer Forschungsprojekte erhoben, untersucht und gemessen, und die Arbeiten z.T. im Rahmen internationaler Kooperationen publiziert. Der Anteil des Autors der vorliegenden Arbeit umfasst dabei insbesondere:

- a) die inhaltliche Erarbeitung der oben genannten Fragestellungen, Erstellung von Erklärungsmodellen, Operationalisierung und Überprüfung der Ansätze.
- b) die Vorbereitung der Datenanalyse. Diese beinhaltet die Anforderung, Sammlung und Harmonisierung der Phänotypdaten der einzelnen Teilkollektive, die Anforderung und Beantragung notwendiger zusätzlicher externer Datensätze (Trainingsstichproben für Schizophrenie- und Depressions-Scores), für die Alkoholstichprobe darüber hinaus die Auswahl der zu typisierenden Individuen.
- c) die Vorverarbeitung der im Labor bestimmten Genotyp-Daten, für die Studie zur Alkoholabhängigkeit zusätzlich die Extraktion der Genotypen aus den Intensitäts-Rohdaten und die Filterung derselben im Rahmen der Qualitätskontrolle sowie die lokale Imputation von zusätzlichen nicht typisierten genetischen Markern.
- d) für die Fragestellungen 1 und 2 die komplette alleinige Durchführung der statistischen Analyse, für Frage 3 die Anleitung und gemeinsame Durchführung der Analyse mit dem Co-1st-Autor der zugehörigen Publikation, für Frage 4 die Anlei-

tung zur Daten-Analyse und die gemeinsame Durchführung derselben mit den beiden Erstautoren der genannten Publikation.

- e) die Interpretation der Ergebnisse zu den genannten Fragestellungen
- f) die Abfassung der jeweils zugehörigen Publikationen (Erstautor bei den Publikationen zu Fragen 1 und 2, geteilte Erstautorenschaft für Frage 3, Seniorautor bei Frage 4)

3 MATERIAL UND METHODEN

3.1 Stichproben

Für die vorliegende Arbeit wurden in Kooperation mit dem Humangenetischen Institut der Universität Bonn Genotypdaten durch genomweite Assoziationsstudien zu Alkoholabhängigkeit und Schizophrenie generiert. Die Patienten und Kontrollkollektive wurden durch unsere Arbeitsgruppe bzw. durch Kooperation mit klinischen Partnern im Rahmen von DFG und BMBF geförderten Projekten rekrutiert und charakterisiert.

Alle Individuen waren ausführlich aufgeklärt worden und hatten danach Ihr schriftliches Einverständnis erteilt. Die Einzelstudien, im Rahmen derer die Rekrutierung erfolgte, wie auch die Studien zur Auswertung der gesammelten genetischen Daten waren von den jeweils zuständigen Ethik-Kommissionen positiv votiert worden. Der Studienablauf erfolgte in Übereinstimmung der Deklaration von Helsinki.

3.1.1 Schizophrenie-Studie

Die Patienten stammen aus der BoMa Stichprobe (Rietschel et al., 2011). Hierbei handelt es sich um 1 540 stationär behandelte Patienten mit Schizophrenie, welche sukzessive in verschiedenen psychiatrischen Kliniken in Deutschland für genetische Studien rekrutiert wurden. Die Lebenszeitsymptome der Patienten wurden mit einem umfangreichen Instrumentarium, der IPGS (Fangerau et al., 2004) erhoben. Die Diagnose wurde basierend auf einem strukturierten SCID (First, New York State Psychiatric Institute, & Biometrics Research Department, 1997) bzw. SADS (Endicott, 1978) -Interview nach DSM-IV gestellt. Hauptquelle für die klinischen Informationen im Rahmen der vorliegenden Studie ist die Operational Criteria Checklist (OPCRIT; McGuffin, 1991). Die Kodierung der Response auf Clozapin erfolgte binär, basierend auf einer ursprünglich vierstufigen Einteilung in 0=keine Veränderung oder sogar Verschlechterung, 1=leichte Besserung der Symptome, 2=deutliche Verbesserung der Symptome, und 3=fast vollständige Remission (Nöthen et al., 1995), wobei Clozapinresponse hier mindestens einen Wert von 2 auf der obigen Skala erforderte.

Als Kontrollstichprobe dienten populationsbasierte Kontrollen aus mehreren großen Kohortenstudien: Heinz Nixdorff Recall Studie, (HNR; Erbel et al., 2012),

popgen (Nöthlings & Krawczak, 2012), aus der Kooperativen Gesundheitsforschung in der Region Augsburg (KORA; Holle, Happich, Löwel, & Wichmann, 2005), sowie einer weiteren populationsbasierten Kontrollstichprobe aus der Region Augsburg/Ingolstadt, bei der nur psychiatrisch Gesunde eingeschlossen wurden (Lucae et al., 2006).

3.1.2 Studien zur Alkoholabhängigkeit

Es wurden ca. 1600 Patienten mit der Diagnose Alkoholabhängigkeit nach DSM-IV rekrutiert. Diese waren in fünf verschiedenen Universitäts-Kliniken in Regensburg, Mannheim, München, Bonn/Essen/Düsseldorf/Homburg und Mainz zur stationären Behandlung aufgenommen worden. Alle Patienten wiesen eine schwere Alkoholabhängigkeit gemäß DSM-IV-Kriterien auf, wobei für die vorliegende Studie eine Konsens-Diagnose von zwei Psychiatern erforderlich war. Die Patienten waren mittels semistrukturierter Interviews erhoben worden.

Wegen finanzieller Ressourcenknappheit erfolgte die Genotypisierung der Patienten und der Kontrollsamples sukzessive: In der ersten Welle wurden zunächst 487 Patienten und 1 358 Kontrollen genomweit typisiert. In der zweiten Welle waren es weitere 900 Patienten. Die Daten von 861 weiteren Kontrollen wurden vom MPI für Psychiatrie in München zur Verfügung gestellt. Der kombinierte Datensatz aus beiden Wellen umfasste nach Qualitätskontrolle 463 044 autosomale SNPs von insgesamt 1 333 Fällen und 2 168 Kontrollen. Die Patienten dieser kombinierten Stichprobe waren im Durchschnitt bei der Erhebung $41,9 \pm 8,4$ Jahre alt und mit $30,2 \pm 8,0$ Jahren erkrankt. Die Kontrollen waren zum Zeitpunkt des Interviews im Mittel $50,5 \pm 12,1$ Jahre alt.

Als Kontrollstichprobe dienten dieselben populationsbasierten Kohorten, die bereits in Abschnitt 3.1.1 beschrieben wurden.

Für die Erweiterung der Patienten-Stichprobe wurden Patienten mit alkoholbedingter Leberzirrhose ($n=400$) bzw. alkoholinduzierter chronischer Pankreatitis ($n=1 110$) eingeschlossen. Als Kontrollstichprobe für diese Teilstichproben konnte eine weitere Kontrollstichprobe einbezogen werden aus der Heinz-Nixdorff-Recall-Studie (Erbel et al., 2012), die vom zuständigen Studienzentrum zur Verfügung gestellt wurde. Alle eingeschlossenen Personen sind mitteleuropäischer Abstammung.

Für die Untersuchung der polygenen Risikoscores in unabhängigen Case-Control Samples mit Alkoholabhängigkeit wurden zusätzlich öffentlich zugängliche Datensätze aus den USA und Australien herangezogen. Dies umfasste namentlich die Datensätze folgender Kollektive: (1) 'Alcohol Research using Australian twins and their families' (OZ-ALC; https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000181.v1.p1&phv=73489&phd=1551&pha=&pht=793&phvf=&phdf=&phaf=&phtf=&dssp=1&consent=&temp=1); (2) 'CIDR: Collaborative Study on the Genetics of Alcoholism' (COGA; <https://www.ncbi.nlm.nih.gov/gap/?term=CIDR%3A+Collaborative+Study+on+the+Genetics+of+Alcoholism>) und (3) 'Study of Addiction: Genetics and Environment' (SAGE; <https://www.ncbi.nlm.nih.gov/gap/?term=Study+of+Addiction%3A+Genetics+and+Environment>). Bei allen drei wurden nur Individuen mit Europäischer Abstammung in die vorliegende Arbeit eingeschlossen

Für die Untersuchung der polygenen Risikoscores für Depression in Case-Control Samples mit Alkoholabhängigkeit wurden die Patienten und Kontrollen aus Abschnitt 2.1.1 mit 597 Patienten mit Major Depression nach DSM-IV zusammengeführt. Die Kontrollstichprobe umfasste final 1292 Individuen.

Als Trainingsstichprobe für die Erstellung der Polygenen Risikoscores für Depression diente die PGC-MDD2-Stichprobe (Wray et al., 2018). Diese ist etwa acht Mal so groß ($n_{\text{Patienten}}=59\,851$, $n_{\text{Kontrollen}}=113\,154$) wie die, welche von Andersen et al. (2017) zu Grunde gelegt wurde. Insbesondere wurde untersucht, ob ein Zusammenhang sichtbar ist, wenn nur eine Teilstichprobe des Alkohol-Samples verwendet wird, die explizit unter Ausschluss komorbider Depression erhoben wurde.

3.2 Probengewinnung und Genetische Analyse

Für die Genotypisierung wurde den Patienten Vollblut entnommen, mittels Standardmethoden DNA extrahiert und über mehrere Wellen mit den jeweils aktuellen verfügbaren Illumina Arrays HumanHap 550v3, Human 610 und Human 660quad bei Illumina, bzw. bei der Life & Brain GmbH in Bonn Individualgenotypen bestimmt.

3.3 Benutzte Programme/Softwarepakete

Folgende Programme und Programmpakete wurden im Rahmen der in diese Arbeit einfließenden Studien benutzt:

- PLINK Versionen 1.07 (Purcell et al., 2007) und 1.9 (Chang et al., 2015); und R Versionen 2 und 3 (R Development Core Team, 2010) für die Qualitätskontrolle (QC), die Erstellung von Polygenen Risiko-Scores und die Assoziationstestung von Einzelmarkern und Scores
- SMARTPCA (Patterson, Price, & Reich, 2006) sowie PLINK 1.9 für die Hauptkomponentenanalyse zur Outlierfilterung und Stratifikationskorrektur
- Magma (de Leeuw, Mooij, Heskes, & Posthuma, 2015) für Gen und Gen-Set bzw. Pathway-basierte Analysen
- PRSice (Euesden, Lewis, & O'Reilly, 2014) für Score-basierte Analysen zu Frage 4

3.4 Statistische Verfahren

3.4.1 Qualitätskontrolle (QC) der Genetischen Daten

3.4.1.1 Technische QC – Filterung von Proben und Markern

Es wurden folgende Kriterien für die Filterung verwendet

- a) Für die DNA-Proben: Missing-Rate $\leq 2\%$ und Übereinstimmung von dem in der Akte vermerkten mit dem aus den Genotypen bestimmten Geschlecht. Im Falle von Duplikaten oder kryptischen Verwandtschaften (Identity by State (IBS) über autosomale Marker $> 1,6$), wurde die Probe mit der höheren Missing-Rate entfernt.
- b) Getestete Marker: Nur die SNPs, die gleichzeitig auf allen für die jeweilige Stichprobe benutzten Chips vorhanden waren, wurden in die Analyse einbezogen. Zum Verbleib im Datensatz waren folgende Kriterien erforderlich: Call rate (CR) $\geq 0,98$; Frequenz des selteneren Allels (minor allele frequency MAF) $\geq 0,01$; Übereinstimmung mit dem Hardy-Weinberg-Gleichgewicht (HWE; $P \geq 1E-6$) in allen Teilstichproben, und kein signifikanter Unterschied in der Allelfrequenz zwischen

den verschiedenen Kontroll-Teilstichproben. Für die stärksten Befunde wurden am Schluss die Clusterplots mit den Farbintensitätsrohwerten für eine abschließende Absicherung zur Überprüfung der Typisierungsqualität herangezogen, und SNPs mit schlechter Clusterqualität wurden entfernt.

3.4.1.2 Filterung von genetischen „Ausreißern“ mittels Hauptkomponenten-Analyse

Hierfür wurde eine Teilmenge der SNPs verwendet mit folgenden Charakteristika: Nur autosomale Marker, paarweise unabhängig ($r^2 < 0,05$) innerhalb eines Fensters von 200 SNPs Breite, Frequenz des seltenen Allels $> 5\%$, Hardy-Weinberg-Gleichgewicht der Genotypverteilung ($P_{HWE} > 0,01$).

Es wurde iterativ gefiltert, indem über die genannten SNPs eine PCA durchgeführt wurde und jeweils alle Proben mit einer Abweichung auf mindestens einer der ersten 10 Hauptkomponenten um mehr als 6 Standardabweichungen vom Mittelwert aus dem Datensatz entfernt wurden, so lange, bis entweder keine Probe mehr auffällig war oder eine Maximalzahl von fünf Durchgängen erreicht war. Abschließend wurde die PCA auf dem gefilterten Datensatz erneut durchgeführt. Die bei diesem letzten Durchgang extrahierten Hauptkomponenten wurden als Kovariaten in das Regressionsmodell eingeschlossen um die Variablen für die nachfolgende Stratifikationskorrektur zu gewinnen.

3.4.2 Testverfahren

3.4.2.1 Einzelmarkeranalyse

Einzelmarker wurden in einem einlogistischen Regressionsmodell auf ihren Zusammenhang mit dem jeweiligen Phänotyp getestet. Dabei wurde jeweils für Populationsstratifikation korrigiert durch Einbezug der ersten beiden Hauptkomponenten aus der PCA

3.4.2.2 Multimarkeranalysen

Das Hauptproblem bei genomweiten Einzelmarkertestungen ist die hohe Anzahl an durchgeführten Tests, für die korrigiert werden muss (mehrere 100 000 bis Millionen). Dies hat zur Folge, dass nur Marker mit sehr starken Effekten gefunden werden können, wenn man nicht eine hohe Zahl falsch positiver Befunde in Kauf

nehmen will. Gleichzeitig geht man aber bei komplexen Erkrankungen von vielen beteiligten Markern mit jeweils nur kleinen Effekten aus. Darüber hinaus ignoriert man bei diesem Verfahren mögliche synergistische Effekte zwischen den beteiligten Markern, die evtl. gemeinsam in einem Gen oder einem Pathway von mehreren miteinander interagierenden Genen enthalten sind.

Um diese Probleme zu lösen, wurden verschiedene Ansätze entwickelt, die zum einen die Anzahl der durchgeführten Tests deutlich verringern, zum zweiten biologische Hintergrund-Information berücksichtigen, wie z.B. welche Varianten zusammen im gleichen Gen vorkommen oder auch, welche dieser Gene zu größeren Einheiten wie biologischen Pathways oder Funktionskategorien zusammengefasst werden können.

Haupt-Herausforderungen bei diesen Verfahren sind

- a) Die Pathways beinhalten unterschiedlich viele Gene. Diese wiederum sind unterschiedlich groß und mit unterschiedlich guter Abdeckung auf den Typisierungschips vertreten. D.h. Gene wie Pathways sind mit unterschiedlich vielen SNPs im Datensatz repräsentiert.
- b) Die auf den Chips vorhandenen Marker weisen eine heterogene Korrelationsstruktur auf, bei manchen Genen sind stark interkorrelierende Marker vorhanden, bei anderen nur schwach zusammenhängende.

Daher beinhalten die aktuellen Verfahren Methoden, um für jeweils wenigstens eine der beiden genannten Fehlerquellen zu korrigieren.

3.4.2.2.1 Gen-basierte Analysen

In der vorliegenden Arbeit wurde hier der Ansatz einer Hauptkomponentenregression gewählt: Zunächst wurde für jedes Gen eine PCA über die enthaltenen SNPs durchgeführt. Nach der Entfernung von Komponenten mit sehr kleinen Eigenwerten wurde dann eine Regression des Phänotyps auf die extrahierten Hauptkomponenten vorgenommen, jeweils mit Stratifikationskorrektur durch Einbindung entsprechender Kovariaten. Dieser Ansatz berücksichtigt implizit auch die z.T. sehr hohe Interkorrelation der untersuchten Marker.

3.4.2.2.2 Genset-basierte Analysen

Hier gibt es grob zwei verschiedene Ansätze, zum einen sog. „self-contained“-Methoden, die testen, ob das Genset, also die Gesamtheit aller im Set enthaltenen Gene, im Schnitt eine überzufällig starke Assoziation mit dem Zielphänotyp aufweist.

Die zweite Gruppe ist die der „kompetitiven“ Vergleiche. Dort wird nicht die Assoziation des Gensets für sich mit dem Phänotyp getestet, sondern sie wird mit der Assoziation der Gesamtheit aller Gene verglichen, die nicht Bestandteil des Sets sind.

In der vorliegenden Arbeit wurde hierbei dem kompetitiven Ansatz der Vorzug gegeben, um eine Aussage zu gewinnen, welches der Gensets im Vergleich zu den anderen am stärksten assoziiert war.

Für den hier gewählten Ansatz wurden die Ergebnisse der genbasierten Analyse in Z-Werte transformiert, und dann wurde für alle 678 in der Reactome Pathway-Datenbank Version 5.1 verzeichneten Gensets der oben beschriebene Vergleich durchgeführt. Die gewählte Methode berücksichtigt auch ungleiche Genset-Größen.

Um adäquate Behandlung des z.T. kompletten Kopplungsungleichgewichts der SNPs in den Genen zu sicherzustellen wurde der Datensatz vorgefiltert und alle SNPs mit einem Varianzinflationsfaktor $VIF > 10$ innerhalb eines Fensters von 200 SNPs Breite entfernt.

3.4.2.2.3 Polygener Risikoscore

Es besteht auch die Möglichkeit die Information der im Datensatz vorhandenen Marker in ihrer Gesamtheit zu nutzen und ein genomweites Maß für die genetische Belastung eines Individuums für eine bestimmte Störung zu konstruieren. Der Polygene Risikoscore (Purcell et al., 2009) stellt ein solches Maß dar.

Für die Konstruktion eines solchen Scores sind zunächst genomweite Ergebnisse aus einer unabhängigen Studie (Trainings- oder Discovery-Sample) erforderlich. Aus diesen lässt sich für jeden Marker ein Risiko-Allel für die Störung bestimmen, sowie ein Effektmaß für den Zusammenhang des Markers mit der Störung. Die Berechnung des Scores für eine Person in der eigentlichen Zielstichprobe

geschieht in folgender Weise: Zunächst wird eine Auswahl der einzubeziehenden Marker anhand ihrer Testergebnisse in der Trainingsstichprobe getroffen (z.B. anhand des dort erzielten p-Werts, indem genau die SNPs betrachtet werden, deren Ergebnis unterhalb einer best. Fehlerwahrscheinlichkeit bleibt). Dann wird für jeden eingeschlossenen Marker errechnet, wie viele Risiko-Allele die betrachtete Person jeweils aufweist. Diese werden über alle einbezogenen Marker aufsummiert, wobei bei jedem einzelnen Marker noch eine Gewichtung mit der in der Trainingsstichprobe ermittelten Effektstärke erfolgen kann, und am Schluss durch die gewichtete Gesamtzahl der für diese Person vorhandenen Genotypen im betrachteten Marker-Set geteilt. Um nicht einzelne Regionen des Genoms mit vielen hochkorrelierten SNPs überzugewichten, trifft man eine Teilauswahl aus dem Gesamtset der angeordneten SNPs, welche paarweise möglichst unabhängig sind.

Die Testung des Polygenen Scores auf Assoziation mit dem Krankheitsstatus in der Zielstichprobe wurde mittels eines logistischen Regressionsmodells Ansatz durchgeführt. Dazu wurden zwei Modelle miteinander verglichen: (1) nur die aus der PCA ermittelten Hauptkomponenten als Prädiktoren und Krankheitsstatus als Antwortvariable, (2) das gleiche Modell, aber mit dem polygenen Score als zusätzlichem Prädiktor. Es wurde dann getestet, ob das zweite Modell eine signifikant bessere Vorhersage des Phänotyps ermöglicht.

Für die Studie zur Schizophrenie diente die Stichprobe aus der Arbeit von Ripke et al (Ripke et al., 2011) als Trainingsstichprobe, da hier genomweite Assoziationsergebnisse für über 21 000 Personen vorlagen. Da speziell bei kleineren Stichproben abhängig vom gewählten Vererbungsmodell die Teststärke mit der Anzahl der einbezogenen SNPs tendenziell ansteigt (Purcell et al., 2009), wurde für die primäre Fragestellung ein einzelner Grenzwert für die SNP-Auswahl angenommen ($p < 0,5$) und nur autosomale, paarweise unabhängige ($r^2 < 0.25$) SNPs einbezogen.

Für die Alkoholstudie wurde die eigene Stichprobe im Rahmen einer Kreuzvalidierung stratifiziert (gleicher Anteil von Patienten und Kontrollen) in zwei Hälften geteilt und jede Hälfte als Trainingsstichprobe für die jeweils andere verwendet. Darüber hinaus wurde auf Basis der Assoziationsergebnisse in der deutschen Gesamtstichprobe von 1 333 Patienten und 2 168 Kontrollen ein polygener Risiko-Score für Alkoholabhängigkeit in den aus dbGaP (Mailman et al., 2007) abgerufenen

Datensätzen zu den Studien COGA, SAGE und OZALC berechnet und getestet. Für die Konstruktion des polygenen Risiko-Scores wurden nur autosomale, paarweise unabhängige ($r^2 < 0.25$) SNPs mit einer MAF > 0.02 benutzt, die in der Trainingsstichprobe jeweils mit der Diagnose mit $P < 0.5$ assoziiert waren.

Für die Studie zum gemeinsamen genetischen Hintergrund von Depression und Alkoholabhängigkeit diente als Trainingsstichprobe die kürzlich erschienene Meta-Analyse des PGC zu Depression (Wray et al., 2018). Dabei wurden alle Individuen, die Teil einer der beiden Zielstichproben waren (Patienten wie auch Kontrollen) ausgeschlossen, sodass am Schluss 59 265 Patienten und 112 092 für die Trainingsstichprobe enthalten waren. Für die Konstruktion des polygenen Risiko-Scores wurden nur autosomale, paarweise unabhängige ($r^2 < 0.1$) SNPs einbezogen. Es wurden verschiedene P-Wert-Grenzen für die Filterung von SNPs nach ihren Assoziationssignalen in Betracht gezogen: 5×10^{-8} , 1×10^{-6} , 1×10^{-4} , 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 1.0. Als Zielstichprobe diente zum einen die komplette Stichprobe von 1 333 alkoholabhängigen Patienten und 2 168 Kontrollen aus Abschnitt 2.1.1, sowie in einem zweiten Schritt eine Teilstichprobe von 331 dieser Patienten, bei denen komorbide Depression explizit ausgeschlossen worden war. Zur Validierung des polygenen Risiko-Scores für Depression wurde eine Stichprobe von 597 Patienten mit Depression sowie 1 295 Kontrollen (siehe Abschnitt 2.1.1) als Zielstichprobe gesetzt.

4 ERGEBNISSE

4.1 Schizophrenie-Studie

4.1.1 Ergebnisse der Publikation: Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients

4.1.1.1 Verteilung klinischer Risikomerkmale in den Teilstichproben

Patienten, bei denen Clozapinbehandlung erforderlich war, waren bereits vor Krankheitsbeginn häufiger sozial schlecht angepasst ($p=8,1 \times 10^{-5}$; OR=1,85; CI=[1,35–2,55]; siehe Tabelle 1). Darüber hinaus begann in dieser Gruppe die Krankheit im Mittel früher ($\bar{x}=23,0$; SD=8,26) als bei den Patienten, die auf Erstmedikation ansprachen ($\bar{x}=25,5$, SD=9,12; $p=3,0 \times 10^{-6}$), und auch häufiger schleichend ($p=4,4 \times 10^{-3}$).

Tabelle 1: Assoziation von Behandlungsresistenz mit prämorbidem klinischen Symptomen (signifikante Items nach Bonferroni-Korrektur grau hinterlegt:

OPCRIT-Item	p-Wert	Adjustierter p-Wert	Odds Ratio [Konfidenzintervall]
3 Sex	0,56	1	1,08 [0,84, 1,39]
4 Age of onset	$3,0 \times 10^{-6a}$	$3,0 \times 10^{-5a}$	NA
5 Mode of onset	$4,4 \times 10^{-3}$	0,04	NA
9 Poor premorbid work adjustment	0,83	1	1,04 [0,74; 1,45]
10 Poor premorbid social adjustment	$8,1 \times 10^{-5}$	$8,1 \times 10^{-4}$	1,85 [1,35; 2,55]
11 Premorbid personality disorder	0,60	1	1,13 [0,69; 1,86]
12 Alcohol / Drug abuse within year of onset	0,28	1	0,84 [0,6; 1,17]
13 Family history of schizophrenia	0,95	1	1,01 [0,74; 1,38]
14 Family history of other psychiatric disorder	0,59	1	1,07 [0,83; 1,39]
16 Definite psychosocial stressor prior to onset	0,48	1	1,11 [0,82; 1,52]

4.1.1.2 Polygener Risiko-Score

In die Berechnung des polygenen Risiko-Scores gingen 58 663 SNPs mit ein. Der polygene Risiko-Score für Schizophrenie ist in Patienten deutlich erhöht gegenüber populationsbasierten Kontrollen ($p=1,1 \times 10^{-41}$). Behandlungsresistente Patienten wiederum haben höheren polygenen Risiko-Score als Patienten bei denen keine Clozapinbehandlung erforderlich war ($p=0,02$). Den höchsten polygenen Risiko-Score für Schizophrenie wiesen Patienten auf, bei denen neben der Clozapin-Resistenz noch zusätzliche klinische Risikomerkmale wie früher und schleichender Krankheitsbeginn, sowie schlechte prämorbid soziale Anpassung vorliegen (Abbildung 3).

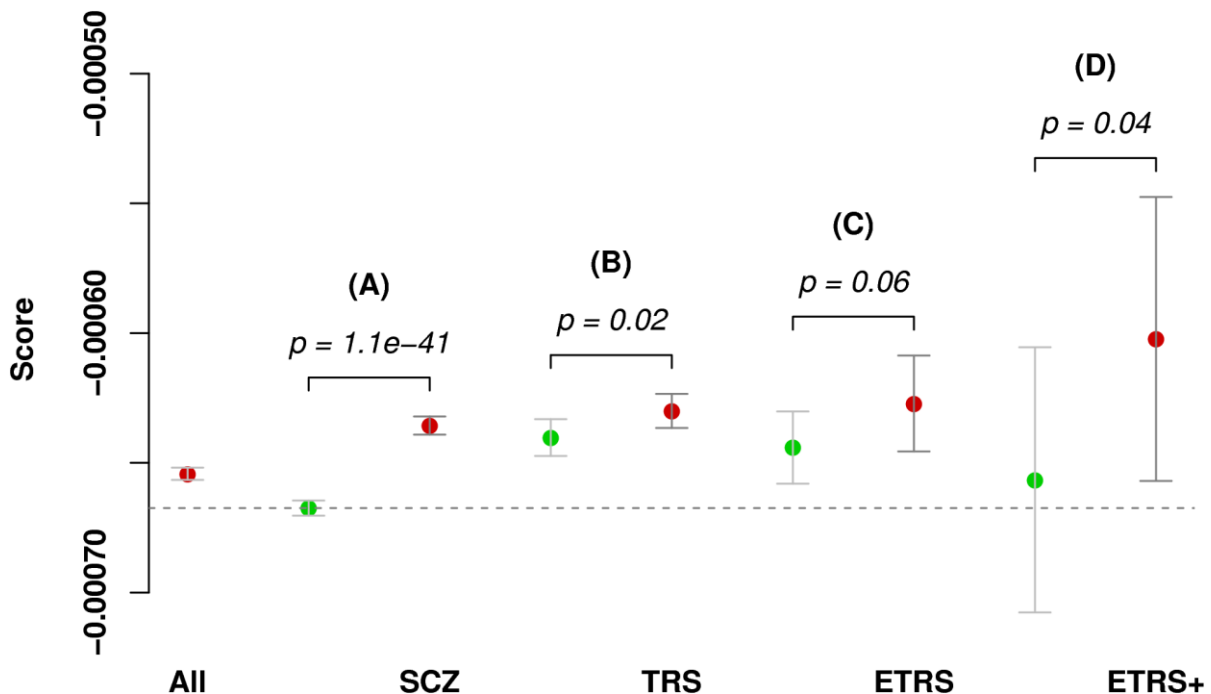


Abbildung 3: Vergleich der polygenen Risiko-Scores für Schizophrenie in den einzelnen Teilstichproben: (A) populationsbasierte Kontrollen vs. Patienten; (B) Patienten ohne vs. Patienten mit behandlungsresistenter Schizophrenie (TRS); (C) Patienten ohne vs. mit extremer TRS; (D) Patienten nur mit ETRS vs. Patienten mit ETRS und zusätzlich schlechter soziale Anpassung vor Krankheitsbeginn sowie frühem und schleichendem Einsetzen der Krankheit; die jeweilige Fallkohorte rot und die zugehörigen Kontrollen grün dargestellt; gezeigt sind geschätzte Randmittel mit 95%-Konfidenzintervallen. Die gestrichelte Linie zeigt den Mittelwert der populationsbasierten Kontrollstichprobe als Referenz

Zusätzlich wurden Sekundär-Analysen mit den oben genannten prämorbid Merkmalen durchgeführt, um zu sehen, ob und inwieweit diese die genetische Gesamtbelastung repräsentieren (Tabelle 2).

Tabelle 2: Assoziation von prämorbid klinischen Merkmalen mit polygenen Risiko-Scores (^alogistische Regression; ^blineare Regression; ^cschleichender Beginn, definiert als 2-stufige Einteilung der OPCRIT-Variablen „mode of onset“ in <6 Monate vs. > 6 Monate)

Untersuchte Gruppe	Poor premorbid social adjustment ^a	Age at onset (quantitativ) ^b	Insidious disease onset ^{a,c}	Family history ^a
Alle Patienten mit GWAS-Daten (n=1540)	0.08	0.005	0.95	0.10
TRS (Clozapinpatienten; n=434)	0.02	0.002	0.15	0.15
ETRS (Clozapin-Non-Responder) (n=45)	0.10	0.06	0.13	0.62

Innerhalb der extrem behandlungsresistenten Patienten, d.h. bei Clozapin-Non-Respondern (ETRS) wird das Assoziationssignal durch Adjustierung für obige klinische Merkmale nicht aufgehoben (Tabelle 3).

Tabelle 3: Einseitige p-Werte aus der logistischen Regression zum Zusammenhang zwischen (extremer) Behandlungsresistenz und polygenen Risikoscores für Schizophrenie nach Korrektur für assoziierte klinische Risikofaktoren aus Tabelle 2 (^aGrenzwert der Einzelstatistik für den Einschluss betreffender genetischer Marker in den Score; ^bin Klammern zum Vergleich ohne Adjustierung aber nach Ausfilterung von Beobachtungen mit fehlenden Daten für mind. eine der prämorbid klinischen Variablen um die gleiche Stichprobengröße zu gewährleisten)

p-Wert cut-off ^a	TRS	ETRS
<0,01	0,58 (0,46) ^b	0,80
<0,05	0,05 (0,01)	0,35
<0,1	0,10 (0,03)	0,20
<0,2	0,04 (0,01)	0,06
<0,3	0,05 (0,02)	0,07
<0,4	0,07 (0,02)	0,07
<0,5	0,06 (0,02)	0,05

4.1.1.3 *Beantwortung der Frage 1: Ist eine erhöhte genetische Belastung für Schizophrenie mit dem Therapieansprechen bzw. dem Krankheitsverlauf schizophrener Patienten assoziiert?*

Die Ergebnisse zeigen, dass eine durch den polygenen Risiko-Score abgebildete erhöhte genetische Belastung für Schizophrenie mit einem schlechteren Therapieansprechen einhergeht.

4.2 Studien zur Alkoholabhängigkeit

4.2.1 Ergebnisse der Publikation: Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster

Die mittlere Call Rate der SNPs (Fälle und Kontrollen kombiniert) lag bei 99,94% (SD = 0,12%). Abbildung 4 zeigt die Scores der Probanden auf den ersten beiden in der PCA extrahierten Hauptkomponenten nach Entfernung der genetischen Outlier.

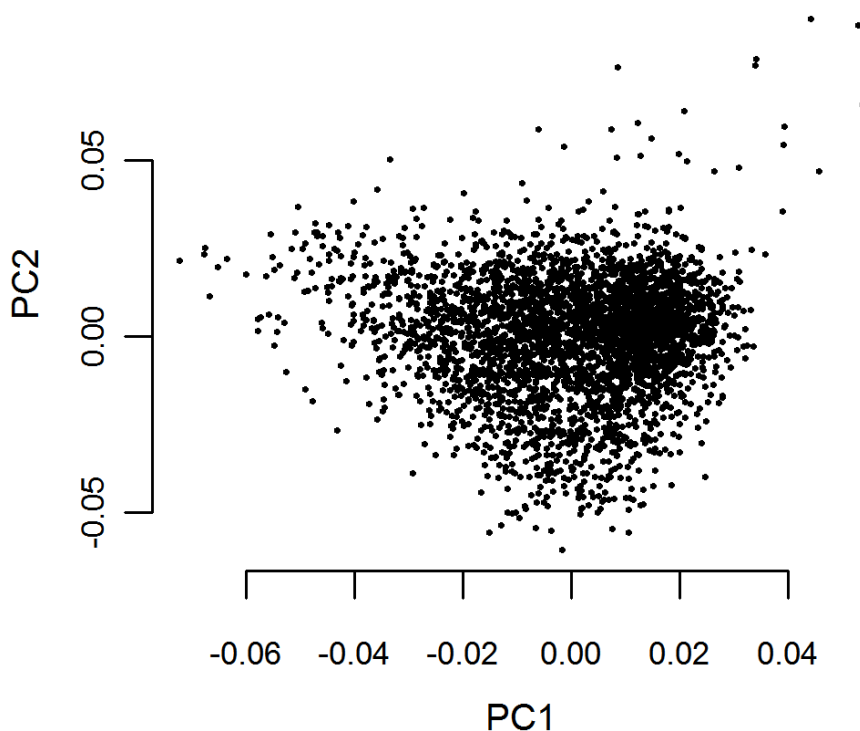


Abbildung 4: Scatterplot der ersten beiden aus der PCA resultierenden Hauptkomponenten

Um das die Auswirkung der QC-Schritte bei der Filterung der SNPs zu visualisieren, sind die P-Werte vor und nach QC in Abbildung 5 als Quantil-Quantil-Diagramm veranschaulicht. Eine ganze Reihe von genomweit signifikanten Befunden im Rohdatensatz ist auf Genotypisierungs-Artefakte zurückzuführen, die durch die Filterung entdeckt und entfernt wurden.

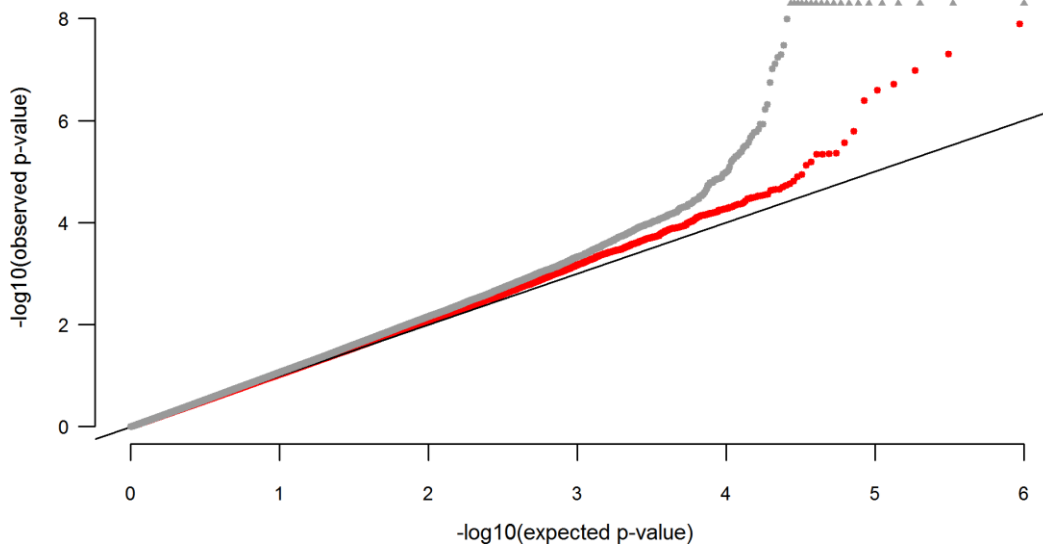


Abbildung 5: QQ-Plot zum Vergleich der genomweiten Einzelmarker-Ergebnisse vor (grau) und nach Qualitätskontrolle (rot). P-Werte $< 10^{-9}$ wurden am oberen Rand angetragen

4.2.1.1 Einzelmarkeranalyse

Die Ergebnisse der Einzelmarker-Analyse sind dargestellt in Abbildung 6. Ein SNP, rs1789891, lag unter der etablierten genomweiten Signifikanzgrenze von 5×10^{-8} ($p=1,3 \times 10^{-8}$, $OR=1,46$). Dieser SNP ist zwischen den Genen *ADH1B* und *ADH1C* lokalisiert und befindet sich im kompletten Kopplungsungleichgewicht mit der funktionellen Variante Arg272Gln.

Ergebnisse

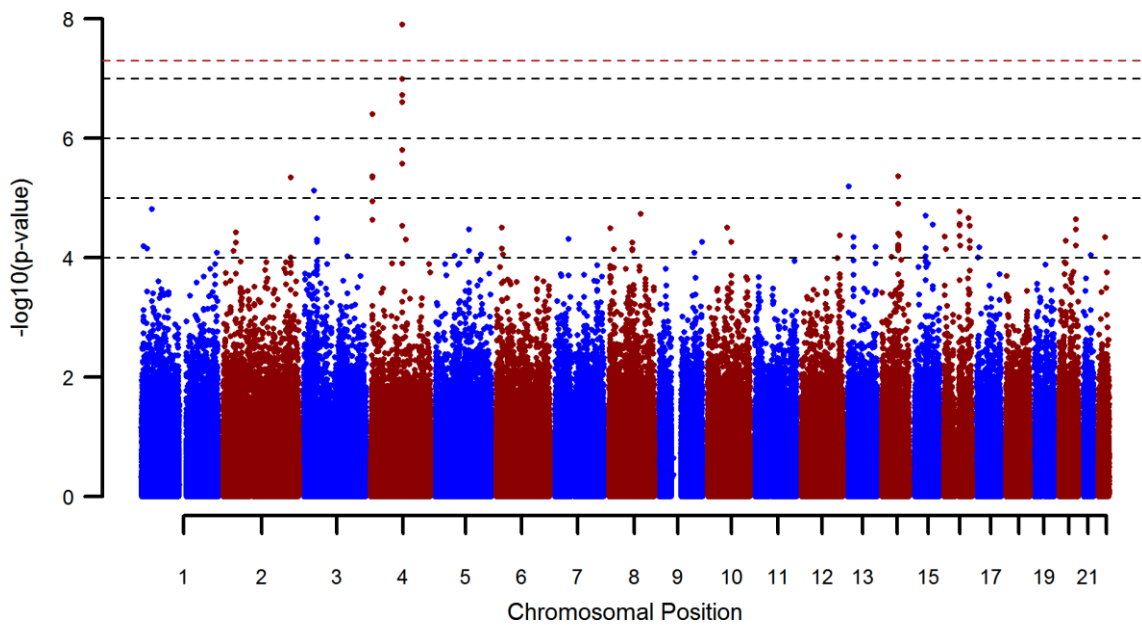


Abbildung 6: Manhattan-Plot zum Überblick über die genomweiten Einzelmarker-Ergebnisse

4.2.1.2 Polygener Score

In die Berechnung des polygenen Risiko-Scores gingen nach Filterung schließlich jeweils über 84 000 SNPs mit ein (außer bei OZALC, da dort ein anderer Chip mit geringerer Abdeckung verwendet worden war. Hier gingen 48 952 SNPs in die Score-Berechnung ein). In jeder der beiden Hälften der deutschen Primärstichprobe wiesen die Patienten einen signifikant höheren polygenen Risiko-Score auf als die populationsbasierten Kontrollen ($p_1=1,3 \times 10^{-6}$, Abbildung 7A; $p_2=9,7 \times 10^{-9}$, Abbildung 7B). In den beiden davon unabhängigen Stichproben von COGA und SAGE konnte diese Ergebnis repliziert werden ($p_{\text{COGA}}=3,9 \times 10^{-2}$, Abbildung 8A; $p_{\text{SAGE}}=1,1 \times 10^{-4}$, Abbildung 8B). In der Australischen Stichprobe dagegen ergab sich kein signifikanter Zusammenhang des polygenen Risiko-Scores mit Alkoholabhängigkeit ($p_{\text{OZALC}}=0,56$, Abbildung 8C).

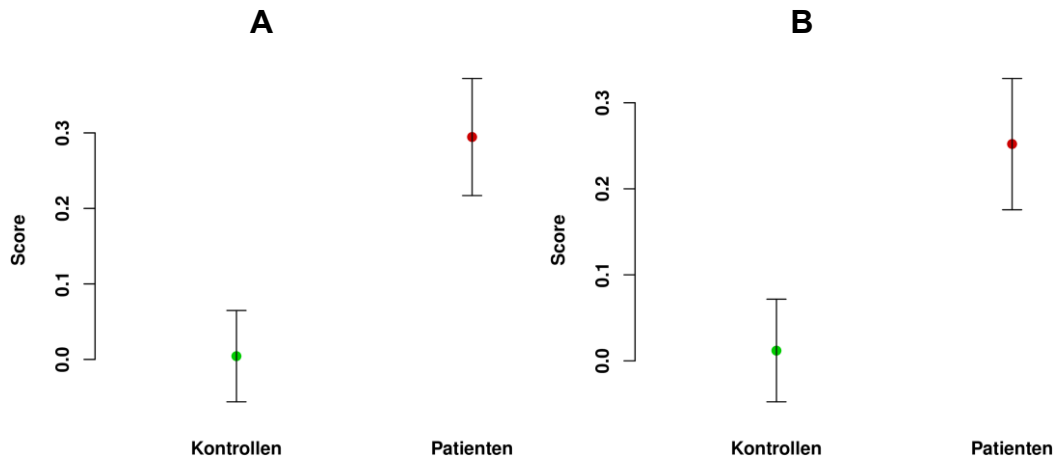


Abbildung 7: Standardisierte polygene Risiko-Scores für Alkoholabhängigkeit in den beiden deutschen durch stratifizierte Randomisierung für die 2-fach-Kreuzvalidierung erzeugten Teilstichproben; dargestellt sind geschätzte Randmittel der Gruppen mit 95%-Konfidenzintervallen, Scores sind standardisiert an den Werten der jeweils dargestellten Kontrollteilstichprobe

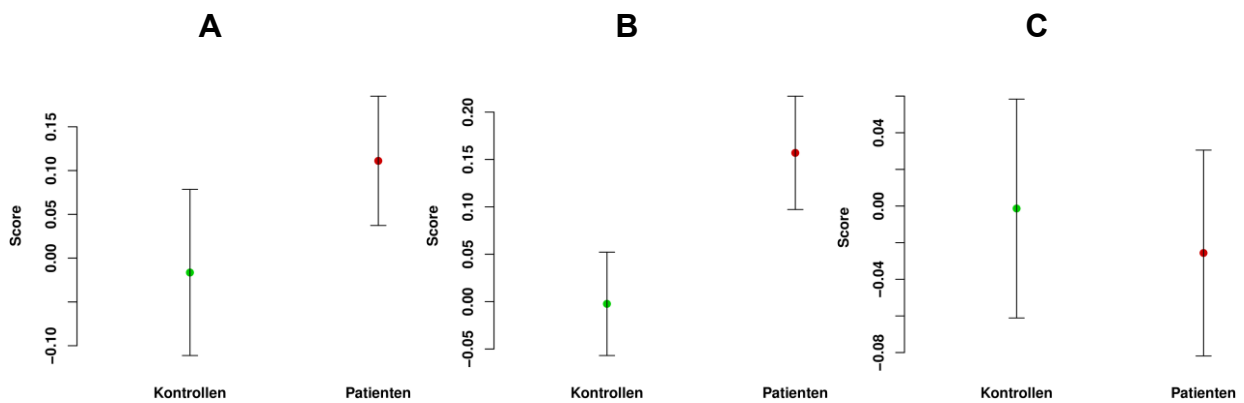


Abbildung 8: Standardisierte polygene Risiko-Scores für Alkoholabhängigkeit in den aus dbGaP abgerufenen Datensätzen zu COGA (A), SAGE (B) und OZALC (C); dargestellt sind geschätzte Randmittel der Gruppen mit 95%-Konfidenzintervallen, Scores sind standardisiert an den Werten der jeweils dargestellten Kontrollteilstichprobe

4.2.1.3 *Beantwortung der Frage 2:* Handelt es sich bei der Alkoholabhängigkeit ebenfalls um eine komplexe genetische Störung mit einem polygenen Anteil und tragen Genvarianten, die häufig in der Bevölkerung vorkommen, zur Entstehung von Alkoholabhängigkeit bei?

Der polygene Score, der auf Basis der deutschen Fall-Kontroll-Stichprobe in unabhängigen Case-Control Samples berechnet wurde, ist dort mit Alkoholabhängig-

keit assoziiert. Dieser Befund zeigt, dass eine Vielzahl häufiger Varianten zur Ätiologie der Alkoholabhängigkeit beitragen, und belegt somit eine polygene Komponente der Alkoholabhängigkeit.

Durch den GWAS Ansatz konnte eine häufig vorkommende Risikovariante identifiziert werden, die im *ADH*-Gen-Cluster lokalisiert ist.

4.2.2 Ergebnisse der Publikation: Genetic contribution to alcohol dependence: investigation of a heterogeneous German sample of individuals with alcohol dependence, chronic alcoholic pancreatitis, and alcohol-related cirrhosis

4.2.2.1 Einzelmarker-Analyse

Über die kombinierte Alkoholstichprobe hinweg konnten keine genomweit signifikanten Signale gefunden werden. Die Variante rs1789891 welche in Abschnitt 4.2.1.1 genomweit signifikant war, erzielt in der kombinierten Stichprobe nur noch ein deutlich abgeschwächtes Signal ($p=1,3 \times 10^{-5}$, OR=1,23). In den Teilstichproben mit ALZ/ACP war kein Signal erkennbar ($p=0.70$, OR=1,03; siehe Tabelle 4)

Tabelle 4: Genotypzahlen, Allelfrequenzen und Hardy-Weinberg-Testwerte für rs1789891 in der erweiterten Stichprobe

	Gesamtstichprobe	Originalstichprobe	Zusatzsample
p-Wert	$1,3 \times 10^{-5}$	$1,6 \times 10^{-8}$	0,64
Odds ratio	1,23	1,47	1,03
MAF (alle)	0,17	0,16	0,17
MAF (Patienten)	0,19	0,19	0,18
MAF (Kontrollen)	0,15	0,14	0,17
Genotypzahlen (alle)	200/1782/4507	94/872/2298	106/910/2209
Genotypzahlen (Patienten)	102/850/1879	53/410/868	49/440/1011
Genotypzahlen (Kontrollen)	98/932/2628	41/462/1430	57/470/1198

4.2.2.2 Genbasierte Analyse

Hier ergab sich nach Bonferroni-Korrektur auf eine Anzahl von 16 835 berücksichtigten Genen ($p_{\text{crit}}=2,6 \times 10^{-6}$) eine genomweite signifikante Assoziation mit dem

Gen *ADH1B* ($p=1,2 \times 10^{-6}$). Diese Assoziation wurde vom ursprünglichen Sample getragen ($6,5 \times 10^{-10}$), der neu hinzugekommene Teil der Stichprobe (ALZ und ACP-Patienten) lieferte kein Signal ($p=0,69$). Da der SNP rs1789891 von beiden Genen, *ADH1B* und *ADH1C*, weniger als 20kb entfernt liegt, wurde dieser SNP auch beiden Genen zugeordnet (Abbildung 9), und er liefert in der genbasierten Analyse den größten Beitrag zum Assoziationssignal.

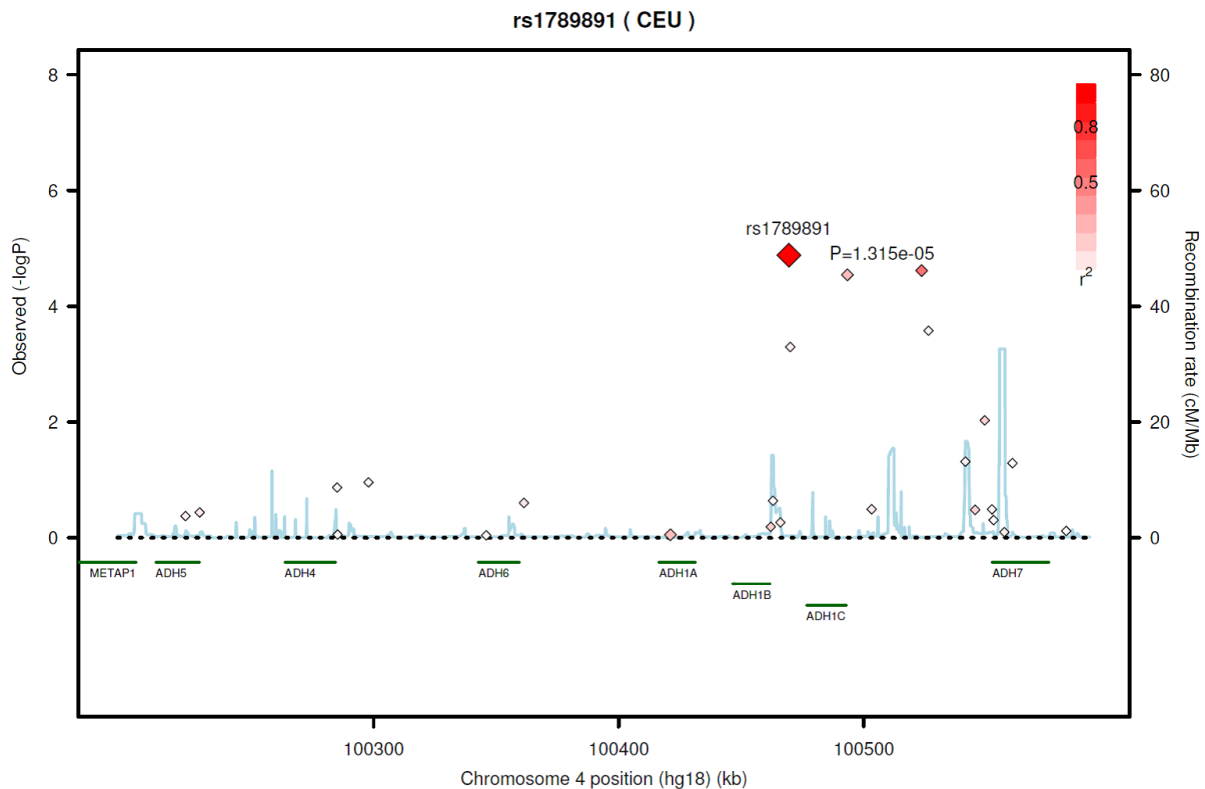


Abbildung 9: Regionaler Assoziationsplot der *ADH1B/ADH1C*-Region in der erweiterten Stichprobe

4.2.2.3 Genset-basierte Analyse

Der stärkste Befund in der kombinierten Stichprobe ergab sich für das Genset „Ethanol Oxidation“, welches wiederum das Gen *ADH1B* enthält ($p=2,2 \times 10^{-4}$).

4.2.2.4 *Beantwortung der Frage 3:* Ist eine Vergrößerung der Stichprobe auch auf Kosten der klinischen Homogenität zielführend, um weitere Gene und Pathways zu finden?

Die Studie belegt, dass eine Erhöhung der Stichprobenanzahl unter Hinzunahme von klinisch heterogenen Samples nicht automatisch zu einer Zunahme der Power, sondern sogar zu einem Verlust der Power und gegenläufigen Effekten führen kann.

4.2.3 Ergebnisse der Publikation: Shared genetic etiology between alcohol dependence and major depressive disorder

4.2.3.1 Assoziation von Alkoholabhängigkeit mit genetischer Belastung für Depression

Es ergab sich eine signifikante Assoziation von Alkoholabhängigkeit mit dem polygenen Risiko-Score für Depression (siehe Abbildung 10), sowohl basierend auf PGC-MDD1 ($P=1,4 \times 10^{-4}$) als auch für den von PGC-MDD2 abgeleiteten polygenen Risiko-Score ($P=6,3 \times 10^{-3}$). Bemerkenswerterweise ergab sich für die vergrößerte Trainingsstichprobe keine bessere Aufklärung der Varianz ($R^2=0,53\%$) als für die ursprünglich Stichprobe ($R^2=0,66\%$).

Im Gegensatz dazu war wie erwartet die Teststärke für die Vorhersage von Depression mit diesem Score mit der größeren Trainingsstichprobe deutlich verbessert ($P=3,8 \times 10^{-5}$, $R^2=1,34\%$), verglichen mit PGC-MDD1 ($P=1,3 \times 10^{-3}$, $R^2=0,81\%$).

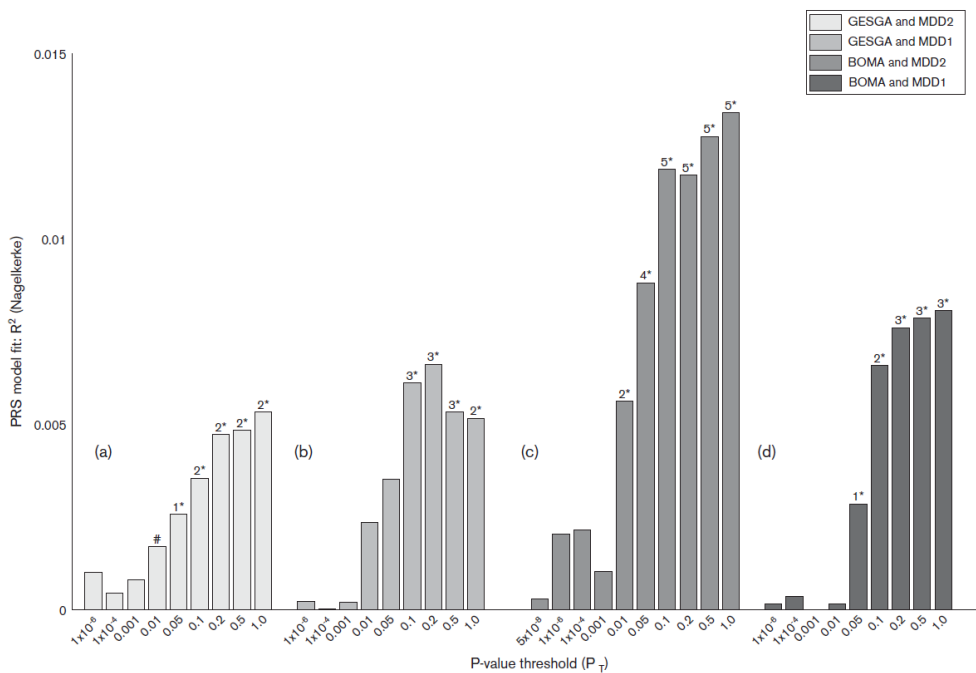


Abbildung 10: Assoziation von Diagnose mit polygenem Score für Depression: a) PGC-MDD2 auf Alkohol, b) PGC-MDD1 auf Alkohol, c) PGC-MDD2 auf BoMa MDD, d) PGC-MDD1 auf MoBa MDD

4.2.3.2 Assoziation nach Stratifizierung für komorbide Depression

In einer Teilstichprobe der alkoholabhängigen Patienten war komorbide Depression explizit ausgeschlossen worden (n=331, "PREDICT"-Teilstichprobe). Wurde ein Test nur über diese Teilstichprobe durchgeführt, war die Assoziation immer noch signifikant (P=0.042, R²=0,40%, siehe Abbildung 11).

Allerdings war beim Rest der Patienten mit Alkoholabhängigkeit, bei denen Depression kein Ausschlusskriterium darstellte, die Vorhersagekraft größer (P=0.0003, R²=0.69%).

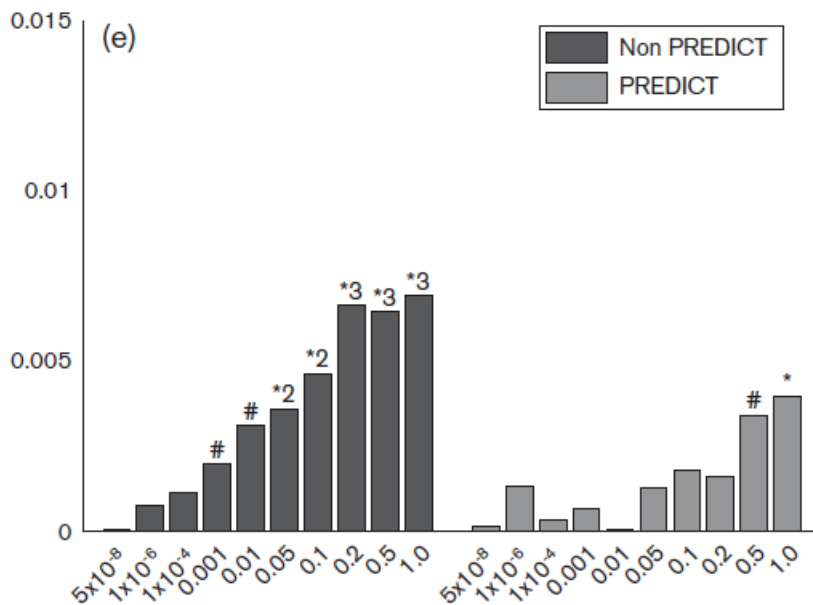


Abbildung 11: Assoziation von Alkoholabhängigkeit mit polygenem Risiko-Score für Depression, stratifiziert nach komorbider Depression

4.2.3.3 Beantwortung der Frage 4: Zwischen Alkoholabhängigkeit und Depression ist auf klinischer Ebene eine Komorbidität belegt, wobei sich die Krankheiten gegenseitig bedingen könnten und/oder beide Krankheiten einen gemeinsamen genetischen Hintergrund haben. Hier soll die Frage beantwortet werden, ob Risikovarianten für Depression auch das Risiko für Alkoholabhängigkeit erhöhen.

Die Ergebnisse zeigen, dass es zahlreiche, häufig vorkommende Varianten mit kleinem Effekt gibt, die sowohl zum Risiko für Depression als auch zum Risiko für Alkoholabhängigkeit beitragen

5 DISKUSSION

5.1 Schizophrenie

5.1.1 Publikation: Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients

Die Ergebnisse dieser Studie zeigen, dass Patienten mit einem schwereren Krankheitsverlauf einen höheren Risikoscore für Schizophrenie aufweisen als Patienten mit einem günstigeren Krankheitsverlauf, bei denen Medikamente besser ansprechen. Die Gruppe der Patienten mit den klinischen Symptomen eines frühen und schleichenden Krankheitsbeginns und schlechten Ansprechens auf Medikamente wiesen den höchsten Risikoscore im Vergleich zu den anderen Patienten auf. Wurden klinische Symptome und Risikoscore als Prädiktoren für den Krankheitsverlauf genutzt, waren die klinischen Symptome weitaus stärkere Prädiktoren, wobei sich die Effekte nur teilweise überlappten. Für eine individuelle Prädiktion ist die Aussagekraft natürlich noch viel zu gering (Area under the Curve = 63%). Aber die genetischen Befunde deuten darauf hin, dass Patienten, die eine Therapieresistenz aufweisen, möglicherweise an einer ätiologisch unterschiedlichen Form der Schizophrenie leiden, als solche, die gut respondieren, eine Beobachtung, die in Einklang mit neueren klinischen Befunden steht (Gillespie, Samanaite, Mill, Egerton, & MacCabe, 2017).

Zu bemerken ist, dass das Trainings-Set, auf dem die Analysen beruhen, eine entscheidende Rolle für die Ergebnisse spielt. Die hier berichteten Befunde beruhen auf Daten, die auf dem PGC-Schizophrenia1 als Trainings-Set basieren. Nach Vergrößerung des Training Sets um weitere Schizophrenie-Patienten wurde die Assoziation des polygenen Risiko-Scores mit der Diagnose Schizophrenie tatsächlich stärker, da das Trainingsset um Patienten mit der Diagnose Schizophrenie angereichert wurde und daher der prädiktive Wert für Schizophrenie zunahm. Allerdings nahm die Assoziation mit dem Therapieansprechen gleichzeitig ab. Eine plausible Erklärung dafür wäre, dass für die Aufstockung des Trainingssets eine populationsbasierte Kohorte von Patienten aus der schwedischen Bevölkerung hinzugefügt wurde. Das Einschlusskriterium für diese schwedische Kohorte war, an mindestens zwei Episoden einer Schizophrenie gelitten zu haben. Es ist davon auszugehen, dass diese Kohorte sich von der ersten Patientengruppe unterscheidet, die das ursprüngli-

che, kleinere Trainingsset darstellte. Die letztere war über Jahrzehnte hinweg für genetische Studien rekrutiert worden und für schwer erkrankte Patienten mit einer hohen familiären Belastung angereichert. Dies könnte dazu geführt haben, dass in der ersten Stichprobe eine Anreicherung von Varianten für einen schweren Verlauf vorhanden war, nicht aber in der zweiten Kohorte. Dieser Befund weist auf die Bedeutung der Zusammensetzung/Homogenität und Größe von Stichproben für genetische Studien hin.

5.2 Alkoholabhängigkeit

5.2.1 Publikation: Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster

In dieser GWA Studie wurde nur ein genomweit signifikantes Resultat erzielt und zwar eine Assoziation mit dem intergenisch, in der Region des *ADH*-Clusters zwischen den *ADH1B* und *ADH1C* Genen lokalisierten SNPs rs1789891. Wie biochemische Untersuchungen zeigen, spielen die Stoffwechsellzyme bei dem enzymatischen Abbau von Alkohol eine Rolle und führen zu dem toxischen Intermediärprodukt Acetaldehyd (Edenberg, 2007). Dieses wird normalerweise rasch durch das Enzym Aldehyddehydrogenase zu Acetat weiter verstoffwechselt. Eine Anhäufung von Acetaldehyd im Blut führt zu der sogenannten „Flushing“-Reaktion, die als äußerst unangenehm erlebt wird, und es wird angenommen, dass zu deren Vermeidung auf weiteren Alkoholkonsum verzichtet wird.

Assoziationen zwischen Alkoholabhängigkeit und Varianten in den *ADH*- und *ALDH*- Genen wurden auf die protektiven Effekte solcher Stoffwechselwirkungen zurückgeführt (Ball, 2008; Edenberg, 2007; Eng, Luczak, & Wall, 2007; Enomoto, Takase, Yasuhara, & Takada, 1991; Li, Zhao, & Gelernter, 2011; Macgregor et al., 2008, p. 200; Tawa et al., 2016). Insbesondere bei asiatischen Populationen ist die Auswirkung dieses Effekts besonders ausgeprägt: in dieser Population hat eine im *ALDH2* Gen liegende Variante, Glu504Lys, eine hohe Prävalenz. Sie führt zu einer starken Reduktion der Aktivität des Enzymes und erhöht so die Konzentration von Acetaldehyd nach Alkoholkonsum im Blut. Ähnliche Effekte werden auch für Varianten in den *ADH* Genen vermutet, wie z.B. für die funktionelle, häufiger in der Bevölkerung vorkommende Variante Arg272Gln, die mit der Trinkmenge assoziiert

gefunden wurde (Macgregor et al., 2008). Diese Variante ist nicht auf den für die Genotypisierung verwendeten Chips vorhanden, konnte aber erfolgreich lokal imputiert werden. Das Ergebnis nach lokaler Imputierung ist im vorliegenden Sample in Übereinstimmung mit dem von Macgregor et al. (2008) berichteten Befund: das dort als protektiv berichtete Allel ist auch in der vorliegenden Stichprobe das protektive Allel. Da die Assoziation mit der funktionellen Variante jedoch nicht stärker ist als mit dem hier berichteten genomweit signifikanten, intergenisch gelegenen SNP und beide SNPs in komplettem Kopplungsungleichgewicht zueinander stehen, kann nicht entschieden werden, auf welcher Variante der Assoziationsbefund beruht. In der vorliegenden Studie lagen insgesamt sechs der 15 Top-SNPs ($p < 10^{-5}$) der GWAS in dieser Region. Eine im Anschluss durchgeführte konditionelle Analyse legt zumindest kleine unabhängige Effekte dieser Varianten nahe. Frühere Studien haben gezeigt, dass Varianten im *ADH*-Cluster unabhängig voneinander wirken können (Macgregor et al., 2008).

Bisherige genetische Untersuchungen d.h. Kandidatengenuntersuchungen (durch biochemische Evidenz und Untersuchungsergebnissen an Mäusen getrieben) und auch genomweite Kopplungsuntersuchungen implizieren konsistent die Bedeutung dieser Region mit der Ätiopathogenese von Alkoholabhängigkeit (Birley et al., 2009).

Dass eine in diesem Cluster liegende Variante auch durch eine systematische GWAS zu Alkoholabhängigkeit identifiziert werden konnte, belegte hier zum ersten Mal, dass der GWAS-Ansatz bei Alkoholabhängigkeit zur Identifikation von Risikogenen prinzipiell sinnvoll sein kann. Dieser Befund wurde danach in unabhängigen Samples repliziert (Park et al., 2013; Gelernter et al., 2014; Way et al., 2015). Darüber hinaus stellt auch in der gerade erschienenen Meta-Analyse des PGC zu Alkoholabhängigkeit (Walters et al., 2018) diese Region immer noch den stärksten Befund dar.

Allerdings ist es, trotz der hohen Prävalenz von Alkoholabhängigkeit weltweit bisher noch nicht gelungen, wirklich ausreichend große Patientenzahlen für GWAS zu erheben, Im Gegensatz dazu gibt es für einen verwandten Alkoholphänotyp, nämlich Alkoholkonsum inzwischen GWAS mit bis zu 500.000 Personen umfassenden Stichproben aus der Normalbevölkerung (Clarke et al., 2017). Alkoholkonsum über-

lappt genetisch zwar nur inkomplett mit Alkoholabhängigkeit (Grant et al., 2009). Dennoch stellt die *ADH*-Region nun auch einen der Hauptbefunde für Alkoholkonsum dar (Clarke et al., 2017).

Da es sich allerdings um einen prinzipiell bekannten Befund handelt und in der vorliegenden Arbeit mit relativ kleinen Samplegrößen keine neuen Varianten identifiziert werden konnten, stellte sich die Frage, ob es sich bei der Alkoholabhängigkeit tatsächlich, wie vermutet, um eine komplexe genetische Störung handelt, d.h., ob tatsächlich eine Kombination weiterer häufiger Gene das Krankheitsrisiko erhöht.

Diese Vermutung wurde durch die vorliegende Studie zum ersten Mal auf molekularer Ebene bestätigt: Häufige Varianten, die in einem polygenen Score zusammengefasst wurden, konnten den Krankheitsstatus in einer unabhängigen Kohorte von Patienten und Kontrollen bestätigen. Allerdings ist die Vorhersagekraft äußerst gering und die Konfidenzintervalle sehr groß, was auf die kleinen Patientenzahlen zurückzuführen ist. Eingedenk dessen war es auch nicht das Bestreben der Arbeit, prominente Vorhersagewerte zu generieren, sondern prinzipiell zu belegen, dass es weitere häufige Risikovarianten für Alkoholabhängigkeit gibt, die bei ausreichender Stichprobengröße identifiziert werden können.

5.2.2 Publikation: Genetic contribution to alcohol dependence: investigation of a heterogeneous German sample of individuals with alcohol dependence, chronic alcoholic pancreatitis, and alcohol-related cirrhosis

Angesichts der im vorigen Abschnitt genannten Schwierigkeiten und in Anbetracht der Tatsache, dass in der vorliegenden Arbeit bereits mit einem eingeschränkt großen Sample ein genomweiter Befund erzielt werden konnte, wurde für eine weitere GWAS zur Alkoholabhängigkeit die Stichprobengröße verdoppelt. Allerdings standen von den klinischen Kooperationspartnern hierfür nur Patienten mit alkoholinduzierten Folgekrankheiten zur Verfügung.

Es stellte sich somit die Frage, ob eine weitere Vergrößerung der bisherigen Stichprobe um 1 510 Patienten und 1 750 Kontrollen – allerdings auf Kosten der klinischen Homogenität – zur Identifikation weiterer Gene und Pathways führen würde.

Trotz des größeren Stichprobenumfangs wurde auf der Ebene einzelner SNPs kein genomweit signifikanter Befund erzielt. Genweit wurde ein einziger genomweit signifikanter Befund erzielt und zwar wiederum mit dem *ADH1B* Gen. Allerdings wurde dieser Befund alleinig von der ursprünglichen oben genannten Stichprobe getragen ($p=6.5 \times 10^{-10}$) während die neue Stichprobe der Patienten mit komorbider Leberzirrhose bzw. Pankreatitis nicht zu dem Befund beitrug ($p=0.69$).

Dies mag – in Anbetracht des ansonsten robusten Beitrags der ADH Gene zur Alkoholabhängigkeit – auf den ersten Blick erstaunen.

Allerdings zeigte sich, dass die Häufigkeit Risikoallels des – im ursprünglichen Sample genomweit assoziierten SNPs rs1789891 – dort im Patientensample höher (19,4% vs. 17,9%) und in den Kontrollen niedriger (14,1% vs. 16,9%) war als in dem Folgesample mit somatischen Alkoholfolgekrankheiten. Der fehlende Assoziationsbefund im Folgesample könnte auf zufällige oder aber auf systematische Unterschiede zwischen a) den Patienten und/oder b) den Kontrollstichproben zurückzuführen sein:

a) Patientens Stichprobe:

- Für die Folgestichprobe aus ACP und ALZ-Patienten war keine explizite AA-Diagnose erforderlich. Daher kann sich die genetische Veranlagung unterscheiden. Diese Patienten wurden in Kliniken rekrutiert, die auf die Behandlung alkoholbedingter somatischer Erkrankungen spezialisiert sind. Bei all diesen Patienten war die jeweilige Erkrankung durch übermäßigen Alkoholkonsum bedingt, und die Mehrheit der Patienten konnte dennoch nicht auf Alkohol verzichten
- Die Unterschiede in der Genotypverteilung von rs1789891 in den Stichproben sind nicht zufällig: *ADH1B* wandelt Ethanol in Azetaldehyd um, dessen negative Effekte vom weiteren Alkohol-Konsum abhalten. Die gleichen Allele, die zu schnellerer Ethanolverstoffwechslung führen, können aber auch zu Gewebeschäden beitragen (Eriksson, 2001)

b) Kontrollstichprobe:

Zufällige Schwankungen als Ursache scheinen hier sehr wahrscheinlich, wenn man die Allelfrequenzen der verschiedenen Kontroll-Teilstichproben vergleicht. Bei näherer Betrachtung speziell der Allelfrequenzen von rs1789891 zeigt sich: Obwohl die Kontrollen für das Folgesample aus der gleichen Population stam-

men wie eine der Teilstichproben der ursprünglichen Kontrollen, liegt hier die Frequenz des Risikoallels von rs1789891 höher (16,9%) als in der entsprechenden Teilstichprobe, die im Ursprungssample benutzt wurde (13,9%)

Die Ergebnisse der Untersuchung zeigen also, dass bei Alkoholabhängigkeit selbst starke Befunde übersehen werden können, wenn die Stichproben klein und klinisch heterogen sind, zumal wenn gegenläufige Effekte zum Tragen kommen können. Da solche gegenläufigen Effekte bei der Entstehung von Alkoholabhängigkeit schon seit längerem vermutet werden (z.B. impulsiver (Rogers, Moeller, Swann, & Clark, 2010) vs. stress-coping Alkoholgenuss (Keyes, Hatzenbuehler, Grant, & Hasin, 2012). und diese Traits auch in der Kontrollbevölkerung in hoher Frequenz vorkommen), sind ausreichend große und gleichzeitig gut charakterisierte Patienten- und Kontrollstichproben für den GWAS Ansatz unabdingbar.

5.2.3 Publikation: Shared genetic etiology between alcohol dependence and major depressive disorder

Die Ergebnisse dieser Studie belegen unabhängig die Ergebnisse einer kürzlich veröffentlichten Studie von Andersen et al. (2017), nämlich, dass Risikogene für Depression auch das Risiko für Alkoholabhängigkeit erhöhen.

Das Problem solcher Untersuchungen liegt allerdings darin, wenn in dem Training und/oder Targetsample die Komorbidität nicht erfasst wurde. Weder das PGC-MDD1 noch das PGC-MDD2 Sample, auf welchen die MDD-Scores beruhen, wurde für Komorbidität für Alkoholabhängigkeit phänotypisiert.

Es ist daher nicht auszuschließen, dass in dem MDD Score auch Risikogene für Alkoholabhängigkeit enthalten sind, die auf Patienten mit komorbider Alkoholabhängigkeit zurückzuführen sind. Ähnlich kann es sich verhalten, wenn Patienten im Alkoholabhängigkeits-Target Sample an komorbider Depression leiden. Für letzteren Fall konnte in der vorliegenden Studie kontrolliert werden, da bei einer Teilstichprobe komorbide Depression explizit ausgeschlossen worden war, wie es auch Andersen et al. (2017) für Subgruppen berichtet hatten. Auch wenn für die andere Teilstichprobe eine komorbide Depression nicht mit Sicherheit ausgeschlossen werden konnte, so stand diese mit Sicherheit nicht im Vordergrund des klinischen Bildes, da dann die Diagnose einer Depression als Zweitdiagnose erstellt worden wäre.

Allerdings zeigte sich in der jetzigen Studie keine Verbesserung des Modells durch die Vergrößerung des Training Samples von $N=8\,148$ Patienten im PGC-MDD1 auf $N=59\,265$ im PGC-MDD2 Sample. Dies belegt wiederum die Notwendigkeit nicht nur genügend großer sondern auch homogener Stichproben für aussagekräftige Ergebnisse.

Eine weitere Einschränkung liegt darin begründet, dass für beide Phänotypen letztlich dieselbe Kontrollstichprobe benutzt wurde, sodass prinzipiell auch spezielle Beschaffenheit des Kontrollsamples die Assoziationen der polygenen Risikoscores mit beiden Phänotypen erklären könnte, ohne dass hier von Pleiotropie, d.h. bestimmte genetische Merkmale sind für mehrere Phänotypen ursächlich, auszugehen wäre, da die polygene Risikoscore-Analyse dies nicht trennen kann. Es gibt mittlerweile jedoch ein Verfahren, das es ermöglicht, genetische Korrelationen zwischen verschiedenen Störungen zu berechnen, ohne dass durch geteilte Kontrollstichproben eine Verzerrung auftritt, namentlich die LD-Score-Regression (Bulik-Sullivan et al., 2015).

Dieses Verfahren basiert auf Summenstatistiken und benötigt aufgrund der geringeren statistischen Effizienz deutlich größere Stichproben ($n > 5\,000$). In der gerade erschienen Meta-Analyse des PGC zu Alkoholabhängigkeit (Walters et al., 2018), in welche die in der vorliegenden Arbeit untersuchten Individuen mit eingegangen sind, konnte der hier erzielte Befund aber mit Hilfe der oben genannten Methode bestätigt werden: Die genetische Korrelation von Alkoholabhängigkeit und Depression liegt dort für Europäer bei 56%. Dies bewegt sich im oberen Bereich der genetischen Korrelationen von Alkoholabhängigkeit mit anderen psychiatrischen Störungen und liegt z.B. deutlich über der mit Schizophrenie (36%) oder Bipolarer Störung (19%).

Darüber hinausgehend kann durch spezielle Verfahren wie z.B. Buhmbox (Han et al., 2016) echte Pleiotropie abgegrenzt werden von heterogener Sample-Zusammenstellung. Mendelsche Randomisierung (Davey Smith & Ebrahim, 2003) erlaubt zusätzlich sogar die Bildung eines Modells zur kausalen Richtung des Zusammenhangs. Basierend auf den aktuellen Stichproben der Arbeitsgruppen zu Depression und Substanzgebrauchsstörungen des PGC wurde hierzu kürzlich eine

Arbeit eingereicht, die genau das versucht, und an der der Autor der vorliegenden Arbeit ebenfalls umfangreich mitgewirkt hat (Polimanti et al., 2018).

Zusammenfassend kann gesagt werden, dass die hier vorliegenden Studien belegen, dass GWAS prinzipiell geeignet sind, um bei psychiatrischen Störungen Gene und Pathways zu identifizieren. Darüber hinaus liefern genomweite Daten die Möglichkeit, Fragen hinsichtlich der genetischen Struktur von komplexen Störungen, genetischen Gemeinsamkeiten mit anderen Störungen und biologischer Subgruppierung mit beantworten zu helfen.

Die Ergebnisse zeigen, wie wichtig die Etablierung ausreichend großer aber gleichzeitig auch gut charakterisierter Stichproben ist. Die Bedeutung des Phänotyps darf hier keinesfalls unbeachtet bleiben. Dies ist aus verschiedenen Gründen wichtig:

1. Wie am Beispiel *ADH1B/ADH1C* dokumentiert können durch Sampleheterogenität selbst robuste Effekte maskiert werden, während entsprechend homogene Stichproben schon bei kleinerer Größe zu validen Aussagen kommen können.
2. Wenn das Trainingssample den vorherzusagenden Ziel-Phänotyp nicht adäquat repräsentiert, wird auch der Prädiktionsgewinn mit vergrößerten Trainingssamples insuffizient, wie am Beispiel des Therapieansprechens bei Schizophrenie zu sehen.
3. Wenn Informationen zur Komorbidität von verschiedenen Diagnosen nicht oder nur unzureichend erfasst sind, ist auch die Erkennung möglicher pleiotroper genetischer Effekte erschwert, wie im Beispiel von Alkoholabhängigkeit und Depression zu sehen. Dagegen kann eine entsprechend genaue Erfassung und Subgruppierung der Stichproben u.U. schon bei vergleichsweise kleiner Anzahl die Identifikation von solchen Zusammenhängen ermöglichen.

Um bei psychiatrisch-genetischen Untersuchungen das jeweils aktuelle genetische Wissen adäquat nutzen zu können ist – über die bloße Erhebung von DSM-Kategorien deutlich hinausgehend – eine detaillierte Erfassung von weiteren klinischen Parametern und speziell Endophänotypen, wie z.B. im Rahmen des Research Domain Criteria Frameworks, erforderlich.

In Anbetracht der heute zur Verfügung stehenden technologischen Möglichkeiten wie z.B. Ecological Momentary Assessments, kostengünstiger Genotypisie-

Diskussion

rungsmöglichkeiten und der großen Zahl von Patienten weltweit, ist dieser Ansatz sicherlich geeignet, um zum besseren Verständnis der Entstehung psychischer Störungen, sowie zu deren Therapie und Prävention wesentlich beizutragen.

6 ZUSAMMENFASSUNG

In den hier vorliegenden Arbeiten wurden genomweite Daten generiert und genutzt, um einen Beitrag zum besseren Verständnis der Beteiligung genetischer Faktoren in der Ätiologie von Alkoholabhängigkeit und Schizophrenie zu leisten. Hierbei wurden folgende Fragen gestellt:

Frage 1: Ist eine erhöhte genetische Belastung für Schizophrenie mit dem Therapieansprechen bzw. dem Krankheitsverlauf schizophrener Patienten assoziiert?

Für die Beantwortung dieser Frage wurde untersucht, ob der Krankheitsverlauf und das Therapieansprechen von schizophrenen Patienten eine Assoziation mit einem erhöhten polygenen Risikoscore für Schizophrenie zeigen, was der Fall war. Zudem konnte gezeigt werden, dass Non-Responder mit zusätzlichen bekannten klinischen Risiko-Faktoren, das heißt frühem und schleichendem Krankheitsbeginn die höchsten genetischen Risikoscores aufwiesen. (Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients)

Frage 2: Handelt es sich bei der Alkoholabhängigkeit ebenfalls um eine komplexe genetische Störung mit einem polygenen Anteil und tragen Genvarianten, die häufig in der Bevölkerung vorkommen zur Entstehung von Alkoholabhängigkeit bei?

Zur Beantwortung dieser Frage wurde eine genomweite Assoziationsuntersuchung zu Alkoholabhängigkeit in einer deutschen Fall-Kontroll-Stichprobe durchgeführt und basierend auf den Ergebnissen ein polygener Risikoscore gebildet. Dieser Score war in zwei unabhängigen Case-Control Stichproben mit Alkoholabhängigkeit assoziiert. Hiermit konnte zum ersten Mal auf molekularer Ebene gezeigt werden, dass häufige Varianten zur Ätiologie der Alkoholabhängigkeit beitragen und somit eine polygene Komponente der Alkoholabhängigkeit belegen. Dabei konnte eine häufig vorkommende Risikovariante identifiziert werden, die im Alkoholdehydrogenase-Gen-Cluster lokalisiert ist, einer Gruppe von Genen, welche für die Verstoffwechslung von Alkohol verantwortlich sind. (*Genome-Wide Significant Association between Alcohol Dependence and a Variant in the ADH Gene Cluster*).

Frage 3: Ist eine Vergrößerung der Stichprobe auch auf Kosten der klinischen Homogenität zielführend, um weitere Gene und Pathways zu finden?

Zur Beantwortung dieser Frage wurde die Größe der Case-Control Stichprobe verdoppelt, allerdings unter Einbezug von Patienten mit alkoholinduzierter Leberzirrhose/Pankreatitis. Die Studie belegt an Hand der Untersuchung einer durch viele Kandidatengen- und Tierstudien gesicherten Assoziation des *Alkoholdehydrogenase*-Genclusters mit Alkoholabhängigkeit, dass eine Erhöhung der Versuchspersonenzahl unter Hinzunahme von klinisch heterogenen Samples nicht automatisch zu einer Zunahme der Teststärke, sondern sogar zu einem Verlust der Power durch gegenläufige Effekte führen kann, und verweist auf die Bedeutung sorgfältiger Charakterisierung der Patienten als auch Kontrollsamples. (*Genetic Contribution to Alcohol Dependence: Investigation of a Heterogeneous German Sample of Individuals with Alcohol Dependence, Chronic Alcoholic Pancreatitis, and Alcohol-Related Cirrhosis*)

Frage 4: Zwischen Alkoholabhängigkeit und Depression ist auf klinischer Ebene eine Komorbidität belegt. Hier soll die Frage beantwortet werden, ob Risikovarianten für Depression auch das Risiko für Alkoholabhängigkeit erhöhen.

Zur Beantwortung dieser Frage wurde ein polygener Risikoscore für Depression berechnet. Es konnte gezeigt werden, dass dieser Score signifikant mit Alkoholabhängigkeit assoziiert ist. Dieser Befund wurde auch in einer Subgruppe alkoholabhängiger Patienten gefunden, bei denen eine Komorbidität mit Depression sicher ausgeschlossen werden konnte. (*Shared genetic etiology between alcohol dependence and major depressive disorder*)

Zusammenfassend kann gesagt werden: Die Arbeiten belegen, dass die Nutzung genomweiter Daten hilfreich ist, um neue Erkenntnisse hinsichtlich genetischer Risikofaktoren zu gewinnen, Einblicke in die Ursachen der Komorbidität psychischer Störungen zu erzielen, sowie die Ursachen klinischer Unterschiede auf molekularer Ebene zu untersuchen, wie hier am Beispiel unterschiedlichen Therapieansprechens. Die Studienergebnisse zeigen auch deutlich, dass das Potential des Ansatzes nur ausgeschöpft werden kann durch ausreichend große aber gleichzeitig auch gut charakterisierte Stichproben.

7 LITERATURVERZEICHNIS

- Andersen, A. M., Pietrzak, R. H., Kranzler, H. R., Ma, L., Zhou, H., Liu, X., ... Han, S. (2017). Polygenic Scores for Major Depressive Disorder and Risk of Alcohol Dependence. *JAMA Psychiatry*.
<https://doi.org/10.1001/jamapsychiatry.2017.2269>
- Ball, D. (2008). Addiction science and its genetics. *Addiction (Abingdon, England)*, *103*(3), 360–367. <https://doi.org/10.1111/j.1360-0443.2007.02061.x>
- Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2004). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, *21*(2), 263–265. <https://doi.org/10.1093/bioinformatics/bth457>
- Baum, A. E., Hamshere, M., Green, E., Cichon, S., Rietschel, M., Noethen, M. M., ... McMahon, F. J. (2008). Meta-analysis of two genome-wide association studies of bipolar disorder reveals important points of agreement. *Molecular Psychiatry*, *13*(5), 466–467. <https://doi.org/10.1038/mp.2008.16>
- Beesdo-Baum, K., Knappe, S., Asselmann, E., Zimmermann, P., Brückl, T., Höfler, M., ... Wittchen, H.-U. (2015). The “Early Developmental Stages of Psychopathology (EDSP) study”: a 20-year review of methods and findings. *Social Psychiatry and Psychiatric Epidemiology*, *50*(6), 851–866. <https://doi.org/10.1007/s00127-015-1062-x>
- Bigdeli, T. B., Ripke, S., Peterson, R. E., Trzaskowski, M., Bacanu, S.-A., Abdellaoui, A., ... Kendler, K. S. (2017). Genetic effects influencing risk for major depressive disorder in China and Europe. *Translational Psychiatry*, *7*(3), e1074. <https://doi.org/10.1038/tp.2016.292>
- Birley, A. J., James, M. R., Dickson, P. A., Montgomery, G. W., Heath, A. C., Martin, N. G., & Whitfield, J. B. (2009). ADH single nucleotide polymorphism associations with alcohol metabolism in vivo. *Human Molecular Genetics*, *18*(8), 1533–1542. <https://doi.org/10.1093/hmg/ddp060>

- Black, J. R. M., & Clark, S. J. (2016). Age-related macular degeneration: genome-wide association studies to translation. *Genetics in Medicine*, *18*(4), 283–289. <https://doi.org/10.1038/gim.2015.70>
- Buckley, P. F., Miller, B. J., Lehrer, D. S., & Castle, D. J. (2009). Psychiatric comorbidities and schizophrenia. *Schizophrenia Bulletin*, *35*(2), 383–402. <https://doi.org/10.1093/schbul/sbn135>
- Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., ... Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*. <https://doi.org/10.1038/ng.3211>
- Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., ... Worthington, J. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, *447*(7145), 661–678. <https://doi.org/10.1038/nature05911>
- Cardno, A. G., Marshall, E. J., Coid, B., Macdonald, A. M., Ribchester, T. R., Davies, N. J., ... Murray, R. M. (1999). Heritability Estimates for Psychotic Disorders: The Maudsley Twin Psychosis Series. *Archives of General Psychiatry*, *56*(2), 162. <https://doi.org/10.1001/archpsyc.56.2.162>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1). <https://doi.org/10.1186/s13742-015-0047-8>
- Clarke, T.-K., Adams, M. J., Davies, G., Howard, D. M., Hall, L. S., Padmanabhan, S., ... McIntosh, A. M. (2017). Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N=112 117). *Molecular Psychiatry*, *22*, 1376.
- Connor, J. P., Haber, P. S., & Hall, W. D. (2016). Alcohol use disorders. *The Lancet*, *387*(10022), 988–998.
- Crespo-Facorro, B., de la Foz, V. O.-G., Ayesa-Arriola, R., Pérez-Iglesias, R., Mata, I., Suarez-Pinilla, P., ... Vázquez-Barquero, J. L. (2013). Prediction of acute

- clinical response following a first episode of non affective psychosis: results of a cohort of 375 patients from the Spanish PAFIP study. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 44, 162–167. <https://doi.org/10.1016/j.pnpbp.2013.02.009>
- Crum, R. M., Green, K. M., Storr, C. L., Chan, Y.-F., Ialongo, N., Stuart, E. A., & Anthony, J. C. (2008). Depressed Mood in Childhood and Subsequent Alcohol Use Through Adolescence and Young Adulthood. *Archives of General Psychiatry*, 65(6), 702. <https://doi.org/10.1001/archpsyc.65.6.702>
- Davey Smith, G., & Ebrahim, S. (2003). ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease?*. *International Journal of Epidemiology*, 32(1), 1–22. <https://doi.org/10.1093/ije/dyg070>
- de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology*, 11(4), e1004219. <https://doi.org/10.1371/journal.pcbi.1004219>
- Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., ... Neale, B. M. (2018). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics*. <https://doi.org/10.1038/s41588-018-0269-7>
- Edenberg, H. J. (2007). The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Research & Health: The Journal of the National Institute on Alcohol Abuse and Alcoholism*, 30(1), 5–13.
- Endicott, J. (1978). A Diagnostic Interview: The Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry*, 35(7), 837. <https://doi.org/10.1001/archpsyc.1978.01770310043002>
- Eng, M. Y., Luczak, S. E., & Wall, T. L. (2007). ALDH2, ADH1B, and ADH1C genotypes in Asians: a literature review. *Alcohol Research & Health: The Journal of the National Institute on Alcohol Abuse and Alcoholism*, 30(1), 22–27.

- Enoch, M.-A., & Goldman, D. (2002). Problem drinking and alcoholism: diagnosis and treatment. *American Family Physician*, 65(3), 441–448.
- Enomoto, N., Takase, S., Yasuhara, M., & Takada, A. (1991). Acetaldehyde metabolism in different aldehyde dehydrogenase-2 genotypes. *Alcoholism, Clinical and Experimental Research*, 15(1), 141–144.
- Erbel, R., Eisele, L., Moebus, S., Dragano, N., Möhlenkamp, S., Bauer, M., ... Jöckel, K.-H. (2012). Die Heinz Nixdorf Recall Studie. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 55(6–7), 809–815. <https://doi.org/10.1007/s00103-012-1490-7>
- Eriksson, C. J. (2001). The role of acetaldehyde in the actions of alcohol (update 2000). *Alcoholism, Clinical and Experimental Research*, 25(5 Suppl ISBRA), 15S–32S.
- Essali, A., Al-Haj Haasan, N., Li, C., & Rathbone, J. (2009). Clozapine versus typical neuroleptic medication for schizophrenia. In The Cochrane Collaboration & A. Essali (Eds.), *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd. Retrieved from <http://doi.wiley.com/10.1002/14651858.CD000059.pub2>
- Euesden, J., Lewis, C. M., & O'Reilly, P. F. (2014). PRSice: Polygenic Risk Score software. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btu848>
- Fangerau, H., Ohlraun, S., Granath, R. O., Nöthen, M. M., Rietschel, M., & Schulze, T. G. (2004). Computer-Assisted Phenotype Characterization for Genetic Research in Psychiatry. *Human Heredity*, 58(3–4), 122–130. <https://doi.org/10.1159/000083538>
- First, M. B., New York State Psychiatric Institute, & Biometrics Research Department. (1997). *Structured clinical interview for DSM-IV axis I disorders: SCID-I*. Washington, DC; New York, N.Y.: American Psychiatric Press; Biometrics Research Dept., New York State Psychiatric Institute.
- Foulds, J. A., Adamson, S. J., Boden, J. M., Williman, J. A., & Mulder, R. T. (2015). Depression in patients with alcohol use disorders: Systematic review and me-

- ta-analysis of outcomes for independent and substance-induced disorders. *Journal of Affective Disorders*, 185, 47–59. <https://doi.org/10.1016/j.jad.2015.06.024>
- Frazer, K. A., Murray, S. S., Schork, N. J., & Topol, E. J. (2009). Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics*, 10(4), 241–251. <https://doi.org/10.1038/nrg2554>
- Gaebel, W., & Wölwer, W. (2010). *Gesundheitsberichterstattung des Bundes Heft 50*. Berlin: Robert Koch Institut.
- Gelernter, J., Kranzler, H. R., Sherva, R., Almasy, L., Koesterer, R., Smith, A. H., ... others. (2014). Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Molecular Psychiatry*, 19(1), 41–49.
- Gillespie, A. L., Samanaite, R., Mill, J., Egerton, A., & MacCabe, J. H. (2017). Is treatment-resistant schizophrenia categorically distinct from treatment-responsive schizophrenia? a systematic review. *BMC Psychiatry*, 17(1). <https://doi.org/10.1186/s12888-016-1177-y>
- Goldstein, D. B. (2009). Common genetic variation and human traits. *New England Journal of Medicine*, 360(17), 1696.
- Grant, J. D., Agrawal, A., Bucholz, K. K., Madden, P. A. F., Pergadia, M. L., Nelson, E. C., ... Heath, A. C. (2009). Alcohol Consumption Indices of Genetic Risk for Alcohol Dependence. *Biological Psychiatry*, 66(8), 795–800. <https://doi.org/10.1016/j.biopsych.2009.05.018>
- Han, B., Pouget, J. G., Slowikowski, K., Stahl, E., Lee, C. H., Diogo, D., ... Raychaudhuri, S. (2016). A method to decipher pleiotropy by detecting underlying heterogeneity driven by hidden subgroups applied to autoimmune and neuropsychiatric diseases. *Nature Genetics*, 48(7), 803–810. <https://doi.org/10.1038/ng.3572>

- Hay, S. I. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016.
- Hilker, R., Helenius, D., Fagerlund, B., Skytthe, A., Christensen, K., Werge, T. M., ... Glenthøj, B. (2017). Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. *Biological Psychiatry*. <https://doi.org/10.1016/j.biopsych.2017.08.017>
- Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics*, 6(2), 95–108. <https://doi.org/10.1038/nrg1521>
- Holle, R., Happich, M., Löwel, H., & Wichmann, H. (2005). KORA - A Research Platform for Population Based Health Research. *Das Gesundheitswesen*, 67(S 01), 19–25. <https://doi.org/10.1055/s-2005-858235>
- Kendler, K. S., Walters, E. E., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1995). The structure of the genetic and environmental risk factors for six major psychiatric disorders in women. Phobia, generalized anxiety disorder, panic disorder, bulimia, major depression, and alcoholism. *Archives of General Psychiatry*, 52(5), 374–383.
- Kessler, R. C., Nelson, C. B., McGonagle, K. A., Edlund, M. J., Frank, R. G., & Leaf, P. J. (1996). The epidemiology of co-occurring addictive and mental disorders: implications for prevention and service utilization. *The American Journal of Orthopsychiatry*, 66(1), 17–31.
- Keyes, K. M., Hatzenbuehler, M. L., Grant, B. F., & Hasin, D. S. (2012). Stress and alcohol: epidemiologic evidence. *Alcohol Research: Current Reviews*, 34(4), 391–400.
- Lango Allen, H., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., ... Hirschhorn, J. N. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, 467, 832.

- Laursen, T. M., Agerbo, E., & Pedersen, C. B. (2009). Bipolar disorder, schizoaffective disorder, and schizophrenia overlap: a new comorbidity index. *The Journal of Clinical Psychiatry*, *70*(10), 1432–1438. <https://doi.org/10.4088/JCP.08m04807>
- Li, D., Zhao, H., & Gelernter, J. (2011). Strong Association of the Alcohol Dehydrogenase 1B Gene (ADH1B) with Alcohol Dependence and Alcohol-Induced Medical Diseases. *Biological Psychiatry*. <https://doi.org/10.1016/j.biopsych.2011.02.024>
- Liu, J. Z., Mcrae, A. F., Nyholt, D. R., Medland, S. E., Wray, N. R., Brown, K. M., ... Martin, N. G. (2010). A Versatile Gene-Based Test for Genome-wide Association Studies. *The American Journal of Human Genetics*, *87*(1), 139–145. <https://doi.org/10.1016/j.ajhg.2010.06.009>
- Lucae, S., Salyakina, D., Barden, N., Harvey, M., Gagné, B., Labbé, M., ... Müller-Myhsok, B. (2006). P2RX7, a gene coding for a purinergic ligand-gated ion channel, is associated with major depressive disorder. *Human Molecular Genetics*, *15*(16), 2438–2445. <https://doi.org/10.1093/hmg/ddl166>
- Lynskey, M. T. (1998). The comorbidity of alcohol dependence and affective disorders: treatment implications. *Drug and Alcohol Dependence*, *52*(3), 201–209.
- Lyons, M. J., Schultz, M., Neale, M., Brady, K., Eisen, S., Toomey, R., ... Tsuang, M. (2006). Specificity of familial vulnerability for alcoholism versus major depression in men. *The Journal of Nervous and Mental Disease*, *194*(11), 809–817. <https://doi.org/10.1097/01.nmd.0000244480.78431.49>
- Macgregor, S., Lind, P. A., Bucholz, K. K., Hansell, N. K., Madden, P. A. F., Richter, M. M., ... Whitfield, J. B. (2008). Associations of ADH and ALDH2 gene variation with self report alcohol reactions, consumption and dependence: an integrated analysis. *Human Molecular Genetics*, *18*(3), 580–593. <https://doi.org/10.1093/hmg/ddn372>
- Maier, W., Lichtermann, D., & Mingos, J. (1994). The relationship between alcoholism and unipolar depression--a controlled family study. *Journal of Psychiatric Research*, *28*(3), 303–317.

- Mailman, M. D., Feolo, M., Jin, Y., Kimura, M., Tryka, K., Bagoutdinov, R., ... Sherry, S. T. (2007). The NCBI dbGaP database of genotypes and phenotypes. *Nature Genetics*, 39(10). <https://doi.org/10.1038/ng1007-1181>
- Marshall, M., Lewis, S., Lockwood, A., Drake, R., Jones, P., & Croudace, T. (2005). Association between duration of untreated psychosis and outcome in cohorts of first-episode patients: a systematic review. *Archives of General Psychiatry*, 62(9), 975–983. <https://doi.org/10.1001/archpsyc.62.9.975>
- McGuffin, P. (1991). A Polydiagnostic Application of Operational Criteria in Studies of Psychotic Illness. Development and Reliability of the OPCRIT System. *Archives of General Psychiatry*, 48(8), 764. <https://doi.org/10.1001/archpsyc.1991.01810320088015>
- McGuffin, P., Rijdsdijk, F., Andrew, M., Sham, P., Katz, R., & Cardno, A. (2003). The Heritability of Bipolar Affective Disorder and the Genetic Relationship to Unipolar Depression. *Archives of General Psychiatry*, 60(5), 497. <https://doi.org/10.1001/archpsyc.60.5.497>
- Meyer, C., Rumpf, H.-J., Hapke, U., Dilling, H., & John, U. (2000). Prevalence of alcohol consumption, abuse and dependence in a country with high per capita consumption: findings from the German TACOS study. *Social Psychiatry and Psychiatric Epidemiology*, 35(12), 539–547. <https://doi.org/10.1007/s001270050277>
- Nielsen, J., Nielsen, R. E., & Correll, C. U. (2012). Predictors of Clozapine Response in Patients With Treatment-Refractory Schizophrenia: Results From a Danish Register Study. *Journal of Clinical Psychopharmacology*, 32(5), 678–683. <https://doi.org/10.1097/JCP.0b013e318267b3cd>
- Nöthen, M. M., Rietschel, M., Erdmann, J., Oberländer, H., Möller, H.-J., Naber, D., & Propping, P. (1995). Genetic variation of the 5-HT_{2A} receptor and response to clozapine. *The Lancet*, 346(8979), 908–909. [https://doi.org/10.1016/S0140-6736\(95\)92756-5](https://doi.org/10.1016/S0140-6736(95)92756-5)
- Nöthlings, U., & Krawczak, M. (2012). PopGen: Eine populationsbasierte Biobank mit Langzeitverfolgung der Kontrollkohorte. *Bundesgesundheitsblatt - Gesundheits-*

- heitsforschung* - *Gesundheitsschutz*, 55(6–7), 831–835.
<https://doi.org/10.1007/s00103-012-1487-2>
- Nurnberger, J. I., Foroud, T., Flury, L., Meyer, E. T., & Wiegand, R. (2002). Is there a genetic relationship between alcoholism and depression? *Alcohol Research & Health: The Journal of the National Institute on Alcohol Abuse and Alcoholism*, 26(3), 233–240.
- O'Donovan, M. C., Craddock, N., Norton, N., Williams, H., Peirce, T., Moskvina, V., ... Owen, M. J. (2008). Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nature Genetics*, 40(9), 1053–1055.
<https://doi.org/10.1038/ng.201>
- Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., ... CRESTAR Consortium. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, 50(3), 381–389. <https://doi.org/10.1038/s41588-018-0059-2>
- Park, B. L., Kim, J. W., Cheong, H. S., Kim, L. H., Lee, B. C., Seo, C. H., ... Choi, I.-G. (2013). Extended genetic effects of ADH cluster genes on the risk of alcohol dependence: from GWAS to replication. *Human Genetics*, 132(6), 657–668. <https://doi.org/10.1007/s00439-013-1281-8>
- Patterson, N., Price, A. L., & Reich, D. (2006). Population Structure and Eigenanalysis. *PLoS Genetics*, 2(12), e190. <https://doi.org/10.1371/journal.pgen.0020190>
- Pe'er, I., Yelensky, R., Altshuler, D., & Daly, M. J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic Epidemiology*, 32(4), 381–385.
<https://doi.org/10.1002/gepi.20303>
- Polimanti, R., Peterson, R. E., Ong, J. S., Macgregor, S., Edwards, A., Clarke, T.-K., ... Derks, E. M. (2018). Evidence of causal effect of major depression on alcohol dependence: Findings from the Psychiatric Genomics Consortium. <https://doi.org/10.1101/412098>

- Power, R. A., Kyaga, S., Uher, R., MacCabe, J. H., Långström, N., Landen, M., ... Svensson, A. C. (2013). Fecundity of Patients With Schizophrenia, Autism, Bipolar Disorder, Depression, Anorexia Nervosa, or Substance Abuse vs Their Unaffected Siblings. *JAMA Psychiatry*, 70(1), 22. <https://doi.org/10.1001/jamapsychiatry.2013.268>
- Prescott, C. A., Aggen, S. H., & Kendler, K. S. (2000). Sex-specific genetic influences on the comorbidity of alcoholism and major depression in a population-based sample of US twins. *Archives of General Psychiatry*, 57(8), 803–811.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., ... Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- Purcell, S., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., & Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, 460(7256), 748–752. <https://doi.org/10.1038/nature08185>
- R Development Core Team. (2010). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Retrieved from <http://www.R-project.org>
- Regier, D. A. (1990). Comorbidity of Mental Disorders With Alcohol and Other Drug Abuse: Results From the Epidemiologic Catchment Area (ECA) Study. *JAMA*, 264(19), 2511. <https://doi.org/10.1001/jama.1990.03450190043026>
- Rietschel, M., Mattheisen, M., Degenhardt, F., Kahn, R. S., Linszen, D. H., Os, J. van, ... Cichon, S. (2011). Association between genetic variation in a region on chromosome 11 and schizophrenia in large samples from Europe. *Molecular Psychiatry*, 17(9), 906–917. <https://doi.org/10.1038/mp.2011.80>
- Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K.-H., Holmans, P. A., ... O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, 511(7510), 421–427. <https://doi.org/10.1038/nature13595>

- Ripke, S., Sanders, A. R., Kendler, K. S., Levinson, D. F., Sklar, P., Holmans, P. A., ... Gejman, P. V. (2011). Genome-wide association study identifies five new schizophrenia loci. *Nature Genetics*, 43(10), 969–976. <https://doi.org/10.1038/ng.940>
- Ripke, S., Wray, N. R., Lewis, C. M., Hamilton, S. P., Weissman, M. M., Breen, G., ... Cichon, S. (2013). A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*, 18(4), 497.
- Rogers, R. D., Moeller, F. G., Swann, A. C., & Clark, L. (2010). Recent Research on Impulsivity in Individuals With Drug Use and Mental Health Disorders: Implications for Alcoholism: IMPULSIVITY, MOOD, AND SUBSTANCE MISUSE DISORDERS. *Alcoholism: Clinical and Experimental Research*, no-no. <https://doi.org/10.1111/j.1530-0277.2010.01216.x>
- Schennach, R., Riedel, M., Musil, R., & Möller, H.-J. (2012). Treatment Response in First-episode Schizophrenia. *Clinical Psychopharmacology and Neuroscience*, 10(2), 78. <https://doi.org/10.9758/cpn.2012.10.2.78>
- Schoepf, D., & Heun, R. (2015). Alcohol dependence and physical comorbidity: Increased prevalence but reduced relevance of individual comorbidities for hospital-based mortality during a 12.5-year observation period in general hospital admissions in urban North-West England. *European Psychiatry*, 30(4), 459–468. <https://doi.org/10.1016/j.eurpsy.2015.03.001>
- Spencer, C. C., Su, Z., Donnelly, P., & Marchini, J. (2009). Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genetics*, 5(5), e1000477.
- Stahl, E., Forstner, A., McQuillin, A., Ripke, S., Ophoff, R., Scott, L., ... Breen, G. (2017). Genomewide association study identifies 30 loci associated with bipolar disorder. <https://doi.org/10.1101/173062>
- Sullivan, P. F. (2012). Don't give up on GWAS. *Molecular Psychiatry*, 17(1), 2.

- Sullivan, P. F., Daly, M. J., & O'Donovan, M. (2012). Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature Reviews Genetics*, *13*(8), 537–551. <https://doi.org/10.1038/nrg3240>
- Swendsen, J. D., & Merikangas, K. R. (2000). The comorbidity of depression and substance use disorders. *Clinical Psychology Review*, *20*(2), 173–189.
- Tawa, E. A., Hall, S. D., & Lohoff, F. W. (2016). Overview of the Genetics of Alcohol Use Disorder. *Alcohol and Alcoholism*, *51*(5), 507–514. <https://doi.org/10.1093/alcalc/agw046>
- Taylor, D. M., Young, C., & Paton, C. (2003). Prior antipsychotic prescribing in patients currently receiving clozapine: a case note review. *The Journal of Clinical Psychiatry*, *64*(1), 30–34.
- The NIDDK IBD Genetics Consortium, McGovern, D. P. B., Gardet, A., Törkvist, L., Goyette, P., Essers, J., ... Seielstad, M. (2010). Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nature Genetics*, *42*(4), 332–337. <https://doi.org/10.1038/ng.549>
- Treutlein, J., Cichon, S., Ridinger, M., Wodarz, N., Soyka, M., Zill, P., ... Rietschel, M. (2009). Genome-wide association study of alcohol dependence. *Archives of General Psychiatry*, *66*(7), 773–784. <https://doi.org/10.1001/archgenpsychiatry.2009.83>
- Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet Psychiatry*, *3*(2), 171–178.
- Walters, R. K., Polimanti, R., Johnson, E. C., McClintick, J. N., Adams, M. J., Adkins, A. E., ... Agrawal, A. (2018). Trans-ancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nature Neuroscience*.
- Wang, J.-C., Foroud, T., Hinrichs, A. L., Le, N. X. H., Bertelsen, S., Budde, J. P., ... Goate, A. M. (2012). A genome-wide association study of alcohol-dependence symptom counts in extended pedigrees identifies C15orf53. *Molecular Psychiatry*, *18*, 1218.

- Way, M., McQuillin, A., Saini, J., Ruparelia, K., Lydall, G. J., Guerrini, I., ... Gurling, H. M. D. (2015). Genetic variants in or near *ADH 1 B* and *ADH 1 C* affect susceptibility to alcohol dependence in a British and Irish population: ADH SNPs protect against ADS. *Addiction Biology*, 20(3), 594–604. <https://doi.org/10.1111/adb.12141>
- Winokur, G., & Coryell, W. (1991). Familial alcoholism in primary unipolar major depressive disorder. *The American Journal of Psychiatry*, 148(2), 184–188. <https://doi.org/10.1176/ajp.148.2.184>
- Woodruff, R. A., Guze, S. B., Clayton, P. J., & Carr, D. (1973). Alcoholism and depression. *Archives of General Psychiatry*, 28(1), 97–100.
- World Health Organization. (2009). *Global health risks: Mortality and burden of disease attributable to selected major risks*. Geneva: WHO.
- Wray, N. R., Purcell, S. M., & Visscher, P. M. (2011). Synthetic Associations Created by Rare Variants Do Not Explain Most GWAS Results. *PLoS Biology*, 9(1), e1000579. <https://doi.org/10.1371/journal.pbio.1000579>
- Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., ... the Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, 50(5), 668–681. <https://doi.org/10.1038/s41588-018-0090-3>
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., ... Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42(7), 565–569. <https://doi.org/10.1038/ng.608>
- Zondervan, K. T., & Cardon, L. R. (2007). Designing candidate gene and genome-wide case–control association studies. *Nature Protocols*, 2(10), 2492–2501. <https://doi.org/10.1038/nprot.2007.366>

8 ENTHALTENE EIGENE VERÖFFENTLICHUNGEN

Frank, J., Lang, M., Witt, S. H., Strohmaier, J., Rujescu, D., Cichon, S., ... others. (2014). Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients. *Molecular Psychiatry*. Retrieved from <http://www.nature.com/articles/mp201456>

Frank, J., Lang, M., Witt, S. H., Strohmaier, J., Rujescu, D., Cichon, S., ... Rietschel, M. (2015). Corrigendum: Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients. *Molecular Psychiatry*. <https://doi.org/10.1038/mp.2015.52>

Frank, J., Cichon, S., Treutlein, J., Ridinger, M., Mattheisen, M., Hoffmann, P., ... Rietschel, M. (2012). Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster. *Addiction Biology*, 17(1), 171–180. <https://doi.org/10.1111/j.1369-1600.2011.00395.x>

Treutlein, J.*, Frank, J.*, Streit, F., Reinbold, C., Juraeva, D., Degenhardt, F., ... Rietschel, M. (2017). Genetic Contribution to Alcohol Dependence: Investigation of a Heterogeneous German Sample of Individuals with Alcohol Dependence, Chronic Alcoholic Pancreatitis, and Alcohol-Related Cirrhosis. *Genes*, 8(7), 183. <https://doi.org/10.3390/genes8070183>

Foo, J. C.*, Streit, F.*, Treutlein, J., Ripke, S., Witt, S. H., Strohmaier, J., ... Frank, J. (2018). Shared genetic etiology between alcohol dependence and major depressive disorder: *Psychiatric Genetics*, 1. <https://doi.org/10.1097/YPG.0000000000000201>

*Autoren haben in gleichem Umfang zur Publikation beigetragen

9 LEBENS LAUF

PERSONALIEN

Name und Vorname: Frank, Josef

Geburtsdatum: 7.2.1972

Geburtsort: Riedenburg

Familienstand: Verheiratet

SCHULISCHER WERDEGANG

1983 – 1992 Albertus-Magnus-Gymnasium Regensburg

8.7.1992 Abitur

UNIVERSITÄRER WERDEGANG

WS1998/99	Beginn des Studiums Psychologie an der Universität Trier
30.4.2000	Vordiplom
2000 – 2003	Hauptstudium
2003	Diplomarbeit Geschlechtsspezifität emotionsbedingten Essverhaltens
30.10.2003	Diplom, Note: 1,7
Seit Februar 2005	Wissenschaftlicher Mitarbeiter an der Abteilung Genetische Epidemiologie in der Psychiatrie des ZI-Mannheim unter Leitung von Frau Prof. Dr. med. Marcella Rietschel

10 DANKSAGUNG

Mein besonderer Dank gilt Frau Prof. Dr. med. Marcella Rietschel für die Überlassung des Themas dieser Arbeit und die ausgezeichneten Möglichkeiten, es zu bearbeiten, sowie für ihre geduldige Betreuung und Hilfe bei meinen vielen Fragen.

Ebenso danke ich Herrn Dr. Jens Treutlein, der mir zum Verständnis vieler Quellen verhalf und mir wichtige sachliche Hinweise insbesondere zu molekulargenetischen Begrifflichkeiten gab.

Weiter danke ich ganz herzlich Herrn Stefan Herms, der mir geduldig bei all meinen Fragen zur Verarbeitung von Genomestudio-Daten weiterhalf.

Ich danke darüber hinaus dem ganzen Team der Abteilung genetische Epidemiologie in der Psychiatrie sowie der Abteilung Biostatistik des ZI-Mannheim, insbesondere Herrn Prof. Dr. Stefan Wellek, der mich nie ohne fundierte Antwort auf meine Fragen gelassen hat.

Mein Dank gilt insbesondere auch den Teilnehmern und Teilnehmerinnen an den vorliegenden Studien, ohne deren Beitrag diese Arbeit nicht möglich gewesen wäre.

Schließlich bedanke ich mich bei meiner Frau und meinen Kindern für ihre Geduld, wenn es bei der Erstellung dieser Arbeit und der zugehörigen Artikelmanuskripte wieder sehr spät wurde.