

# **Molekulargenetische Analyse seltener, kodierender Sequenzvarianten innerhalb eines X-chromosomalen Risikolocus für die Schizophrenie**

Dissertation

zur Erlangung des Doktorgrades (Dr. med.)

der Medizinischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

**Isabelle Patrizia Claus**

aus Moers

2025

Angefertigt mit der Genehmigung  
der Medizinischen Fakultät der Universität Bonn

1. Gutachter: Prof. Dr. med. Markus M. Nöthen
2. Gutachter: Prof. Dr. Glen Kristiansen

Tag der mündlichen Prüfung: 28.05.2025

Aus dem Institut für Humangenetik

Für meine Familie



## **Inhaltsverzeichnis**

<b>Abkürzungsverzeichnis</b>	6
<b>1. Deutsche Zusammenfassung</b>	8
1.1 Einleitung	8
1.2 Material und Methoden	16
1.3 Ergebnisse	20
1.4 Diskussion	23
1.5 Zusammenfassung	30
1.6 Literaturverzeichnis der deutschen Zusammenfassung	31
<b>2. Veröffentlichung</b>	42
<b>3. Erklärung zum Eigenanteil</b>	43
<b>4. Danksagung</b>	44

## Abkürzungsverzeichnis

BRCC3	BRCA1/BRCA2-Containing Complex Subunit 3
CADD	Combined Annotation Dependent Depletion Score
CLIC2	Chloride Intracellular Channel 2
CMC4	C-X9-C Motif-Containing Protein 4
CNV	Kopienzahlvariante oder Copy Number Variant
CUBA	Core Unit for Bioinformatics Analysis
DNA	Desoxyribonukleinsäure
ExAC	Exome Aggregation Consortium
F8	Coagulation Factor 8
FUNDC2	FUN14 Domain-Containing Protein 2
GATK	Genome Analysis Toolkit
GWAS	Genomweite Assoziationsstudie(n) oder Genome-Wide Association Study
InDels	kleinere Insertionen und Deletionen
MAF	Minor-Allelfrequenz oder Minor Allele Frequency
MTCP1	Mature T-Cell Proliferation 1
NGS	Next Generation Sequencing
OR	Odds Ratio
PCR	Polymerasekettenreaktion oder Polymerase-Chain-Reaction
PGC	Psychiatric Genomics Consortium
RAB39	Ras-Related Protein Rab-39
SCHEMA	Schizophrenia Exome Sequencing Meta-Analysis Consortium
SCZ	Schizophrenie

SKAT-O	Optimal Unified Sequence Kernel Association Test
smMIP	single-molecule Molecular Inversion Probe
SNP	Einzelnukleotidpolymorphismus oder Single Nucleotide Polymorphism
SNV	Einzelnukleotidvariante oder Single Nucleotide Variant
TAD	Topologically Associated Domain
VBP1	Von Hippel-Lindau Binding Protein 1

# 1. Deutsche Zusammenfassung

## 1.1 Einleitung

### 1.1.1 Schizophrenie als schwerwiegende psychiatrische Erkrankung

Die Schizophrenie (SCZ) ist eine der schwerwiegendsten neuropsychiatrischen Erkrankungen (Kahn et al. 2015; Owen et al. 2016; Jauhar et al. 2022) und tritt mit einer Lebenszeitprävalenz von etwas unter 1 % (McGrath et al. 2008) in der Allgemeinbevölkerung auf. Weltweit ist die SCZ mit hohen persönlichen und gesellschaftlichen Kosten verbunden (Solmi et al. 2023). So verlieren beispielsweise Betroffene mit einer SCZ durchschnittlich 10-20 Jahre ihres Lebens durch die Krankheit (Hjorthøj et al. 2017), welches sich nicht ausschließlich durch die auf 10 % erhöhte Suizidrate erklären lässt (Plana-Ripoll et al. 2019). Betroffene leben in der Regel viele Jahre mit dieser beeinträchtigenden Erkrankung. Dies resultiert aus dem Beginn im jungen Erwachsenenalter und dem häufig chronischen Verlauf (Jauhar et al. 2022). Zudem blieben in der Vergangenheit trotz jahrzehntelanger intensiver Forschung maßgebliche therapeutische Fortschritte in der Psychiatrie aus (Tricklebank et al. 2021). Die Gruppe der antipsychotischen Medikamente bildet nach wie vor einen Grundpfeiler der Therapie bei der SCZ, obwohl sie wenig wirksam in Hinblick auf bestimmte Kategorien der klinischen Symptomatik und mit einer Vielzahl von Nebenwirkungen, insbesondere kardiovaskulären, verbunden sind (Kahn et al. 2015; Owen et al. 2016). Daher bringt die Diagnose SCZ selbst für Betroffene mit einem günstigeren Krankheitsverlauf in der Regel lebenslange Konsequenzen (Jauhar et al. 2022) mit sich, was u.a. durch eine Arbeitslosenquote unter Betroffenen mit einer SCZ von 70-90 % verdeutlicht wird (Marwaha et al. 2007).

Trotz aller wissenschaftlicher Anstrengungen der letzten Jahrzehnte ist die SCZ nach wie vor eine primär klinische Diagnose, basierend auf einer ausführlichen psychiatrischen Anamnese, da es bisher keinen genetischen, biochemischen oder bildgebenden Parameter gibt, der die Diagnose sichern kann (Owen et al. 2016).

### 1.1.2 Klinisches Erscheinungsbild mit geschlechtsspezifischen Unterschieden

Die vielfältigen klinischen Symptome der Schizophrenie sind komplex und können verschiedene Ebenen der Wahrnehmung, der Ich-Umwelt-Grenzen, der Kognition, des Affektes und der Psychomotorik betreffen. In der Praxis werden diese daher in üblicherweise drei Kategorien unterteilt: 1) positive Symptome wie Halluzinationen, Wahnvorstellungen und eine desorganisierte Sprache oder Verhaltensweisen, 2) negative Symptome wie eine affektive Verflachung oder ein sozialer und emotionaler Rückzug und 3) kognitive Symptome wie eine Beeinträchtigung des Gedächtnisses oder Störungen der exekutiven Funktionsfähigkeit (American Psychiatric Association 2000).

Innerhalb des letzten Jahrzehntes hat das Bewusstsein für geschlechtsspezifische Aspekte in der Medizin (Mauvais-Jarvis et al. 2020) und Wissenschaft (Shansky und Murphy 2021; Tannenbaum et al. 2019) stetig zugenommen. Weltweit konnten Forschende aufzeigen, dass diese Unterschiede sich auf verschiedene Merkmale einer Erkrankung, wie z. B. die Epidemiologie oder das klinische Erscheinungsbild, auswirken können (Wierenga et al. 2024). Dennoch sind die zugrundeliegenden biologischen Mechanismen bisher weitestgehend unverstanden, was zu einer Wissenslücke (Rechlin et al. 2022) mit potentiell schwerwiegenden Auswirkungen für beide Geschlechter führt. Auch für Erkrankungen des Gehirns sind solche geschlechtsspezifischen Aspekte mittlerweile wissenschaftlich anerkannt (Bölte et al. 2023).

Speziell für die SCZ bestand in der Vergangenheit lange die Ansicht, dass Frauen und Männer gleichermaßen häufig betroffen sind. Erst relativ aktuelle Metanalysen (McGrath et al. 2008; Jongsma et al. 2019) konnten eine moderat erhöhte Inzidenz für Männer im Vergleich zu Frauen belegen (Männer-zu-Frauen-Verhältnis 1.4:1 bzw. 1.7:1). Hinsichtlich des klinischen Erscheinungsbildes sind die geschlechtsspezifischen Unterschiede zwischen männlichen und weiblichen Betroffenen mit einer SCZ vielfältig (Gaebel und Wölwer 2010) und erste Beschreibungen gehen sogar auf den Erstbeschreiber Emil Kräpelin (Kräpelin 1919) zurück: So erkrankten Frauen tendenziell etwas später in ihrem Leben an einer SCZ und haben einen zweiten, schwächeren Inzidenzgipfel nach der Menopause.

Auch haben Patientinnen im Vergleich zu Patienten mit einer SCZ insgesamt eine bessere Prognose in Bezug auf den klinischen Verlauf und den Ausgang der Erkrankung, wozu beispielsweise ein besseres therapeutisches Ansprechen und eine damit verbundene, besser erhaltene psychosoziale Funktionsfähigkeit beitragen (Falkenburg und Tracy 2014; Abel et al. 2010). Am deutlichsten stellt sich dies in einer aktuellen Auswertung der durch die Weltgesundheitsorganisation durchgeführten *Global Burden of Disease* Studie (Solmi et al. 2023) dar, die beobachten konnte, dass sich das Geschlechterverhältnis bei der SCZ im Laufe der gesamten Lebensspanne umkehrt. Dies deutet darauf hin, dass sich neben dem Ersterkrankungsalter sogar das Gesamtüberleben bei der SCZ nach biologischem Geschlecht unterscheidet, mit einer stärkeren, krankheitsbedingten Reduzierung der Lebenserwartung bei betroffenen Männern.

Trotz potenzieller klinischer Implikationen, wie unterschiedlichen Arzneimittelnebenwirkungen und hormonellen Veränderungen sowohl innerhalb des Menstruationszyklus sowie im Verlauf des weiblichen Lebens (González-Rodríguez et al. 2020; Seeman 2021), wurden diese geschlechtsspezifischen Unterschiede in den jüngsten Leitlinien für die klinische Praxis kaum berücksichtigt (Deutsche Gesellschaft für Psychiatrie und Psychotherapie Psychosomatik und Nervenheilkunde 2019; American Psychiatric Association 2020).

Eine denkbare Hypothese hinsichtlich des biologischen Ursprungs dieser geschlechtsspezifischen Unterschiede besagt, dass genetische Risikofaktoren auf dem X-Chromosom eine mögliche Rolle in dem komplexen Zusammenspiel spielen, welches zur Entwicklung einer SCZ beiträgt (Bache und DeLisi 2018).

### 1.1.3 Aktuelle Fortschritte im Bereich der Genetik der Schizophrenie

Anhand bisheriger Studien ist bekannt, dass psychiatrische Erkrankungen in der Regel multifaktorieller Ursache sind, d.h. sowohl Umwelt- als auch genetische Faktoren tragen zur Krankheitsentstehung bei (Andreassen et al. 2023). Ausgehend von einer lange beobachteten familiären Häufung der SCZ schätzen aktuelle Zwillingsstudien die sogenannte *Erblichkeit* oder *Heritabilität* auf bis zu 80 % (Wray und Gottesman 2012). Diese beschreibt den Anteil an der Krankheitsentstehung, welcher im Durchschnitt aller Betroffenen auf erbliche Faktoren zurückzuführen ist.

Damit gilt die SCZ als eine der am stärksten durch erbliche Faktoren beeinflussten psychiatrischen Erkrankungen (Andreassen et al. 2023). Um die Natur der erblichen Faktoren aufzuklären, hat insbesondere seit etwa Mitte der 2000er Jahre die molekulargenetische Forschung grundlegende Einblicke in die genetische Architektur dieser Erkrankungen ermöglicht (Andreassen et al. 2023; Sullivan et al. 2024; Owen et al. 2023). Diese Fortschritte wurden maßgeblich befördert durch: 1. eine stetig wachsende Anzahl an Studienteilnehmenden, insbesondere durch den Zusammenschluss einzelner Studien in großangelegten, internationalen Konsortien wie dem *Psychiatric Genomics Consortium* (PGC); 2. den technologischen Fortschritt und die daraus resultierenden kontinuierlich abnehmenden Genotypisierungs-/Sequenzierkosten und 3. die Entwicklung ausgefeilter statistischer Methoden für die Analyse großer Datensätze.

Aktuelle genetische Studien zu psychiatrischen Erkrankungen greifen in der Regel auf einen der folgenden genom- oder exomweiten Ansätze zurück (Rees und Owen 2020): 1. Genomweite Assoziationsstudien (GWAS), welche die Allelhäufigkeiten von Einzelnukleotidpolymorphismen (SNPs) zwischen Betroffenen und Kontrollpersonen vergleichen, mit dem Ziel, häufige genetische Risikovarianten zu identifizieren, die jeweils für sich genommen, das Risiko zu erkranken nur in einem geringen Maße erhöhen (Uffelmann et al. 2021); 2. die genomweite Analyse spezifischer Kopienzahlvarianten (CNVs), sprich größerer, submikroskopischer Duplikationen und Deletionen, welche mehrere tausend Basenpaare umfassen und seltene Risikovarianten mit einer hohen Effektstärke darstellen können (Rees und Kirov 2021) und 3. die Untersuchung von seltenen kodierenden Varianten (RCVs) mittels Verfahren des sogenannten *Next Generation Sequencings* (NGS), welche sowohl Einzelnukleotidvarianten (SNVs) als auch kleinere Insertionen und Deletionen (InDels) mit moderater bis hoher Effektstärke umfassen.

Die bisherigen Befunde stützen einmal mehr die komplexe und polygene, d.h. durch eine Vielzahl unterschiedlicher Gene beeinflusste Natur der SCZ (Sullivan et al. 2024). Bis heute wurden Tausende von Risikovarianten in Hunderten von verschiedenen Genen identifiziert, welche sich über das gesamte menschliche Erbgut verteilen. Alleine die aktuell umfassendste vom PGC durgeführte systematische genetische Studie an 76.755 Betroffenen mit einer SCZ und 243.649 Kontrollpersonen konnte 342 unabhängige SNPs in 287 genomischen Loci identifizieren (Trubetskoy et al. 2022).

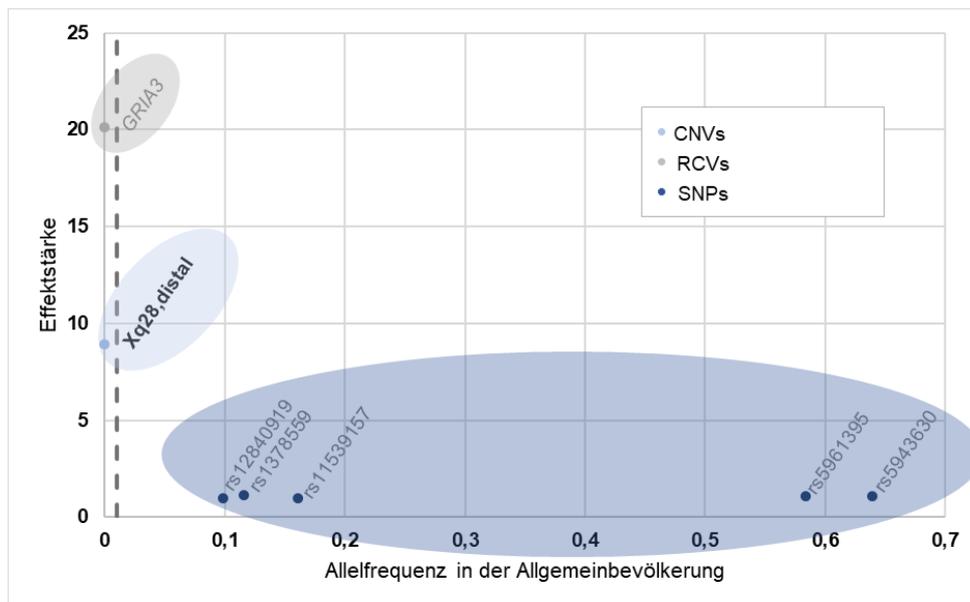
Dem genetischen Hintergrund der SCZ liegen daher hauptsächlich diese häufigen Risikovarianten mit jeweils geringer Effektstärke zugrunde, die jedoch in ihrer Summe zu einem potentiell krankheitsrelevanten Risiko beitragen können. Daneben liegen in einer Subgruppe von Betroffenen mit einer SCZ seltene genetische Veränderungen vor, die das Risiko, an einer SCZ zu erkranken, in einem stärkeren Maße erhöhen (Sullivan et al. 2024).

Seitdem Deletionen in der Chromosomensubbande 22q11.2 erstmals im Zusammenhang mit einem erhöhten Risiko für die SCZ beschrieben wurden (Murphy et al. 1999), stellen CNVs die bisher stärksten identifizierten individuellen genetischen Risikofaktoren dar und kommen insgesamt betrachtet bei schätzungsweise bis zu 2-3 % der Betroffenen mit einer SCZ vor (Rees et al. 2014). Zusätzlich gehen diese neuropsychiatrischen CNVs oftmals auch mit weiteren Auffälligkeiten wie z. B. einer kognitiven Beeinträchtigung und/oder angeborenen Herzfehlern einher (Rees und Kirov 2021). Da aktuelle Studien unter anderem anhand einer Analyse elektronischer Gesundheitsdaten nahelegen, dass nur etwa 5 % der Fälle mit einem klinisch relevanten, neuropsychiatrischen CNV eine entsprechende molekulargenetische Diagnose erhalten (Shimelis et al. 2022), ist zudem davon auszugehen, dass derzeit ein relevanter Anteil der Betroffenen mit einem seltenen genetischen Syndrom in der Psychiatrie unterdiagnostiziert bleibt. Folglich wurde für die SCZ bereits die Verwendung einer Array-basierten CNV-Untersuchung als routinemäßiger diagnostischer Test bei Vorliegen bestimmter Symptomkonstellationen diskutiert (Baker et al. 2014). Das Verständnis der zugrundeliegenden biologischen Mechanismen ist aber noch begrenzt, da die identifizierten CNVs in der Regel mehrere Gene umfassen, und damit die ursächlichen Gene, die den entsprechenden Assoziationen in ihrer biologischen Wirkung zugrunde liegen, in der Regel noch nicht bekannt sind (Rees und Kirov 2021).

Unabhängig davon standen aufgrund technischer und methodischer Einschränkungen weder geschlechtsabhängige Effekte (Blokland et al. 2022) noch X-chromosomale Risikofaktoren im Fokus der bisherigen systematischen genetischen Studien (Khramtsova et al. 2019). Die Gründe hierfür reichen beispielhaft von einer schlechteren Abdeckung des X-Chromosoms mit untersuchten SNPs über eine aufwendigere statistische Analyse X-chromosomaler Daten bis hin zur Nichtberücksichtigung der Geschlechtinformation in Fall- und Kontrollgruppen (Khramtsova et al. 2019).

Dabei konnte eine aktuelle Studie ausgehend von monogenen Erkrankungen eine Anreicherung von Genen auf dem X-Chromosom feststellen, die für die Funktion des Gehirnes relevant sind (Leitão et al. 2022). Auch für andere neuropsychiatrische Erkrankungen mit einer höheren Inzidenz für das männliche Geschlecht, wie z. B. die Intelligenzminderung oder Autismus-Spektrum-Störungen, sind mittlerweile eine Reihe von krankheitsursächlichen Varianten auf dem X-Chromosom für monogene Formen dieser Erkrankungen bekannt (Martin et al. 2021; Turner et al. 2019).

Neuere bioinformatische Analysemethoden sowie Genotypisierungs- und Sequenzieretechniken ermöglichten denn auch für die SCZ eine Identifikation der ersten genetischen Risikofaktoren auf dem X-Chromosom (Marshall et al. 2017; Singh et al. 2022; Trubetskoy et al. 2022). Diese spiegeln das gesamte Spektrum der genetischen Variabilität wieder, welches der Entstehung dieser multifaktoriellen Erkrankung zugrunde liegt, und sind zur Übersicht in Abb. 1 dargestellt.



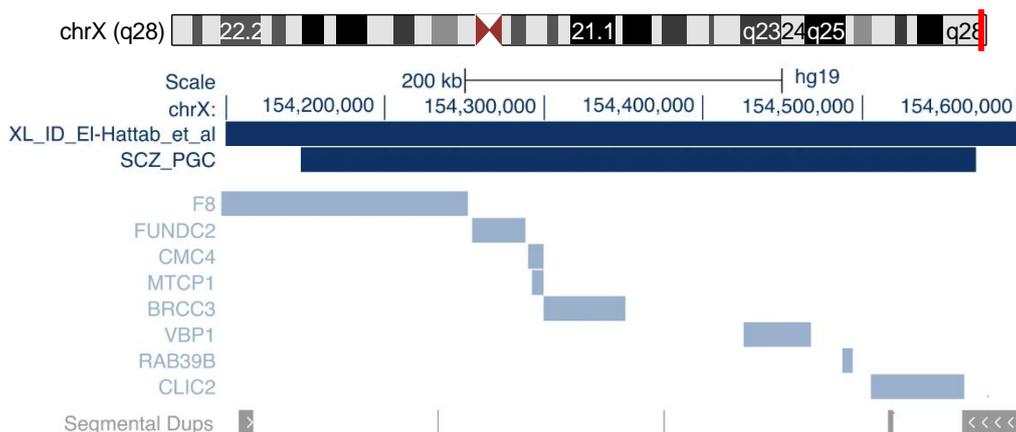
**Abb. 1:** Allelisches Spektrum der X-chromosomalen Risikofaktoren für die Schizophrenie. Diese Abbildung wurde in Anlehnung an (Sullivan et al. 2012) erstellt und fasst die Ergebnisse für das X-Chromosom der derzeit umfassendsten genetischen Analysen für die Schizophrenie auf Ebene der Kopienzahlvarianten (CNVs) (Marshall et al. 2017), seltener codierender Varianten (RCVs) (Singh et al. 2022) und Einzelnukleotidpolymorphismen (SNPs) (Trubetskoy et al. 2022) zusammen. Die graue Linie kennzeichnet einen Schwellenwert der Allelfrequenz von  $\leq 0,01$  % für seltene Varianten.

So konnte z. B. auf der Ebene der häufigen SNPs mit einer geringen Effektstärke die aktuellste GWAS für die SCZ die Anzahl der genomweiten signifikanten Loci für das X-Chromosom auf insgesamt fünf erhöhen (Trubetsky et al. 2022).

Hinsichtlich seltener RCVs mit moderater bis hoher Effektstärke identifizierte eine der derzeit umfassendsten Exomsequenzierungsstudien für die SCZ in einer wegweisenden Publikation des *Schizophrenia Exome Sequencing Meta-Analysis* (SCHEMA) Konsortiums das X-chromosomale *GRIA3*-Gen als eines von insgesamt zehn exomweit signifikanten Genen (Singh et al. 2022).

Im Jahr 2017 wurde zusätzlich erstmals ein potenzieller risikoassoziierter CNV auf dem X-Chromosom bei 18/21.094 Betroffenen mit einer SCZ und 2/20.227 Kontrollpersonen entdeckt (Marshall et al. 2017). Die Autoren des PGCs fanden heraus, dass Duplikationen des Xq28,distal Locus das SCZ-Risiko für beide Geschlechter erhöhen und die berichtete Odds Ratio (OR) betrug insgesamt 8,9, wobei bei Männern ein stärkerer Effekt beobachtet wurde (OR Frauen = 6,3; OR Männer =  $\infty$ ) (Marshall et al., 2017). Dabei umfassen diese Duplikationen die folgenden acht proteinkodierenden Gene: *Coagulation Factor 8 (F8)*, *FUN14 Domain-Containing Protein 2 (FUNDC2)*, *C-X9-C Motif-Containing Protein 4 (CMC4)*, *Mature T-Cell Proliferation 1 (MTCP1)*, *BRCA1/BRCA2-Containing Complex Subunit 3 (BRCC3)*, *Von Hippel-Lindau Binding Protein 1 (VBP1)*, *Ras-Related Protein Rab-39 (RAB39)* und *Chloride Intracellular Channel 2 (CLIC2)*.

Zudem brachten weitere Studien dieselbe genomische Region in den Zusammenhang mit einer X-chromosomalen Intelligenzminderung (Ballout et al. 2020; Ballout und El-Hattab 2021; El-Hattab et al. 2011; El-Hattab et al. 2015). Alle dieser bisher berichteten Träger\*innen einer Duplikation des Xq28,distal Locus teilen dieselben genomischen Bruchpunkte, welche in Abb. 2 veranschaulicht sind. Daher ist es eine Herausforderung, zu ermitteln, welches Gen bzw. welche Gruppe von Genen innerhalb dieser Region möglicherweise einen Beitrag zu dem erhöhten SCZ-Risiko leistet. Basierend auf den Ergebnissen von Marshall et al. könnten zusätzliche genetische Hinweise aus unabhängigen Studien dazu beitragen, das krankheitsrelevante Gen/die krankheitsrelevante Gruppe von Genen zu identifizieren.



**Abb. 2:** Schematische Abbildung des Xq28,distal Locus basierend auf dem UCSC Genome Browser (Kent et al. 2002). Dargestellt sind die genomischen Positionen der bereits identifizierten Duplikationen im Zusammenhang mit einer Schizophrenie (Marshall et al. 2017) und einer X-chromosomalen Intelligenzminderung (El-Hattab et al. 2011) sowie die beteiligten Gene.

Frühere Studien haben nahegelegt, dass Gendosis-Effekte zur Pathogenese der SCZ beitragen: So konnten mehrere Beispiele aufgezeigt werden, bei denen sowohl eine erhöhte als auch eine erniedrigte Gendosis das Risiko für eine SCZ erhöht, wie etwa die 1q21.1 oder 7q11.21(*ZNF92*) Regionen (Rees et al. 2016; Marshall et al. 2017).

Bei Duplikationen konnte die Wissenschaft zeigen, dass der Pathomechanismus prinzipiell entweder durch eine tatsächliche Überexpression von Kandidatengenen oder durch eine Haploinsuffizienz (z. B. durch die Unterbrechung eines dosisabhängigen Gens) vermittelt werden kann (Rice und McLysaght 2017). Im Speziellen ist auf der Grundlage bisheriger Studien ebenfalls bekannt, dass sowohl Duplikationen als auch seltene proteintrunkierende Varianten innerhalb desselben Gens zu einem erhöhten SCZ-Risiko führen können (Marshall et al. 2017; Singh et al. 2022). So wurde beispielsweise ein Zusammenhang zwischen einem erhöhten SCZ-Risiko und seltenen Duplikationen, die das *RB1CC1*-Gen betreffen, festgestellt (Degenhardt et al., 2013; Marshall et al., 2017), während dieses Gen eines der weiteren durch das SCHEMA Konsortium berichteten Gene mit exomweiter Signifikanz für die SCZ ist (Singh et al. 2022).

#### 1.1.4 Ziel der vorliegenden Dissertation

Um weiter einzugrenzen zu können, welches der Gene bzw. welche Gruppe von Genen innerhalb des Xq28,distal Locus zu dem potentiell erhöhten SCZ-Risiko beisteuert, war das Ziel der vorliegenden Dissertation, den Beitrag von seltenen und potenziell funktionellen Sequenzvarianten in dieser Region durch eine groß angelegte Sequenzierstudie zu analysieren. Die Auswertung der Ergebnisse erfolgte unter der Prämisse, dass beide Arten der genetischen Varianten, Duplikationen des Xq28,distal Locus wie auch funktionsmindernde RCVs, letztlich zu einer Verringerung der Genexpression führen können. Dies führte zu der Hypothese, dass phänotyprelevante Gene durch eine mögliche Anreicherung von seltenen und potenziell funktionsrelevanten Sequenzvarianten bei den Betroffenen mit einer SCZ im Vergleich zu der Kontrollgruppe identifiziert werden können. Dazu wurden gezielt die acht proteinkodierenden Gene des Xq28,distal Locus in einer umfangreichen europäischen SCZ-Fall-Kontroll-Kohorte (1,935 Betroffene mit einer SCZ und 1,905 Kontrollpersonen) mit der *single-molecule Molecular Inversion Probe* (smMIP) Methode sequenziert (Hiatt et al. 2013; O’Roak et al. 2012) und anschließend unter Verwendung verschiedener statistischer Methoden, wie dem *Optimal Unified Sequence Kernel Association Test* (SKAT-O), und Berücksichtigung potentieller geschlechtsspezifischer Effekte RCVs auf eine solche Anreicherung in der SCZ-Kohorte hin untersucht.

## 1.2 Material und Methoden

### 1.2.1 Kurzbeschreibung der eingeschlossenen Kohorten

Die Studie, die dieser Dissertation zugrunde liegt, wurde in Übereinstimmung mit der Deklaration von Helsinki des Weltärztebundes und den geltenden gesetzlichen Vorgaben erstellt. Insbesondere lagen die Zustimmung der zuständigen institutionellen Ethikkomitees sowie das schriftliche Einverständnis der Teilnehmenden vor Studieneinschluss vor. Die Rekrutierung erfolgte durch langjährige klinische Kooperationspartner des Instituts für Humangenetik an unterschiedlichen Standorten in Deutschland. Die entsprechenden Einschlusskriterien beider Studiengruppen sind in der Veröffentlichung über den Xq28,distal Locus beschrieben (Claus et al. 2025).

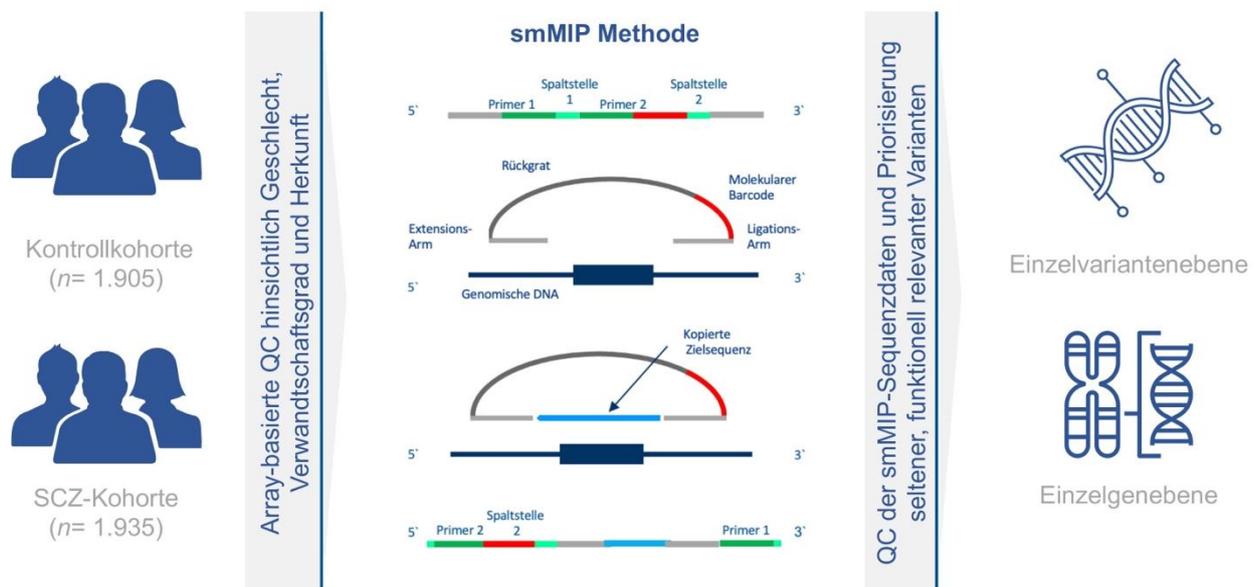
### 1.2.2 Gezielte Hochdurchsatz-Sequenzierung des Xq28,distal Locus

Die gezielte Sequenzierung mittels smMIPs stellt ein einfach zu skalierendes und kosteneffektives Hochdurchsatz-Sequenzierverfahren dar, welches die parallele Sequenzierung verschiedener Kandidatengene in einer großen Anzahl an Proben bei minimalem Einsatz von Desoxyribonukleinsäure (DNA) ermöglicht (Hiatt et al. 2013; O’Roak et al. 2012). Die in dieser Arbeit untersuchten acht proteinkodierenden Gene des Xq28,distal Locus waren Teil eines übergeordneten Resequenzierungsprojektes von Kandidatengenen für die SCZ, für die die folgenden Arbeitsschritte zeitgleich ausgeführt und die statistische Auswertung projektspezifisch an die besonderen Gegebenheiten der X-chromosomalen Vererbung angepasst wurden. Dabei lassen sich die einzelnen Laborabschnitte zusammenfassend wie folgt beschreiben: Zunächst werden die smMIP Sonden mittels der frei verfügbaren Software MIPgen (Boyle et al. 2014) entworfen. Da die einzelnen smMIPs jeweils mit einer individuellen Erkennungssequenz (sogenannten molekularen Barcodes) versehen sind, können diese in einem smMIP-Pool zusammengegeben werden und nach einer Aktivierung mittels Phosphorylierung parallel an zahlreiche genomische Zielregionen binden. Die dabei entstehenden Lücken werden daraufhin im Rahmen einer Polymerase/Ligase-Reaktion durch zugegebene Desoxynukleosidtriphosphate komplementär aufgefüllt, so dass die smMIPs kurzzeitig als zirkuläre DNA vorliegen. Im Anschluss entfernt eine Exonuklease-Behandlung nichtgebundene smMIPs und die verbliebende einzelsträngige genomische DNA, bevor die smMIPs wieder linear aufgetrennt werden.

Die Zielregionen werden danach mittels Polymerase-Kettenreaktion (PCR) mit für die Proben spezifischen molekularen Barcodes versehenen Primern vervielfältigt, so dass nach einer Aufreinigung eine große Anzahl an Studienteilnehmenden in einem übergeordneten Pool zusammengefasst und gleichzeitig sequenziert werden kann.

Zusätzlich wird zuvor in einem einmaligen, sogenannten *Rebalancing* Schritt eine ausreichende Abdeckung der genomischen Zielregionen sichergestellt, indem nach einem Testlauf in einer kleinen Stichprobe die Konzentration der entsprechenden smMIPs in Abhängigkeit von der durchschnittlichen Sequenziertiefe dieses Laufes angepasst oder die betreffende smMIP unphosphoryliert als Kompetitor hinzugegeben wurde.

Die Probenvorbereitung erfolgte unter Einhaltung der Herstellervorgaben und orientierte sich an bereits im Institut für Humangenetik etablierten und zuvor veröffentlichten Protokollen, die geringfügig modifiziert wurden (Thieme et al. 2021; Mathey et al. 2022). Die Sequenzierung wurde durch die NGS Core Facility der Medizinischen Fakultät der Universität Bonn auf einer Illumina HiSeq 2500 Plattform (Illumina, San Diego, CA, Vereinigte Staaten) durchgeführt. Eine schematische Darstellung der smMIP Methode und der nachfolgenden bioinformatischen Auswertung befindet sich in Abb. 3.



**Abb. 3:** Schematische Darstellung der smMIP-Methode und der bioinformatischen Aufbereitung (modifiziert nach (Hiatt et al. 2013)). *Abkürzungen:* SCZ, Schizophrenie; QC, Qualitätskontrolle; smMIP; single-molecule molecular inversion probe.

### 1.2.3 Bioinformatische Aufbereitung und Priorisierung der Varianten

Für eine erste Prozessierung der smMIP-Sequenzdaten wurde basierend auf der Herangehensweise der Entwickler der smMIP-Methode (Hiatt et al. 2013; O’Roak et al. 2012) und im Einklang mit den Richtlinien des *Genome Analysis Toolkits* (GATK) (van der Auwera et al. 2013) im Institut für Humangenetik eine bioinformatische Auswertepipeline entwickelt, welche bereits in vorangegangenen Publikationen beschrieben wurde (Thieme et al. 2021; Mathey et al. 2022). Die nachfolgende Qualitätskontrolle und statistische Analyse wurden in enger Zusammenarbeit mit der Core Unit für die bioinformatische Datenanalyse (CUBA) der Medizinischen Fakultät der Universität Bonn ausgeführt.

Dabei basierte die Qualitätskontrolle auf zwei übergeordneten Schritten: Im ersten Abschnitt wurden bereits verfügbare SNP-Array-Daten verwendet, um Proben auszuschließen, welche 1. zweitgradig oder enger miteinander verwandt sind (Verwandtschaftskoeffizient  $>0.088$ ) 2. ein unterschiedliches berichtetes und anhand der Array-Daten ermitteltes biologisches Geschlecht aufweisen und 3. nicht europäischer biogeographischer Herkunft sind (Abweichung von  $\pm 3$  Standardabweichung vom Mittelwert unter Einbeziehung der Phase 3 Daten des 1000 Genome Projektes (Auton et al. 2015)). Im zweiten Abschnitt erfolgte eine Auswertung der mit den smMIPs generierten Sequenzdaten, bei der sowohl Proben (Gesamtabdeckung von  $\geq 10X$  für mindestens 90 % der genomischen Zielregionen) als auch Varianten (basierend auf den harten GATK Filterkriterien (van der Auwera et al. 2013)) von geringer Qualität ausgeschlossen wurden. Zusätzlich wurden das Heterozygoten/Homozygoten- und Transversionen/Transitionen-Verhältnis untersucht und ebenfalls Proben ausgeschlossen, die die mehr als drei Standardabweichungen von dem errechneten Mittelwert innerhalb der mit den smMIPs generierten Sequenzdaten abwichen. Mit dem Ziel, den Beitrag seltener, kodierender Sequenzvarianten innerhalb des Xq28,distal Locus zu untersuchen, wurde anschließend der Fokus auf nichtsynonyme SNVs und InDels gelegt. Die verbliebenden Varianten wurden im Wesentlichen weiter nach selten (Minor-Allelfrequenz (MAF) von  $\leq 1$  % bzw.  $\leq 0.1$  % innerhalb der gesamten aktuell untersuchten Studienkohorte und anhand der Frequenzangaben des *Exome Aggregation Consortium* (ExAC) für die nicht-finnischen, nicht-psychiatrisch erkrankten Europäer (Lek et al. 2016)) und potentiell funktionell relevant (*Combined Annotation Dependent Depletion* (CADD) Score  $\geq 20$ ) priorisiert.

Der CADD-Score ist hierbei ein gängiges Instrument zur Einschätzung der Pathogenität von SNVs und gibt basierend auf maschinellem Lernen einen zusammengesetzten Wert aus, in den insbesondere verschiedene Parameter der Proteinstruktur und phylogenetischen Konservierung einfließen. Außerdem wurden alle Varianten, welche in den anschließenden SKAT-O Test eingingen, mit Hilfe des *Integrative Genomics Viewer* Version 2.16.1 (Robinson et al. 2011) visuell überprüft und einer Bestätigung mittels Sanger-Sequenzierung unterzogen.

Weitere technische Details des Laborablaufes der smMIP-Methode und der bioinformatischen Auswertung sind in der Veröffentlichung, die dieser Dissertation zugrunde liegt, aufgelistet (Claus et al. 2025).

#### 1.2.4 Statistische Analyse

Die statistische Analyse befasste sich maßgeblich mit zwei Hauptebenen: 1. Auf Ebene der einzelnen Varianten wurde für jede identifizierte Variante die Anzahl der Allele zwischen der SCZ- und der Kontrollkohorte mit einem zweiseitigen Fisher-Exakt-Test verglichen: (i) für die gesamte im Rahmen dieser Dissertation untersuchte Studienkohorte (ii) für Betroffene und Kontrollpersonen getrennt nach biologischem Geschlecht und (iii) in Kombination mit den Daten, die vom SCHEMA Konsortium (Singh et al. 2022) für die entsprechenden Varianten veröffentlicht wurden. 2. Auf Ebene der acht einzelnen proteinkodierenden Gene des Xq28,distal Locus wurde die statistische Analyse unter Verwendung des SKAT-O Tests (Ionita-Laza et al. 2013; Lee et al. 2012a; Lee et al. 2012b; Wu et al. 2011) durchgeführt. Hierbei wurde das Ziel verfolgt, eine mögliche erhöhte Mutationslast von seltenen und potentiell schädigenden Varianten innerhalb eines oder mehrerer der untersuchten Gene nachzuweisen. Um die Besonderheiten der X-chromosomalen Vererbung zu adressieren und sich so den zugrundeliegenden biologischen Mechanismen bestmöglich anzunähern, wurde bei dieser Testung das statistische Modell verwendet, welches von den Entwicklern des SKAT-O Test speziell für das X-Chromosom zur Verfügung gestellt wurde (Ma et al. 2015). Dabei wurde die Geschlechtsinformation als Kovariable berücksichtigt.

#### 1.3 Ergebnisse

Die Ergebnisse dieser Arbeit wurden nach einem Peer-Review-Verfahren als Originalartikel im *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* veröffentlicht (Claus et al. 2025).

### 1.3.1 Qualitätskontrolle mittels Array- und Sequenzierungsdaten

Insgesamt erfüllten 457 Proben nicht die Vorgaben der zweistufigen Qualitätskontrolle. Bezüglich des Verwandtschaftsgrades und der biogeographischen Herkunft ergab die Analyse der bereits vorliegenden SNP-Array-Daten, dass 196 der verwendeten Proben nicht die erforderlichen Kriterien erfüllten.

Zwischen dem berichteten und dem technisch ermittelten biologischem Geschlecht konnten keine Diskrepanzen festgestellt werden. Im Anschluss wurde auf der Ebene der smMIP-Sequenzdaten bestimmt, dass 187 Proben nicht die Anforderungen hinsichtlich einer ausreichenden Abdeckung der genomischen Zielregionen innerhalb des Xq28,distal Locus erfüllten und weitere 74 Proben als Ausreißer hinsichtlich des Heterozygoten/Homozygoten-Verhältnisses identifiziert wurden. Entsprechend umfasste das initiale Datenset nach der Qualitätskontrolle 71 unterschiedliche exonische Varianten in 3383 Proband\*innen (davon 1757 Betroffene mit einer SCZ und 1626 Kontrollpersonen).

### 1.3.2 Priorisierung und Verifikation der identifizierten Varianten

Von den 71 Varianten wurden 63 anhand des MAF-Grenzwertes ( $\leq 0,1\%$ ) für den SKAT-O Test als selten klassifiziert. Davon wurden wiederum 13 verschiedene Varianten in den Genen *F8*, *BRCC3*, *VBP1*, *FUNDC2* und *RAB39* mittels des CADD-Score-Grenzwertes von  $\geq 20$  als potentiell funktionell relevant eingestuft. Diese 13 Varianten wurden in 18 Studienteilnehmenden identifiziert. Anzumerken ist jedoch, dass für eine spezifische Variante im *BRCC3*-Gen die visuelle Überprüfung bei 3 von 5 Studienteilnehmenden, bei denen diese festgestellt wurde, auf ein technisches Artefakt hinwies. Dies bestätigte sich auch im Rahmen der Sanger-Sequenzierung.

Zusammenfassend wurden 13 verschiedene Varianten innerhalb von fünf Genen identifiziert, welche bei vier Betroffenen mit einer SCZ und 11 Kontrollpersonen vorlagen.

Eine detaillierte Aufstellung der priorisierten Varianten ist in Tabelle 1 der Veröffentlichung über den Xq28,distal Locus dargestellt (Claus et al. 2025).

### 1.3.3 Einzelvariantenanalyse mittels zweiseitigem Fisher-Exakt-Test

In der Einzelvariantenanalyse mittels eines zweiseitigen Fisher-Exakt-Tests waren keine der untersuchten 69 Varianten mit einem initialen MAF-Grenzwert von  $\leq 1\%$  (nominell) mit einem erhöhten Risiko für eine SCZ assoziiert. Die Mehrheit der Varianten (11/13), welche in den anschließenden SKAT-O Test eingegangen sind, wurde jeweils in einer oder einem Studienteilnehmenden detektiert, was einzeln betrachtet, ihre Seltenheit bestätigt.

Auch konnte keine weitere signifikante Anreicherung in der geschlechtsspezifischen Analyse oder in der kombinierten Analyse gefunden werden, welche systematisch die vom SCHEMA Konsortium zur Verfügung gestellte Anzahl der Allele zu den Daten der vorliegenden Studienkohorte hinzufügte. Auf der Einzelvariantenebene lag der niedrigste ermittelte  $p$ -Wert für diese Varianten bei 0,209 im Rahmen der kombinierten Analyse für die Varianten *F8:c.599A>G;p.(Glu200Gly)* und *FUNDC2:c.85C>T;p.(Arg29Cys)*. Die Ergebnisse der Einzelvariantenanalyse sind im Anhang der Veröffentlichung über den Xq28,distal Locus beigefügt (Claus et al. 2025).

### 1.3.4 Einzelgenanalyse mittels des SKAT-O Testes

Der SKAT-O Test ergab keine signifikant erhöhte Mutationslast innerhalb der acht proteinkodierenden Gene des Xq28,distal Locus (alle  $p$ -Werte  $> 0,05$ ). Der geringste  $p$ -Wert wurde für das *BRCC3*-Gen beobachtet ( $p = 0,103$ ;  $p$ -korrigiert = 0,515), welches eine Untereinheit des BRCA1/BRCA2- enthaltenen Komplexes kodiert. Da der SKAT-O Test grundsätzlich mehrere Richtungen eines kausalen Effektes berücksichtigt, könnte dies auf einen potenziell protektiven Effekt hindeuten. Anhand der verwendeten strengen Kriterien für die Qualitätskontrolle und Priorisierung, konnten in den Genen *CLIC2*, *MTCP1* und *CMC4* keine seltenen und potentiell funktionell relevanten Varianten identifiziert werden. Folglich konnten diese Gene bei der entsprechenden statistischen Analyse nicht berücksichtigt werden.

#### 1.4 Diskussion

Im Rahmen der vorliegenden Dissertation wurde erstmals eine detaillierte Analyse des Beitrags seltener kodierender Sequenzvarianten innerhalb eines potentiellen genetischen Risikolocus für die SCZ auf dem X-Chromosom erstellt, welcher zuvor in der derzeit größten genomweiten CNV-Analyse für diese Erkrankung berichtet wurde (Marshall et al. 2017).

Dabei zeigte keine der untersuchten seltenen und potenziell funktionellen Varianten oder eines der acht proteinkodierenden Gene des Xq28,distal Locus eine statistisch signifikante Anreicherung in Betroffenen mit einer SCZ im Vergleich zu Kontrollpersonen, obwohl diese Analyse in einer umfangreichen SCZ-Fall-Kontroll-Kohorte und unter Berücksichtigung möglicher geschlechtsspezifischer Effekte durchgeführt wurde (Claus et al. 2025).

Folglich stützt dieses Gesamtergebnis auch in der Zusammenschau mit den bisherigen systematischen genetischen Studien zur SCZ (wie z. B. der aktuell umfassendsten vom SCHEMA-Konsortium durchgeführten Meta-Analyse von Exom-Sequenzierungsdaten (Singh et al. 2022)) nicht die Hypothese, dass seltene und potenziell funktionelle Sequenzvarianten im Xq28,distal Locus das Risiko für die Entwicklung einer SCZ erhöhen.

Einerseits könnte die nur geringe Anzahl von priorisierten Varianten auf direkte Auswirkungen des Studiendesigns und/oder technische Einschränkungen der verwendeten Methoden zurückzuführen sein. Diese lassen sich zu vier wesentlichen Limitationen der vorliegenden Arbeit zusammenfassen und sind gemeinsam mit weiteren, weniger relevanten Aspekten in der Publikation über den Xq28,distal Locus aufgeführt (Claus et al. 2025):

1. Mit die bedeutsamste Einschränkung stellt voraussichtlich die Stichprobengröße der untersuchten Studienkohorte und die damit verbundene begrenzte statistische Aussagekraft dar. Wie die Arbeit des SCHEMA-Konsortiums belegen kann, ist eine Untersuchung von zehntausenden Studienteilnehmenden notwendig, um eine exomweite Signifikanz für gerade einmal zehn krankheitsrelevante Gene zu erreichen (Singh et al., 2022). Basierend auf den bisherigen Ergebnissen resultiert daher weiterhin die Möglichkeit, dass funktionell relevante Varianten innerhalb des Xq28,distal Locus eventuell so selten sind, dass sie folglich nur in noch größeren Studienkollektiven mit Sicherheit zu detektieren sind.

2. Zeigen die bisherigen genetischen Befunde auch am Beispiel des Xq28,distal Locus einmal mehr die Schwächen der primär klinischen Definition von psychiatrischen Erkrankungen auf: So konnte mit Hilfe der großen Zahl identifizierter genetischer Risikofaktoren gezeigt werden, dass die genetische Übereinstimmung zwischen verschiedenen Erkrankungen aus dem psychiatrischen Formenkreis wie auch zwischen psychiatrischen und neurologischen Erkrankungen weitreichender und komplexer ist, als es bisher angenommen wurde (Andreassen et al. 2023).

In Hinblick auf häufige Risikovarianten überlappt die SCZ dabei am stärksten mit der bipolaren Störung. Bei seltenen Risikovarianten besteht die größte Schnittmenge zu kindlichen Entwicklungsstörungen des Nervensystems, wie etwa der Intelligenzminderung oder Autismus-Spektrum-Störung (Owen et al. 2023). Für die bei der SCZ beteiligten CNVs, wie dem Xq28,distal Locus, konnten beispielsweise phänotypische Analysen eine starke Anreicherung für prämorbid kognitive Defizite in dieser speziellen Gruppe von Betroffenen feststellen (Foley et al. 2020). Einige Autoren gehen aufgrund dessen von einem phänotypischen und genetischen Kontinuum aus (Owen et al. 2023) und sprechen aktuell von einer Schizophrenie-Spektrum-Störung. Zukünftig könnte die psychiatrische Genetik zur Entwicklung von molekularen Subtypen beisteuern, um perspektivisch eine präzisere Versorgung von Betroffenen mit einer solchen Erkrankung zu ermöglichen.

3. Neben einer unzureichenden Berücksichtigung des Geschlechterverhältnisses innerhalb der verschiedenen Studiengruppen stellt auch der im Verhältnis überproportionale Einschluss von Studienteilnehmenden mit einer europäischen Abstammung im Vergleich zu Personen mit einer anderen biogeographischen Herkunft eine weitere Form einer möglichen Teilnahmeverzerrung und eine der zentralen Herausforderungen von genetischen Studien in der Psychiatrie dar (Derks et al. 2022). Dies trifft auch auf die vorliegenden Untersuchungen über den Xq28,distal Locus zu, da die Analysen auf Personen europäischer Abstammung begrenzt wurden, um eine genetisch möglichst homogene Studienkohorte zu gewährleisten.

Um die Übertragbarkeit der Ergebnisse zwischen Menschen mit verschiedener Abstammung zu evaluieren und mögliche bevölkerungsspezifische Risikofaktoren zu identifizieren, sollten künftige Studien zur SCZ den Beitrag von seltenen Sequenzvarianten innerhalb des Xq28,distal Locus bei Studienteilnehmenden mit unterschiedlicher biogeographischer Herkunft analysieren. In ersten vergleichenden Analysen zwischen Studienteilnehmenden mit ostasiatischer und europäischer Abstammung konnte festgestellt werden, dass die genetische Korrelation zwischen den einzelnen Studienkohorten bei der SCZ deutlich stärker ausgeprägt ist als beispielsweise bei der Depression. Somit ergibt sich das Bild, dass der Grad dieser populationsbasierten Effekte womöglich zwischen den verschiedenen Erkrankungen aus dem psychiatrischen Formenkreis variiert (Lam et al. 2019; Liu et al. 2023; Derks et al. 2022).

Ungeachtet dessen sind weiterhin internationale Anstrengungen erforderlich, damit die Ergebnisse der psychiatrisch-genetischen Forschung weltweit gleichermaßen zugänglich und anwendbar sind (Martin et al. 2019; Peterson et al. 2019).

4. Hinsichtlich der primär technischen Limitationen ist vorrangig anzumerken, dass durch die Verwendung der beschriebenen strengen Filterkriterien eine potentielle Überrepräsentation von Loss-of-function-Varianten und damit eine unzureichende Berücksichtigung eines zugrundeliegenden Gain-of-function-Mechanismus entstanden ist. Zumal ein solcher Mechanismus grundsätzlich eher im Einklang mit dem Konzept eines duplizierten Gens ist. Im Prinzip priorisiert der CADD-Score überwiegend Varianten, die einem Selektionsdruck unterliegen, mit dem primären Ziel, zwischen pathogenen und benignen genetischen Veränderungen zu unterscheiden. Dabei ist der CADD grundsätzlich nicht selektiv und kann prinzipiell sowohl auf Gain-of-function- als auch auf Loss-of-function-Varianten angewendet werden. Dennoch bezieht der dem CADD zugrundeliegende Algorithmus Scores wie z. B. SIFT und PolyPhen (Flanagan et al. 2010) mit ein, die vorwiegend negative Auswirkungen auf die Proteinfunktion in Betracht ziehen. Entsprechend konnten weiterführende Analysen aufzeigen, dass die derzeit verfügbaren Metaprädiktoren, wie der CADD-Score, die Pathogenität einer Variante genauer vorhersagen können, wenn diese auf einem Loss-of-Function-Mechanismus beruht (Sevim Bayrak et al. 2021).

Deshalb hat der CADD-Score wahrscheinlich nur eine begrenzte Aussagekraft für die in dieser Dissertation untersuchte Forschungsfrage, welche sich mit einer wiederkehrenden Duplikation befasst.

Dies gilt ebenso für weitere häufig verwendete In-silico-Vorhersagetools wie REVEL (Hopkins et al. 2023). Für neuere Scores, wie AlphaMissense (Cheng et al. 2023), kann dieser Aspekt bisher nicht abschließend beurteilt werden. Letztere auf künstlicher Intelligenz basierende Methoden stellen einen vielversprechenden Ansatz dar, um zukünftig auch die verschiedenen molekularen Krankheitsmechanismen besser in die Bewertung von Varianten zu integrieren.

Zudem könnten krankheitsrelevante Sequenzvarianten auch in den nicht kodierenden Bereichen des Xq28,distal Locus zu finden sein, welche in der vorliegenden Arbeit nicht betrachtet wurden. Diese Varianten sind aufgrund eines Mangels an Informationen in den bestehenden Datenbanken und verfügbaren Trainingsdatensätzen mit den gegenwärtigen bioinformatischen Vorhersagetools ebenfalls teils schwierig zu interpretieren.

Insgesamt können mit den derzeit verwendeten technischen Methoden schätzungsweise nur etwa 40 % der in den Zwillingsstudien ermittelten Heritabilität potentiell erfasst werden (Owen et al. 2023). Viele dieser technischen Limitationen könnten jedoch in Zukunft voraussichtlich mit der Anwendung neu aufkommender Technologien wie z. B. einer Genom- und Long-Read-Sequenzierung in großen und ethnisch vielfältigen Kohorten umgangen werden. Auf diesem Wege kann das Wissen über den Beitrag seltener genetischer Varianten zur Entwicklung der SCZ für die gesamte Weltbevölkerung erweitert werden. Erste Studien dieser Art wurden bereits durchgeführt und liefern erste vielversprechende Ergebnisse (Maury et al. 2024; Ormond et al. 2024; Feng et al. 2025), auch wenn der Zeitpunkt für eine umfassende Bewertung noch zu früh ist.

Andererseits könnte die geringe Anzahl der entdeckten Varianten auch bedeuten, dass seltene und potenziell funktionell relevante Sequenzvarianten innerhalb des Xq28,distal Locus tatsächlich nur eine begrenzte Rolle hinsichtlich eines möglicherweise erhöhten SCZ-Risikos spielen. Dies könnte auf einen eher allgemeineren Beitrag des Xq28,distal Locus bei der neuronalen Entwicklung hindeuten, welche über alternative Pathomechanismen vermittelt wird, die durch das vorliegende Studiendesign nicht erfasst wurden.

Zum Beispiel konnte im Rahmen einer Studie über das *RAB39B*-Gen eine erhöhte Genexpression in hippocampalen Neuronen gentechnischveränderter Mäuse und in Lymphozyten nachgewiesen werden, die aus dem Blut von einzelnen Träger\*innen einer Duplikation in Xq28,distal isoliert wurden (Vanmarsenille et al. 2014).

Neben einer solchen tatsächlichen Überexpression der beteiligten Gene besteht ein weiterer denkbarer Pathomechanismus darin, dass die Duplikation ihren Effekt durch die Dysregulation mehrerer Gene auch außerhalb des CNVs ausübt. So könnten Gene im Umfeld der Duplikation durch spezifische dreidimensionale funktionelle Einheiten, sogenannte *topologically associated domains* (TADs), miteinander verbunden oder voneinander getrennt sein. Strukturelle Varianten, zu denen in Abhängigkeit von ihrer genomischen Größe auch Deletionen und Duplikationen zählen, können diese Domänen unterbrechen. Im Ergebnis kann dies dazu führen, dass alternative TAD-Grenzen und somit neue, möglicherweise krankheitsursächliche Gen-Enhancer-Interaktionen entstehen (Spielmann et al. 2018). Dies konnte für das X-Chromosom beispielhaft im Zusammenhang mit einer 46,XY-Gonadendysgenese gezeigt werden (Meinel et al. 2023).

Entsprechend könnte gerade auch in Hinblick auf die Identifikation übergeordneter biologischer Mechanismen, auf denen die geschlechtsspezifischen Unterschiede der SCZ basieren, ein weiterer vielversprechender Ansatz darin bestehen, nicht nur Unterschiede auf Ebene der DNA-Sequenz, sondern auch in der Genexpression mittels RNA-Sequenzierung zu untersuchen. Jedoch legten dabei bislang nur wenige Studien dieser Art den Fokus auf eine gezielte Analyse der Genexpression zwischen weiblichen und männlichen Betroffenen mit einer SCZ (Carceller et al. 2024; Benjamin et al. 2024; Hoffman et al. 2022), obwohl insbesondere das weibliche Geschlechtshormon Östrogen über die Bindung an seinen nukleären Rezeptor direkt die Genexpression auf vielfältige Weise beeinflussen kann (Iqbal et al. 2024). Die dabei entstehenden neuroprotektiven Effekte (Arevalo et al. 2015) werden von einigen Autoren unter dem Begriff der „Östrogen-Hypothese“ zusammengefasst (Gogos et al. 2015). In diesem Kontext könnte eine solche Integration verschiedener sogenannter *Omics*-Ebenen zudem dabei behilflich sein, die zuvor aufgeführten technischen Limitationen zu überwinden, indem sie eine genauere Beurteilung der funktionellen Relevanz der identifizierten Varianten und somit eine Priorisierung der komplexen genetischen Befunde erleichtert.

Daneben existieren bereits auf klinischer Ebene Hinweise auf eine mögliche Beteiligung dieser Region an neuropsychiatrischen Erkrankungen: Äquivalente Duplikationen der Xq28,distal Region sind in beiden Geschlechtern wiederholt im Zusammenhang mit einem sehr seltenen neuropsychiatrischen Syndrom beschrieben (Ballout und El-Hattab 2021; Ballout et al. 2020; El-Hattab et al. 2015; El-Hattab et al. 2011). Klinisch ist dieses Syndrom dabei durch einen variablen Grad einer X-chromosomalen Intelligenzminderung, faziale Dysmorphien und weitere neuropsychiatrische Symptome gekennzeichnet. Interessanterweise wiesen 16 der zuvor 35 in der Literatur im Zusammenhang mit einer Intelligenzminderung bzw. Entwicklungsverzögerung klinisch charakterisierten Duplikations-träger\*innen neuropsychiatrische Symptome auf, welche von Merkmalen einer Autismus-Spektrum-Störung bis hin zur Manifestation einer Psychose reichten (Ballout et al. 2020; Ballout und El-Hattab 2021; El-Hattab et al. 2011; El-Hattab et al. 2015). Die Symptomatik ist hierbei in hemizygot betroffenen Patienten im Vergleich zu heterozygoten Patientinnen tendenziell ausgeprägter. Zusätzlich konnte beobachtet werden, dass bei der Mehrzahl der weiblichen Betroffenen eine verschobene X-Inaktivierung vorliegt (El-Hattab et al. 2015). Weiterhin ist zu beachten, dass die Gesamtanzahl der identifizierten Duplikations-träger\*innen nach wie vor unzureichend ist, um verlässliche Angaben zur Penetranz von neuropsychiatrischen Erkrankungen machen zu können, obwohl parallel zur Einreichung der dieser Dissertation zugrundeliegenden Veröffentlichung weitere Betroffene mit einer Duplikation in Xq28,distal und neuropsychiatrischen Auffälligkeiten publiziert wurden (Levy et al. 2024). Dies ist insbesondere für Familien relevant, die die entsprechende Diagnose bereits pränatal erhalten (Ballout et al. 2020; Levy et al. 2024). Hierfür sind neben den großangelegten, systematischen genetischen Studien auch phänotypische Studien notwendig, die eine bestimmte Gruppe, wie z. B. Betroffene mit demselben CNV, tiefergehend klinisch charakterisieren (Rees und Owen 2020).

Im Gegensatz dazu wurde die reziproke Deletion in der Literatur bisher nur bei weiblichen Betroffenen dokumentiert. Daher postulieren mehrere Autoren diesbezüglich eine Letalität bereits während der Embryonalphase für männliche Deletionsträger (El-Hattab et al. 2015; Marshall et al. 2017).

Darüber hinaus wurden zwei der Gene des Xq28,distal Locus unabhängig voneinander mit einer X-chromosomalen Intelligenzminderung in Verbindung gebracht. Sowohl Missense-Varianten innerhalb der im Gehirn exprimierten Gene *RAB39B* und *CLIC2* (Gianandrea et al. 2010; Takano et al. 2012) sowie eine weiter distal gelegene, überlappende Duplikation, die diese beiden Gene umfasst (Andersen et al. 2014), wurden als krankheitsursächlich für einen neuropsychiatrischen Phänotyp identifiziert. Dieser ist dem der Duplikationsträger\*innen der gesamten Xq28,distal Region sehr ähnlich. *CLIC2* kodiert dabei für ein intrazelluläres Transportprotein, das als Kalziumkanal den Kalziumhaushalt und über eine Modulation der Ryanodin-Rezeptor-Aktivität grundlegende zelluläre Prozesse mitreguliert (Board et al. 2004). Hingegen ist das *RAB39B*-Protein ein Mitglied der RAS-Onkogen-Familie und eine kleine GTPase, die durch ihre Rolle in einer Reihe von neuronalen Prozessen wie dem vesikulären Transport zur neuronalen Funktionsfähigkeit und der Ausbildung neuer Synapsen beiträgt (Cheng et al. 2002).

Zusammenfassend liefern die Ergebnisse der vorliegenden Arbeit keinen Hinweis auf einen Beitrag seltener und potentiell funktionell relevanter Sequenzvarianten innerhalb des Xq28,distal Locus zum Risiko, an einer SCZ zu erkranken.

Jedoch ist es in Anbetracht der begrenzten statistischen Aussagekraft und der technischen sowie methodischen Limitationen auf Grundlage dieser Dissertation zu früh, vollständig auszuschließen, dass diese zur Entwicklung einer SCZ beitragen.

Angesichts der dargelegten Assoziationen mit unterschiedlichen neuropsychiatrischen Erkrankungsbildern könnte diese Region dennoch wertvolle weitere Einblicke liefern. Aus diesem Grund sind weitere Untersuchungen des Xq28,distal Locus sowie systematische Studien von X-chromosomalen Risikofaktoren im Allgemeinen in der Forschung zur Genetik der SCZ nach wie vor erforderlich.

Ein tieferes Verständnis geschlechtsabhängiger und geschlechtsspezifischer Risikofaktoren könnte in Zukunft die biologischen Mechanismen weiter entschlüsseln, die den Unterschieden in der klinischen Symptomatik zwischen weiblichen und männlichen Betroffenen mit einer SCZ zugrunde liegen. Dies kann den Weg für den Einzug präziser und wissenschaftlich fundierter Empfehlungen in die klinischen Leitlinien für eine Erkrankung ebnen, die gleichermaßen schwerwiegende Folgen für beide Geschlechter hat.

## 1.5 Zusammenfassung

Obwohl geschlechtsspezifische Unterschiede im klinischen Erscheinungsbild der Schizophrenie (SCZ) allgemein anerkannt sind, standen genetische Risikofaktoren auf dem X-Chromosom bisher nicht im Fokus aktueller systematischer Studien zur komplexen genetischen Architektur dieser polygenen Erkrankung. Trotzdem konnte die weltweit größte Analyse von Kopienzahlvarianten, einer Gruppe von seltene genetischen Risikofaktoren mit einer hohen Effektstärke, erstmals einen potentiellen X-chromosomalen Risiko-Locus bei der SCZ identifizieren. Duplikationen in Xq28,distal erhöhen das Risiko für eine SCZ in beiden Geschlechtern und umfassen acht proteinkodierende Gene. Zudem sind Duplikationen in Xq28,distal in Betroffenen mit einer X-chromosomalen Intelligenzminderung beschrieben, welche oft von psychiatrischen Symptomen begleitet wird. Aufgrund gemeinsamer genetischer Bruchpunkte dieser bereits bekannten Duplikationen, ist es derzeit unklar, welches oder welche der entsprechenden Gene möglicherweise krankheitsrelevant sind. Ziel dieser Dissertation war es daher, den Beitrag seltener und potentiell funktionell relevanter Sequenzvarianten in dieser Region zu untersuchen, um eine mögliche Anreicherung dieser Varianten in krankheitsrelevanten Genen nachzuweisen.

Dazu wurden die beteiligten Gene in einer umfangreichen Kohorte von 1935 Betroffenen mit einer SCZ und 1905 Kontrollpersonen mit sogenannten *single-molecule Molecular Inversion Probes* (smMIPs), einer Methode der gezielten Hochdurchsatz-Sequenzierung, systematisch resequenziert und auf eine solche Anreicherung unter Berücksichtigung potentieller geschlechtsspezifischer Effekte hin getestet. Insgesamt ließen sich 13 seltene, potentiell funktionelle Varianten feststellen. Allerdings zeigte keine dieser Varianten oder eines der untersuchten Gene eine statistisch signifikante Anreicherung. Zusammenfassend ist dies einerseits auf methodische Limitationen wie die in erster Linie durch die Stichprobengröße begrenzte statistische Aussagekraft zurückzuführen; andererseits könnte dies in Anbetracht der klinischen Hinweise auch auf alternative Pathomechanismen hindeuten, welche mit dem vorliegenden Studiendesign nicht berücksichtigt werden konnten (wie z. B. eine tatsächliche Überexpression krankheitsrelevanter Gene). Perspektivisch könnte in Zukunft ein besseres Verständnis geschlechtsspezifischer bzw. geschlechtsabhängiger Risikofaktoren dazu beitragen, die biologischen Grundlagen der klinischen Unterschiede zwischen männlichen und weiblichen Betroffenen mit einer SCZ umfassender zu verstehen.

## 1.6 Literaturverzeichnis der deutschen Zusammenfassung

Abel KM, Drake R, Goldstein JM. Sex differences in schizophrenia. *Int Rev Psychiatry* 2010; 22: 417–428

American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (4th ed., text rev.)*. Washington, D.C.: American Psychiatric Association Publishing 2000

American Psychiatric Association. *The American Psychiatric Association Practice Guideline for the Treatment of Patients With Schizophrenia, Third Edition*. Washington, D.C.: American Psychiatric Association Publishing 2020

Andersen EF, Baldwin EE, Ellingwood S, Smith R, Lamb AN. Xq28 duplication overlapping the int22h-1/int22h-2 region and including RAB39B and CLIC2 in a family with intellectual and developmental disability. *Am J Med Genet A* 2014; 164: 1795–1801

Andreassen OA, Hindley GF, Frei O, Smeland OB. New insights from the last decade of research in psychiatric genetics: discoveries, challenges and clinical implications. *World Psychiatry* 2023; 22: 4–24

Arevalo M-A, Azcoitia I, Garcia-Segura LM. The neuroprotective actions of oestradiol and oestrogen receptors. *Nat Rev Neurosci* 2015; 16: 17–29

Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, et al. A global reference for human genetic variation. *Nature* 2015; 526: 68–74

Bache WK, DeLisi LE. The Sex-Chromosome Hypothesis of Schizophrenia: Alive, Dead, or Forgotten? A Commentary and Review. *Mol Neuropsychiatry* 2018; 4: 83–89

Baker K, Costain G, Fung WLA, Bassett AS. Chromosomal microarray analysis-a routine clinical genetic test for patients with schizophrenia. *Lancet Psychiatry* 2014; 1: 329–331

Ballout RA, Dickerson C, Wick MJ, Al-Sweel N, Openshaw AS, Srivastava S, Swanson LC, Bramswig NC, Kuechler A, Hong B, et al. Int22h1/Int22h2-mediated Xq28 duplication syndrome: de novo duplications, prenatal diagnoses, and additional phenotypic features. *Hum Mutat* 2020; 41: 1238–1249

Ballout RA, El-Hattab AW. The int22h1/int22h2-mediated xq28 duplication syndrome: An intersection between neurodevelopment, immunology, and cancer. *Genes* 2021; 12: 860

Benjamin KJM, Arora R, Feltrin AS, Perteu G, Giles HH, Stolz JM, D'Ignazio L, Collado-Torres L, Shin JH, Ulrich WS, et al. Sex affects transcriptional associations with schizophrenia across the dorsolateral prefrontal cortex, hippocampus, and caudate nucleus. *Nat Commun* 2024; 15: 3980

Blokland GA, Grove J, Chen C-Y, Cotsapas C, Tobet S, Handa R, Ripke S, Neale BM, Corvin A, Walters JT, et al. Sex-Dependent Shared and Nonshared Genetic Architecture Across Mood and Psychotic Disorders. *Biol Psychiatry* 2022; 91: 102–117

Board PG, Coggan M, Watson S, Gage PW, Dulhunty AF. CLIC-2 modulates cardiac ryanodine receptor Ca<sup>2+</sup> release channels. *Int J Biochem Cell Biol* 2004; 36: 1599–1612

Bölte S, Neufeld J, Marschik PB, Williams ZJ, Gallagher L, Lai M-C. Sex and gender in neurodevelopmental conditions. *Nat Rev Neurol* 2023; 19: 136–159

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

Carceller H, Hidalgo MR, Escartí MJ, Nacher J, La Iglesia-Vayá M de, García-García F. The impact of sex on gene expression in the brain of schizophrenic patients: a systematic review and meta-analysis of transcriptomic studies. *Biol Sex Differ* 2024; 15: 59

Cheng H, Ma Y, Ni X, Jiang M, Guo L, Ying K, Xie Y, Mao Y. Isolation and characterization of a human novel RAB (RAB39B) gene. *Cytogenet Genome Res* 2002; 97: 72–75

Cheng J, Novati G, Pan J, Bycroft C, Žemgulytė A, Applebaum T, Pritzel A, Wong LH, Zielinski M, Sargeant T, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* 2023; 381: eadg7492

Claus I, Sivalingam S, Koller AC, Weiß A, Mathey CM, Sindermann L, Klein D, Henschel L, Ludwig KU, Hoffmann P, et al. Contribution of Rare and Potentially Functionally Relevant Sequence Variants in Schizophrenia Risk-Locus Xq28,distal. *Am J Med Genet B Neuropsychiatr Genet* 2025; 198: e33011

Derks EM, Thorp JG, Gerring ZF. Ten challenges for clinical translation in psychiatric genetics. *Nat Genet* 2022; 54: 1457–1465

Deutsche Gesellschaft für Psychiatrie und Psychotherapie Psychosomatik und Nervenheilkunde. S3-Leitlinie Schizophrenie. Langfassung 2019, Version 1.0, zuletzt geändert am 15. März 2019: DGPPN e.V. (Hrsg.) für die Leitliniengruppe

EI-Hattab AW, Fang P, Jin W, Hughes JR, Gibson JB, Patel GS, Grange DK, Manwaring LP, Patel A, Stankiewicz P, et al. Int22h-1/int22h-2-mediated Xq28 rearrangements: Intellectual disability associated with duplications and in utero male lethality with deletions. *J Med Genet* 2011; 48: 840–850

EI-Hattab AW, Schaaf CP, Fang P, Roeder E, Kimonis VE, Church JA, Patel A, Cheung SW. Clinical characterization of int22h1/int22h2-mediated Xq28 duplication/deletion: New cases and literature review. *BMC Med Genet* 2015; 16: 0–11

Falkenburg J, Tracy DK. Sex and schizophrenia: a review of gender differences. *Psychosis* 2014; 6: 61–69

Feng Y-CA, Chen WJ, Lin M-C, Hsu JS, Cheng C-F, Liu C-M, Hwu H-G, Huang Y-T, Lu T-P, Wang S-H. Paternal age, de novo mutation, and age at onset among co-affected schizophrenia sib-pairs: whole-genome sequencing in multiplex families. *Mol Psychiatry* 2025

Flanagan SE, Patch A-M, Ellard S. Using SIFT and PolyPhen to Predict Loss-of-Function and Gain-of-Function Mutations. *Genet Test Mol Biomarkers* 2010; 14: 533–537

Foley C, Heron EA, Harold D, Walters J, Owen M, O'Donovan M, Sebat J, Kelleher E, Mooney C, Durand A, et al. Identifying schizophrenia patients who carry pathogenic genetic copy number variants using standard clinical assessment: retrospective cohort study. *Br J Psychiatry* 2020; 216: 275–279

Gaebel W, Wölwer W. Gesundheitsberichterstattung des Bundes -Themenheft 50 "Schizophrenie". Berlin: Robert Koch-Institut

Giannandrea M, Bianchi V, Mignogna ML, Sirri A, Carrabino S, D'Elia E, Vecellio M, Russo S, Cogliati F, Larizza L, et al. Mutations in the Small GTPase Gene RAB39B Are Responsible for X-linked Mental Retardation Associated with Autism, Epilepsy, and Macrocephaly. *Am J Hum Genet* 2010; 86: 185–195

Gogos A, Sbisa AM, Sun J, Gibbons A, Udawela M, Dean B. A Role for Estrogen in Schizophrenia: Clinical and Preclinical Findings. *Int J Endocrinol* 2015; 2015: 615356

González-Rodríguez A, Guàrdia A, Álvarez Pedrero A, Betriu M, Cobo J, Acebillo S, Monreal JA, Seeman MV, Palao D, Labad J. Women with Schizophrenia over the Life Span: Health Promotion, Treatment and Outcomes. *Int J Environ Res Public Health* 2020; 17: 5594

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

Hjorthøj C, Stürup AE, McGrath JJ, Nordentoft M. Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. *Lancet Psychiatry* 2017; 4: 295–301

Hoffman GE, Ma Y, Montgomery KS, Bendl J, Jaiswal MK, Kozlenkov A, Peters MA, Dracheva S, Fullard JF, Chess A, et al. Sex Differences in the Human Brain Transcriptome of Cases With Schizophrenia. *Biol Psychiatry* 2022; 91: 92–101

Hopkins JJ, Wakeling MN, Johnson MB, Flanagan SE, Laver TW. REVEL Is Better at Predicting Pathogenicity of Loss-of-Function than Gain-of-Function Variants. *Hum Mutat* 2023; 2023: 8857940

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

Iqbal J, Huang G-D, Xue Y-X, Yang M, Jia X-J. Role of estrogen in sex differences in memory, emotion and neuropsychiatric disorders. *Mol Biol Rep* 2024; 51: 415

Jauhar S, Johnstone M, McKenna PJ. Schizophrenia. *Lancet* 2022; 399: 473–486

Jongsma HE, Turner C, Kirkbride JB, Jones PB. International incidence of psychotic disorders, 2002-17: a systematic review and meta-analysis. *Lancet Public Health* 2019; 4: e229-e244

Kahn RS, Sommer IE, Murray RM, Meyer-Lindenberg A, Weinberger DR, Cannon TD, O'Donovan M, Correll CU, Kane JM, van Os J, et al. Schizophrenia. *Nat Rev Dis Primers* 2015; 1: 15067

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

Khramtsova EA, Davis LK, Stranger BE. The role of sex in the genomics of human complex traits. *Nat Rev Genet* 2019; 20: 173–190

Kräpelin E. *Dementia praecox and paraphrenia*. Chicago: Chicago Medical Book Company 1919

Lam M, Chen C-Y, Li Z, Martin AR, Bryois J, Ma X, Gaspar H, Ikeda M, Benjamin B, Brown BC, et al. Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nat Genet* 2019; 51: 1670–1678

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

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

Leitão E, Schröder C, Parenti I, Dalle C, Rastetter A, Kühnel T, Kuechler A, Kaya S, Gérard B, Schaefer E, et al. Systematic analysis and prediction of genes associated with monogenic disorders on human chromosome X. *Nat Commun* 2022; 13: 6570

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016; 536: 285–291

Levy M, Elron E, Shohat M, Lifshitz S, Kahana S, Shani H, Grossman A, Amar S, Narkis G, Sagi-Dain L, et al. Exploring inheritance, and clinical penetrance of distal Xq28 duplication syndrome: insights from 47 new unpublished cases. *J Hum Genet* 2024; 69: 337–343

Liu D, Meyer D, Fennessy B, Feng C, Cheng E, Johnson JS, Park YJ, Rieder M-K, Ascillio S, Pins A de, et al. Schizophrenia risk conferred by rare protein-truncating variants is conserved across diverse human populations. *Nat Genet* 2023; 55: 369–376

Ma C, Boehnke M, Lee S, the GoT2D investigators. Evaluating the Calibration and Power of Three Gene-Based Association Tests of Rare Variants for the X Chromosome. *Genet Epidemiol* 2015; 39: 499–508

Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, Antaki D, Shetty A, Holmans PA, Pinto D, et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet* 2017; 49: 27–35

Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* 2019; 51: 584–591

Martin HC, Gardner EJ, Samocha KE, Kaplanis J, Akawi N, Sifrim A, Eberhardt RY, Tavares ALT, Neville MDC, Niemi MEK, et al. The contribution of X-linked coding variation to severe developmental disorders. *Nat Commun* 2021; 12: 627

Marwaha S, Johnson S, Bebbington P, Stafford M, Angermeyer MC, Brugha T, Azorin J-M, Kilian R, Hansen K, Toumi M. Rates and correlates of employment in people with schizophrenia in the UK, France and Germany. *Br J Psychiatry* 2007; 191: 30–37

Mathey CM, Maj C, Scheer AB, Fazaal J, Wedi B, Wieczorek D, Amann PM, Löffler H, Koch L, Schöffl C, et al. Molecular Genetic Screening in Patients With ACE Inhibitor / Angiotensin Receptor Blocker-Induced Angioedema to Explore the Role of Hereditary Angioedema Genes. *Front Genet* 2022; 13: 1–11

Maury EA, Jones A, Seplyarskiy V, Nguyen TTL, Rosenbluh C, Bae T, Wang Y, Abyzov A, Khoshkhoo S, Chahine Y, et al. Somatic mosaicism in schizophrenia brains reveals prenatal mutational processes. *Science* 2024; 386: 217–224

Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, Brinton RD, Carrero J-J, DeMeo DL, Vries GJ de, Epperson CN, Govindan R, Klein SL, et al. Sex and gender: modifiers of health, disease, and medicine. *Lancet* 2020; 396: 565–582

McGrath J, Saha S, Chant D, Welham J. Schizophrenia: A concise overview of incidence, prevalence, and mortality. *Epidemiol Rev* 2008; 30: 67–76

Meinel JA, Yumiceba V, Künstner A, Schultz K, Kruse N, Kaiser FJ, Holterhus P-M, Claviez A, Hiort O, Busch H, et al. Disruption of the topologically associated domain at Xp21.2 is related to 46,XY gonadal dysgenesis. *J Med Genet* 2023; 60: 469–476

Murphy KC, Jones LA, Owen MJ. High Rates of Schizophrenia in Adults With Velo-Cardio-Facial Syndrome. *Arch Gen Psychiatry* 1999; 56: 940–945

O’Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, Carvill G, Kumar A, Lee C, Ankenman K, et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 2012; 338: 1619–1622

Ormond C, Ryan NM, Hedman AM, Cannon TD, Sullivan PF, Gill M, Hultman C, Heron EA, Johansson V, Corvin A. Whole genome sequencing study of identical twins discordant for psychosis. *Transl Psychiatry* 2024; 14: 313

Owen MJ, Legge SE, Rees E, Walters JTR, O'Donovan MC. Genomic findings in schizophrenia and their implications. *Mol Psychiatry* 2023; 28: 3638–3647

Owen MJ, Sawa A, Mortensen PB. Schizophrenia. *Lancet* 2016; 388: 86–97

Peterson RE, Kuchenbaecker K, Walters RK, Chen C-Y, Popejoy AB, Periyasamy S, Lam M, Iyegbe C, Strawbridge RJ, Brick L, et al. Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell* 2019; 179: 589–603

Plana-Ripoll O, Pedersen CB, Agerbo E, Holtz Y, Erlangsen A, Canudas-Romo V, Andersen PK, Charlson FJ, Christensen MK, Erskine HE, et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. *Lancet* 2019; 394: 1827–1835

Rechlin RK, Splinter TFL, Hodges TE, Albert AY, Galea LAM. An analysis of neuroscience and psychiatry papers published from 2009 and 2019 outlines opportunities for increasing discovery of sex differences. *Nat Commun* 2022; 13: 2137

Rees E, Kendall K, Pardiñas AF, Legge SE, Pocklington A, Escott-Price V, MacCabe JH, Collier DA, Holmans P, O'Donovan MC, et al. Analysis of Intellectual Disability Copy Number Variants for Association With Schizophrenia. *JAMA Psychiatry* 2016; 73: 963–969

Rees E, Kirov G. Copy number variation and neuropsychiatric illness. *Curr Opin Genet Dev* 2021; 68: 57–63

Rees E, Owen MJ. Translating insights from neuropsychiatric genetics and genomics for precision psychiatry. *Genome Med* 2020; 12: 43

Rees E, Walters JTR, Georgieva L, Isles AR, Chambert KD, Richards AL, Mahoney-Davies G, Legge SE, Moran JL, McCarroll SA, et al. Analysis of copy number variations at 15 schizophrenia-associated loci. *Br J Psychiatry* 2014; 204: 108–114

Rice AM, McLysaght A. Dosage-sensitive genes in evolution and disease. *BMC Biol* 2017; 15: 78

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nature Biotechnol* 2011; 29: 24–26

Seeman MV. Sex differences in schizophrenia relevant to clinical care. *Expert Rev Neurother* 2021; 21: 443–453

Sevim Bayrak C, Stein D, Jain A, Chaudhary K, Nadkarni GN, van Vleck TT, Puel A, Boisson-Dupuis S, Okada S, Stenson PD, et al. Identification of discriminative gene-level and protein-level features associated with pathogenic gain-of-function and loss-of-function variants. *Am J Hum Genet* 2021; 108: 2301–2318

Shansky RM, Murphy AZ. Considering sex as a biological variable will require a global shift in science culture. *Nature Neurosci* 2021; 24: 457–464

Shimelis H, Oetjens MT, Walsh LK, Wain KE, Znidarsic M, Myers SM, Finucane BM, Ledbetter DH, Martin CL. Prevalence and Penetrance of Rare Pathogenic Variants in Neurodevelopmental Psychiatric Genes in a Health Care System Population. *Am J Psychiatry* 2022; 180: 65–72

Singh T, Poterba T, Curtis D, Akil H, Al Eissa M, Barchas JD, Bass N, Bigdeli TB, Breen G, Bromet EJ, et al. Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* 2022; 604: 509–516

Solmi M, Seitidis G, Mavridis D, Correll CU, Dragioti E, Guimond S, Tuominen L, Dargél A, Carvalho AF, Fornaro M, et al. Incidence, prevalence, and global burden of schizophrenia - data, with critical appraisal, from the Global Burden of Disease (GBD) 2019. *Mol Psychiatry* 2023; 28: 5319–5327

Spielmann M, Lupiáñez DG, Mundlos S. Structural variation in the 3D genome. *Nat Rev Genet* 2018; 19: 453–467

Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: The emerging picture and its implications. *Nat Rev Genet* 2012; 13: 537–551

Sullivan PF, Yao S, Hjerling-Leffler J. Schizophrenia genomics: genetic complexity and functional insights. *Nat Rev Neurosci* 2024; 25: 611–624

Takano K, Liu D, Tarpey P, Gallant E, Lam A, Witham S, Alexov E, Chaubey A, Stevenson RE, Schwartz CE, et al. An x-linked channelopathy with cardiomegaly due to a CLIC2 mutation enhancing ryanodine receptor channel activity. *Hum Mol Genet* 2012; 21: 4497–4507

Tannenbaum C, Ellis RP, Eyssel F, Zou J, Schiebinger L. Sex and gender analysis improves science and engineering. *Nature* 2019; 575: 137–146

Thieme F, Henschel L, Hammond NL, Ishorst N, Hausen J, Adamson AD, Biedermann A, Bowes J, Zieger HK, Maj C, et al. Extending the allelic spectrum at noncoding risk loci of orofacial clefting. *Hum Mutat* 2021; 42: 1066–1078

Tricklebank MD, Robbins TW, Simmons C, Wong EHF. Time to re-engage psychiatric drug discovery by strengthening confidence in preclinical psychopharmacology. *Psychopharmacology* 2021; 238: 1417–1436

Trubetsky V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, Bryois J, Chen CY, Dennison CA, Hall LS, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* 2022; 604: 502–508

Turner TN, Wilfert AB, Bakken TE, Bernier RA, Pepper MR, Zhang Z, Torene RI, Retterer K, Eichler EE. Sex-Based Analysis of *De Novo* Variants in Neurodevelopmental Disorders. *Am J Hum Genet* 2019; 105: 1274–1285

Uffelmann E, Huang QQ, Munung NS, Vries J de, Okada Y, Martin AR, Martin HC, Lappalainen T, Posthuma D. Genome-wide association studies. *Nat Rev Methods Primers* 2021; 1: 59

van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al. From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013; 43: 11.10.1-11.10.33

Vanmarsenille L, Giannandrea M, Fieremans N, Verbeeck J, Belet S, Raynaud M, Vogels A, Ounap K, Jacqueline V, Briault S, et al. Increased Dosage of RAB39B Affects Neuronal Development and Could Explain the Cognitive Impairment in Male Patients with Distal Xq28 Copy Number Gains. *Hum Mutat* 2014; 35: 377–383

Wierenga LM, Ruigrok A, Aksnes ER, Barth C, Beck D, Burke S, Crestol A, van Drunen L, Ferrara M, Galea LAM, et al. Recommendations for a Better Understanding of Sex and Gender in the Neuroscience of Mental Health. *Biol Psychiatry Glob Open Sci* 2024; 4: 100283

Wray NR, Gottesman II. Using summary data from the Danish National Registers to estimate heritabilities for schizophrenia, bipolar disorder, and major depressive disorder. *Front Genet* 2012; 3: 1–12

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

## 2. Veröffentlichung

Dieser Publikationsdissertation liegt die folgende, unabhängig begutachtete Veröffentlichung zugrunde:

Claus I, Sivalingam S, Koller AC, Weiß A, Mathey CM, Sindermann L, Klein D, Henschel L, Ludwig KU, Hoffmann P, Heimbach A, Heilmann-Heimbach S, Vedder H, Kammerer-Ciernioch J, Stürmer T, Streit F, Maaser-Hecker A, Nenadić I, Baune BT, Hartmann AM, Konte B, Giegling I, Heilbronner U, Wagner M, Philipsen A, Schmidt B, Rujescu D, Bunness A, Schulze TG, Rietschel M, Forstner AJ, Nöthen MM, Degenhardt F. Contribution of Rare and Potentially Functionally Relevant Sequence Variants in Schizophrenia Risk-Locus Xq28,distal. *Am J Med Genet B Neuropsychiatr Genet* 2025; 198: e33011

<https://doi.org/10.1002/ajmg.b.33011>

### **3. Erklärung zum Eigenanteil**

Diese Arbeit wurde im Institut für Humangenetik unter der Leitung von Prof. Dr. med. Markus M. Nöthen durchgeführt.

Die Konzeption der Studie erfolgte gemeinsam mit der ehemaligen Betreuerin PD Dr. med. Franziska Degenhardt.

Die Datenerhebung und experimentellen Versuche wurden nach einer kurzen Einarbeitung durch verschiedene Mitarbeitende des Instituts für Humangenetik überwiegend durch mich durchgeführt mit Ausnahme a) der Rekrutierung der Studienteilnehmenden und b) der letzten Schritte der Sequenzierung, welche durch die NGS Core Facility der Medizinischen Fakultät der Universität Bonn durchgeführt wurden.

Die statistische Auswertung erfolgte in enger Zusammenarbeit mit der Core Unit for Bioinformatics Analysis (CUBA) der Medizinischen Fakultät der Universität Bonn unter Beteiligung von Dr. Sugirthan Sivlingam und Andreas Bunes.

Die Interpretation der Ergebnisse erfolgte federführend durch mich mit Unterstützung der bisher namentlich erwähnten Personen und Prof. Dr. med. Andreas J. Forstner.

Bei der Erstellung dieser Arbeit verwendete ich die professionelle Version von QuillBot zur Rechtschreibprüfung und die frei verfügbare Version von DeepL und ChatGPT, um die Lesbarkeit und Sprache des Manuskripts zu überprüfen. Die Vorschläge von ChatGPT habe ich kritisch bewertet und die entsprechenden Passagen ggf. überarbeitet und übernehme die volle Verantwortung für den Inhalt der veröffentlichten Dissertationsschrift.

Ich versichere, die Dissertationsschrift selbständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

## 4. Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen, den Personen, aber insbesondere allen Ärztinnen und Ärzten, zu danken, die mich maßgeblich auf meinem beruflichen und privaten Lebensweg begleitet haben.

Mein ganz besonderer Dank gilt meinem Doktorvater, Herrn Prof. Dr. med. Markus M. Nöthen, Prof. Dr. med. Andreas Forstner und Dr. rer. nat. Per Hoffmann, die mit fachlichen Diskussionen und Ihrer Unterstützung während der Erstellung und Überarbeitung des Manuskriptes sowie dieser Arbeit einen maßgeblichen Anteil daran hatten, das Ziel dieses Weges zu erreichen.

Daneben danke ich meiner ehemaligen Doktormutter PD Dr. med. Franziska Degenhardt dafür, dass sie mir damit den Weg in die faszinierende Welt der Humangenetik eröffnet hat.

Ein großes Dankeschön gilt außerdem allen weiteren Kolleginnen und Kollegen am Institut für Humangenetik sowie bei Life&Brain; gerade an diejenigen, die nicht mit auf die Autorenliste aufgenommen werden konnten und mir doch mit der Beantwortung einer scheinbar noch so belanglosen Frage auf diesem Weg geholfen haben. Auch möchte ich allen Patientinnen und Patienten sowie allen Kontrollpersonen danken, die Teil dieser Studie waren. Ebenso gilt mein Dank den klinischen Kollaborationspartnerinnen und Kollaborationspartnern, die die Rekrutierung der Studienteilnehmenden durchgeführt haben.

Vielen Dank sagen möchte ich ebenfalls meinen lieben Freundinnen Dr. Lisa Sindermann, Tina Künneth und Katrin Nisters.

In tiefer Dankbarkeit werde ich auch immer Herrn Prof. Dr. med. Rüdiger Krauspe verbunden sein.

Zuletzt möchte ich mich von ganzem Herzen bei meiner Mutter Dr. med. Babette Claus, und meinen Großeltern Dr. med. Hannelore Albert und Prof. Dr. med. Lothar Albert für die unendliche Unterstützung bedanken. Dieser Dank gilt gleichermaßen meinem Vater Uwe Claus.