

Aus der Lübecker Interdisziplinären Plattform für Genomanalytik | LIGA

Leitung: Prof. Dr. med. Lars Bertram

**Die Identifizierung von Risiko-modifizierenden Varianten mittels
genomweiter Technologien für die isolierte Dystonie**

Inauguraldissertation zur Erlangung der Doktorwürde der Universität zu Lübeck

-Aus der Sektion Medizin-

Vorgelegt von Olena Ohlei

aus Kolybajewka, Ukraine

Lübeck, den 16. November 2021



UNIVERSITÄT ZU LÜBECK

1. Berichterstatter: Prof. Dr. med. Lars Bertram
2. Berichterstatter: Prof. Dr. Bernhard Haubold

Tag der mündlichen Prüfung: 17.05.2022

Zum Druck genehmigt. Lübeck, den 17.05.2022

-Promotionskommission der Sektion Medizin-

Inhaltsverzeichnis

INHALTSVERZEICHNIS.....	III
ABBILDUNGSVERZEICHNIS.....	VI
TABELLENVERZEICHNIS	VII
ABKÜRZUNGSVERZEICHNIS	VIII
1 EINLEITUNG	1
1.1 Klinische Aspekte der Dystonie	1
1.1.1 Klinische Präsentation und Symptomatik	1
1.1.2 Klassifikation und Epidemiologie	1
1.1.2.1 Fokale Dystonie	3
1.1.2.2 Segmentale Dystonie	6
1.1.2.3 Multifokale Dystonie	7
1.1.2.4 Generalisierte Dystonie.....	7
1.2 Genetische Aspekte der Dystonie.....	7
1.2.1 Einführung in die Methoden der Genomforschung.....	7
1.2.2 Monogene Formen der Dystonie.....	8
1.2.3 Polygene Formen der Dystonie.....	9
1.3 Methoden zur Identifizierung von genetischen Risikofaktoren	10
1.3.1 Genetische Assoziationsanalysen	10
1.3.1.1 Kandidatengen-basierte Assoziationsstudien	11
1.3.1.2 Genomweite Assoziationsstudien (GWAS)	12
1.3.2 Genomweite Assoziationsstudien (GWAS) zur Dystonie	14
1.4 Wissenschaftliche Fragestellung / Hypothesen	15
2 MATERIAL UND METHODEN.....	17
2.1 Erstellung einer Field-Synopsis zum Stand genetischer Assoziationsstudien der Dystonie	17
2.1.1 Literatursuche, Ein- und Ausschlusskriterien.....	17
2.1.2 Eingabe der Daten und die Qualitätskontrolle	19
2.1.3 Metaanalyse eingeschlossener Daten	19
2.2 Durchführung von Genomweiten Assoziationsanalysen in Dystonie-Datensätzen	22
2.2.1 Verwendete Datensätze und Einbindung in die DFG-Forschungsgruppe „ProtectMove“	23
2.2.2 Genomweite Genotypisierungsexperimente mittels „Global Screening Array“	25
2.2.3 Analysen der autosomalen Chromosomen.....	26
2.2.3.1 Prozessierung und Imputation der Rohgenotypdaten	26
2.2.3.2 Vorbereitende Analysen und Datenfilterungen	26
2.2.3.3 Durchführung der GWAS und GWAS-Metaanalysen	27
2.2.4 Analyse des X-Chromosoms.....	29
2.2.4.1 Prozessierung und Imputation der Rohgenotypdaten auf dem X-Chromosom.....	29
2.2.4.2 Vorbereitende Analysen und Datenfilterungen für das X-Chromosom	30

2.2.4.3	Durchführung von Assoziationsanalysen und Metaanalysen am X-Chromosom.....	31
2.2.5	Post-GWAS-Analysen	31
2.2.5.1	Bestimmung, funktionelle Zuordnung und Annotation genetischer Loci	32
2.2.5.2	Gen-basierte GWAS-Analysen	32
2.2.5.3	Gen-Property Analysen	33
2.2.5.4	Heritabilitätsbestimmungen	33
2.2.5.5	Korrelationsanalysen mit anderen komplex-genetischen Phänotypen	34
2.3	Vergleich der neuen GWAS-Ergebnisse mit denen der Field-Synopsis bereits publizierter Daten .	34
3	ERGEBNISSE	35
3.1	Field-Synopsis zum Stand genetischer Assoziationsstudien der Dystonie	35
3.1.1	Literatursuche, Ein- und Ausschlusskriterien	36
3.1.2	Eingabe der Daten und Qualitätskontrolle	36
3.1.3	Metaanalyse eingeschlossener Daten	36
3.1.4	Zusammenfassung der Ergebnisse.....	40
3.2	Durchführung von genomweiten Assoziationsstudien in Dystonie-Datensätzen	41
3.2.1	Vorliegende Datensätze und Einbindung in die DFG-Forschungsgruppe „FOR2488“	41
3.2.2	Genomweite Genotypisierungsexperimente mittels „Global Screening Array“	41
3.2.3	Analysen der autosomalen Chromosomen.....	42
3.2.3.1	Qualitätskontrolle der Genotypisierungsdaten und Genotyp-Imputation	42
3.2.3.2	Vorbereitende Analysen und Datenfilterungen	42
3.2.3.3	Durchführung der GWAS und der GWAS-Metaanalysen	44
3.2.4	Analyse des X-Chromosoms.....	53
3.2.4.1	Qualitätskontrolle der Genotypisierungsdaten und Genotyp-Imputation	53
3.2.4.2	Vorbereitende Analysen und Datenfilterungen	54
3.2.4.3	Durchführung der XWAS und XWAS-Metaanalysen	55
3.2.5	Post-GWAS-Analysen	56
3.2.5.1	Bestimmung, funktionelle Zuordnung und Annotation genetischer Loci	56
3.2.5.2	Gen-basierte GWAS-Analysen	57
3.2.5.3	Gen-Property-Analysen	59
3.2.5.4	Heritabilitätsbestimmungen	61
3.2.5.5	Korrelationsanalysen mit anderen komplex-genetischen Phänotypen	63
3.3	Vergleich der neuen GWAS-Ergebnisse mit denen der Field-Synopsis	66
4	DISKUSSION	69
4.1	Field-Synopsis	69
4.1.1	Zusammenfassung	69
4.1.2	Einordnung.....	70
4.2	GWAS der isolierten Dystonie und klinischen Subgruppen.....	71
4.2.1	Zusammenfassung	71
4.2.2	Einordnung.....	73
4.3	Korrespondenz der Ergebnisse aus beiden Teilen der Arbeit	76
5	AUSBLICK	78

5.1	Genomweite Analyse weiterer Phänotypen der Dystonie	78
5.2	Genomweite Analyse in Mutationsträgern.....	78
5.3	Genomweite Analysen anderer genetischer Varianten	79
6	ZUSAMMENFASSUNG	80
	LITERATURVERZEICHNIS	81
	ANHANG A	VIII
	DANKSAGUNG.....	XLI

Abbildungsverzeichnis

Abbildung 1.1 Beispielhafte Darstellung von Blepharospasmus..	4
Abbildung 1.2 Beispielhafte Darstellung von Meige-Syndrom.....	6
Abbildung 1.3 Eine Beziehung zwischen den Allelfrequenzen und dem Krankheitseffekt.	11
Abbildung 1.4 Beispiel einer genomweiten Assoziationsstudie mit einem Fall-Kontroll-Design.	13
Abbildung 1.5 Beispielhafte Darstellung einer indirekten Assoziation.	14
Abbildung 3.1 Flussdiagramm der Literatursuche, Datenextraktion und Analysestrategien, die für die Field-Synopsis angewendet wurden.....	35
Abbildung 3.2 Plot der Eigenvektoren der Hauptkomponenten 1 und 2 der PCA-Analysen.	43
Abbildung 3.3 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der isolierten Dystonie.....	46
Abbildung 3.4 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der kranialen Dystonie.....	49
Abbildung 3.5 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der Dystonie der oberen Extremitäten.....	51
Abbildung 3.6 Genlocus-Zoom Plot der neu entdeckten Region auf chr. 6p21.33 für die kraniale Dystonie..	57
Abbildung 3.7 Manhattan- und QQ-Plots der genomweiten Assoziationsergebnisse auf Gen-Ebene mit der segmentalen und multifokalen Dystonie.	58
Abbildung 3.8 Gewebeexpressionsanalyse mit 54 Gewebearten aus GTEx-Projekt (v8) und isolierte Dystonie, fokale Dystonie und zervikale Dystonie.	60

Tabellenverzeichnis

Tabelle 2.1 Allelverteilung unter der Annahme eines additiven genetischen Modells.....	20
Tabelle 3.1 Ergebnisse von Random-Effekt Metaanalysen auf der Basis eines additiven Modells bei allen Arten von isolierter (auch bekannt als "primärer") Dystonie.....	37
Tabelle 3.2 Ergebnisse von Random-Effekt Metaanalysen auf der Basis eines additiven Modells in diagnostischen Subgruppen der isolierten Dystonie.....	39
Tabelle 3.3 Demografische Merkmale von Dystonie-Patienten und Kontrollpersonen, die in diese Studie einbezogen wurden.....	41
Tabelle 3.4 Überblick über die deutschen und amerikanischen Datensätze, die in die GWAS-Analysen einbezogen wurden.....	44
Tabelle 3.5 Eine Übersicht der genomweit signifikanten <i>lead</i> -SNPs ($p < 5 \times 10^{-8}$) und Gene ($p < 2,6 \times 10^{-6}$) im genomischen Locus aus den Metaanalysen zur kranialen Dystonie und zur Dystonie der oberen Extremitäten sowie segmentalen und multifokalen Dystonie.....	52
Tabelle 3.6 Ergebnisse der SNP-Level-QC für die genotypisierten SNPs auf dem X-Chromosom.....	53
Tabelle 3.7 Anzahl der Proben deutscher und amerikanischer Herkunft, welche die QC nach der Imputation bestanden haben und welche in der XWAS und in den Metaanalysen eingeschlossen sind.....	54
Tabelle 3.8 Ergebnisse der SNP-Level-QC für die imputierten SNPs auf dem X-Chromosom basierend auf den Daten aus den Gruppen der deutschen und amerikanischen Studienteilnehmern.....	54
Tabelle 3.9 Ergebnisse der SNP-basierten Heritabilitätsanalysen.....	62
Tabelle 3.10 Ergebnisse der genetischen Korrelationsanalysen für die isolierte Dystonie, die fokale Dystonie und zervikale Dystonie.....	65
Tabelle 3.11 Vergleich der Ergebnisse der Dystonie Metaanalysen aus der Field-Synopsis mit den Ergebnissen der GWAS, welche in dieser Arbeit erstellt wurde.....	68
Tabelle 4.1 Power-Schätzungen für die in diesem Bereich durchgeführten isolierten Dystonie Metaanalysen. Die Stichprobe umfasst 2838 Dystonie Patienten und 5212 gesunde Probanden.....	74

Abkürzungsverzeichnis

- BMI Körpermasseindex (engl. *body mass index*)
- CNVs Kopienzahlvariation (engl. *copy number variants*)
- D' der „Lewontin's“ LD-Koeffizient
- DeNoPa Eigenname der „De Novo Parkinson“ Studie
- DNA Desoxyribonukleinsäure
- DYSTRACK Eigenname der „Dystonia Translational Research and Therapy Consortium“ Studie
- EPIPARK Eigenname der „Epidemiologie nicht-motorischer Symptome beim Parkinson-Syndrom“ Studie
- FDR Falschfindungsrate (engl. *False Discovery Rate*)
- FUMA engl. *Functional Mapping and Annotation*
- GCTA engl. *Genome-wide Complex Trait Analysis*
- GDAC engl. *Genome Data Analysis Core Unit*
- GHC engl. *Group Health Cooperative*
- GSA engl. *Global Screening Array*
- GWAS genomweite Assoziationsstudie
- H₀ Nullhypothese
- H₁ Alternativhypothese
- h² SNP-basierte Heritabilität
- HGNC HUGO Gene Nomenclature Committee
- HRC engl. *Haplotype Reference Consortium*
- HWE Hardy-Weinberg-Equilibrium
- indel-SNP SNP mit inserierten oder deletierten Allelen
- kb Kilobasenpaar
- KI Konfidenzintervall
- LD Kopplungsungleichgewicht (engl. *linkage disequilibrium*)
- LDSC LD-Score-Regression
- LIGA Lübecker Interdisziplinäre Plattform für Genomanalytik
- lncRNA Lange nicht-kodierende RNA (engl. *long non-coding RNA*)
- MAF Frequenz des seltenen Allels (engl. *minor allele frequency*)
- Mb Megabasenpaare
- MDD Schwere depressive Störung (engl. *Major depressive disorder*)
- MDS Akronym der „International Parkinson and Movement Disorder Society“
- MHC Haupt-Histokompatibilitäts-Komplex (engl. *major histocompatibility complex*)

n Anzahl

NCBI Akronym der „National Center for Biotechnology Information“

ncRNA Nichtcodierende Ribonukleinsäure

NHW Personen europäischer Abstammung (engl. *non-Hispanic Whites*)

OR Quotenverhältnis (engl. *odds ratio*)

P6-Projekt Teilprojekt der FOR2488-Forschungsgruppe; Projekttitle: Identifizierung von Penetranz- und Risiko-modifizierenden Varianten in Dystoniegenen mittels genomweiter Technologien

PAR Pseudoautosomale Regionen

PCA Hauptkomponentenanalyse (engl. *principal component analysis*)

PEG Eigenname der „Parkinson's Environment and Gene“ Studie

ProtectMove Akronym der FOR2488-Forschungsgruppe; voller Titel: Reduced penetrance in hereditary movement disorders

QC Qualitätskontrolle

QQ-Plot Quantil-Quantil-Plot (engl. *quantile-quantile-plot*)

QTL engl. *quantitative trait locus*

r^2 LD-Korrelationskoeffizient

RE-Modell Random-Effekt Modell

SFB Sonderforschungsbereich

SNP DNA-Einzelbasenaustausche (engl. *single nucleotide polymorphism*)

WTCCC Akronym der „Wellcome Trust Case Control Consortium“

XDP X-linked-dystonia-parkinsonism

XWAS X-Chromosom-weite Assoziationsanalyse

α Signifikanzschwelle

λ genomische Inflationsfaktor *Lambda*

1 Einleitung

Die vorliegende Arbeit befasst sich mit der Identifizierung von genetischen Risikofaktoren, die sowohl die isolierte Dystonie als auch deren klinischen Subgruppen beeinflussen. Auf genetischer Ebene wurden hierbei hauptsächlich verschiedene genomweite Assoziationsstudien (GWAS) durchgeführt. In diesem Kapitel werden die genetischen Aspekte der monogenen und polygenen Formen der Dystonie beschrieben. Im Anschluss werden die Methoden zur Identifizierung von genetischen Risikofaktoren näher erläutert und die wissenschaftliche Fragestellung formuliert.

1.1 Klinische Aspekte der Dystonie

Unter Dystonie fasst man ein breites Spektrum an Bewegungsstörungen zusammen. Diese weisen eine Fehlfunktion bei der Bewegungs- und Haltungskontrolle auf. Tremor und Myoklonie-ähnliche Bewegungen können bei der Dystonie als ein Symptom auftreten. Der Begriff Dystonie kann dabei drei verschiedene Entitäten beschreiben: ein Symptom, eine eigenständige Krankheit oder ein Syndrom (sekundäre Dystonie) im Rahmen einer anderen Grunderkrankung (1). Für die in dieser Arbeit durchgeführten Analysen werden ausschließlich Fälle von isolierter Dystonie berücksichtigt. Dabei handelt es sich um Dystonieformen, bei denen Dystonie (neben Tremor) das einzige Symptom der Bewegungsstörung ist (2). Im Nachfolgenden wird der Begriff „Dystonie“ daher synonym zu „isolierte Dystonie“ verwendet. Formen der (isolierten) Dystonie wurden für die Analysen dieser Arbeit weiterhin folgendermaßen unterteilt: 1) fokale Dystonie (d.h. zervikal, laryngeal, kranial sowie die Dystonie der oberen und unteren Extremitäten), 2) segmentale Dystonie, 3) multifokale Dystonie und 4) generalisierte Dystonie (2).

1.1.1 Klinische Präsentation und Symptomatik

Die Dystonie zählt zu einer Gruppe der Bewegungsstörungen, bei denen es sich um neurologische Syndrome mit einem Übermaß oder einem Mangel an willentlichen Bewegungen handelt. Die klinische Symptomatik der Dystonie umfasst anhaltende, unwillkürliche Muskelkontraktionen, die häufig zu verzerrenden und sich wiederholenden Bewegungen oder abnormalen Körperhaltungen führen (3). Dystonien werden oft durch freiwillige Aktionen ausgelöst oder verschlimmert und sind mit einer überschüssigen Muskelaktivierung assoziiert (2).

1.1.2 Klassifikation und Epidemiologie

Die Klassifikation der Dystonie ist in fünf klinische Merkmale und anhand ihrer Ätiologie unterteilt (2). Die klinischen Merkmale sind:

- Alter zum Beginn der Erkrankung

- Symptomverteilung
- Zeitlicher Verlauf
- Koexistenz anderer Bewegungsstörungen
- Andere neurologische Manifestationen

Das Alter zum Beginn der Erkrankung wird in folgende Gruppen aufgeteilt: die frühkindliche Altersgruppe (0 – 2 Jahre), die kindliche Altersgruppe (3 - 12 Jahre), die jugendliche Altersgruppe (13 - 20 Jahre), die frühe erwachsene Altersgruppe (21 - 40 Jahre) oder die späte erwachsene (> 40 Jahre) Altersgruppe.

Symptome der Dystonie treten im Körper in der Schädelregion, der Halsregion, am Kehlkopf, am Rumpf, an den oberen oder unteren Gliedmaßen auf. Diese verschiedenen Gebiete können einzeln oder in verschiedenen Kombinationen betroffen sein. Die Verteilung der Symptome kann sich im Verlauf der Erkrankung auf bisher unbeteiligte Körperregionen ausbreiten. Dabei werden fokale (eine Körperregion betroffen), segmentale (zwei oder mehr zusammenhängende Körperregionen betroffen), multifokale (zwei oder mehr nicht zusammenhängende Körperregionen betroffen), generalisierte (Rumpf und mindestens zwei weitere Körperregionen betroffen) Dystonien sowie die Hemidystonie (mehrere betroffene Körperregionen auf derselben Körperseite) unterschieden.

Der zeitliche Krankheitsverlauf kann entweder statisch oder progressiv sein. Zudem wird das zeitlich auftretende Erkrankungsmuster in der tätigkeitsspezifischen Dystonie unterschieden, die nur bei bestimmten Tätigkeiten wie der Musikerdystonie oder dem Schreibkrampf auftreten. Des Weiteren werden die Erkrankungsmuster der Dystonie in die täglichen Fluktuationen (zum Beispiel die dopa-responsive Dystonie) und das paroxysmale Muster aufgeteilt. Das paroxysmale Muster bedeutet ein plötzliches Auftreten der Symptome, die in der Regel durch einen bestimmten Auslöser hervorgerufen werden (2).

Dystonien können isoliert oder in Kombination mit anderen Bewegungsstörungen auftreten. Bei der isolierten Form tritt die Dystonie als einziges motorisches Symptom auf, mit der Ausnahme des Tremors (2). Die Dystonie in Kombination mit anderen Bewegungsstörungen wie Parkinsonismus oder Myoklonus wird als kombinierte Dystonie klassifiziert (2). Außerdem kann die Dystonie mit anderen neurodegenerativen oder systemischen Erkrankungen kombiniert auftreten und wird dann als komplexe Dystonie bezeichnet (2,4).

Die zweite Achse der Dystonie Klassifikation befasst sich mit der Ätiologie, die z. T. noch nicht vollständig geklärt ist. Es wird zwischen der Pathologie des Nervensystems und vererbbarer sowie der erworbenen Dystonie unterschieden (2). Die vererbte Dystonie basiert auf nachgewiesenen monogenen genetischen Mutationen mit folgenden Vererbungsmuster: autosomal dominant,

autosomal rezessiv, X-chromosomal rezessiv und mitochondrial (5). Die Dystonie kann auch durch nicht-genetische Ursachen ausgelöst werden. Dazu gehören perinatale Hirnverletzungen, Infektionen (virale Hirnhautentzündungen, HIV, Tuberkulose, Syphilis und andere), Medikamente, Toxine (Mangan, Kobalt, Schwefelkohlenstoff und andere), vaskuläre Ursachen (Ischämie, Hämorrhagie und arteriovenöse Malformation), neoplastische Ursachen (Hirntumor und paraneoplastische Enzephalitis), Hirnverletzungen (Schädeltrauma, Gehirnoperationen und elektrische Verletzung) und psychische Ursachen (2). Des Weiteren kann die Dystonie sowohl „sporadisch“ entstehen, d.h. aufgrund einer bisher unbekanntem Ursache. Es wird davon ausgegangen, dass ein nicht unerheblicher Teil der sporadischen Dystonie durch die kombinierten Effekte von genetischen Faktoren (mit-)verursacht wird (5). Daher wird diese Form der Dystonie in dieser Arbeit als polygene Dystonie bezeichnet, um sie von denen monogenen Formen der Erkrankung (s.o.) abzugrenzen.

In dieser Arbeit geht es um die Identifikation neuer genetischer Risikofaktoren der sporadischen/polygenen Dystonie unter Verwendung genomweiter Assoziationsanalysen. Hierfür wurden verschiedene Datensätze von deutschen und amerikanischen Probanden und dazugehörigen Kontrollpersonen im Rahmen der DFG-Forschungsgruppe FOR2488 „ProtectMove“ (www.protect-move.de) untersucht. Die Prävalenz der isolierten Dystonie liegt in der deutschen Bevölkerung bei ca. 142 pro eine Million Einwohner (6) und in der amerikanischen Bevölkerung bei ca. 329 pro eine Million Einwohner (7).

1.1.2.1 Fokale Dystonie

Die fokale Form der isolierten Dystonie ist durch eine übermäßige und unwillkürliche Aktivität der Muskulatur nur auf eine Körperregion begrenzt. Typische Beispiele für Körperregionen sind Spasmen der Nackenmuskeln (zervikale Dystonie), Kehlkopfmuskeln (laryngeale Dystonie) und Augenringmuskeln (Blepharospasmus). Insgesamt werden fünf Formen der fokalen Dystonie in dieser Arbeit untersucht: die zervikale Dystonie, die laryngeale Dystonie, die kraniale Dystonie, die aufgabenspezifische Dystonie und die Dystonie der oberen Extremitäten. Die fokale Dystonie ist die häufigste Form der isolierten Dystonie mit der Prävalenz von ca. 101 pro eine Million Einwohner in Deutschland (6) und ca. 295 pro eine Million Einwohner in den USA (7).

Zervikale Dystonie (*Torticollis spasmodicus*)

Die zervikale Dystonie ist eine der häufigsten Formen der fokalen isolierten Dystonie (6). Frauen erkranken häufiger als Männer, das Geschlechtsverhältnis ist 1.6:1 (8). Die klinischen Symptome sind durch übermäßige und unwillkürliche tonische oder phasische Spasmen der Hals- und Nackenmuskulatur gekennzeichnet, die zu abnormalen Bewegungen des Kopfes und

Nackenschmerzen führen. Zusätzlich zur anormalen Kopfposition kann es zu zitternden Bewegungen oder krampfhaftem Zucken des Kopfes kommen. Die Kopfbewegungen werden in den Torticollis (rechts/links Drehung), der Laterocollis (die Neigung zu einer Seite), der Retrocollis (Rückwärtsbeugung nach oben) und der Anterocollis (nach vorn auf die Brust gebeugte Bewegung) unterschieden (8–10).

Laryngeale Dystonie

Die laryngeale Dystonie ist durch eine abnorme Aktivität der Kehlkopfmuskeln (innerer Larynxmuskeln) gekennzeichnet. Die häufigste Form der laryngealen Dystonie ist die spasmodische Dysphonie (10). Dabei handelt es sich um eine seltene Sprachstörung, die sich in der Mitte des Lebens entwickelt. Die Symptome der spasmodische Dysphonie sind ein unkontrollierter Abbruch der Stimme und eine ausgeprägte Anstrengung beim Sprechen. Die Progression ist im ersten Jahr langsam und bleibt danach chronisch. Die spasmodische Dysphonie gehört zu der aufgabenspezifischen Dystonie, bei der Spasmen typischerweise während des Sprechens auftreten aber mit singen, schreien, weinen oder flüstern nachlassen (10,11). Weniger häufig tritt eine Überaktivität der Stimmlippenabduktionsmuskeln auf. In seltenen Fällen können beide Muskelgruppen betroffen sein. Die laryngeale Dystonie kann zusammen mit einem Tremor der Stimme auftreten. Die spasmodische Dysphonie und der Tremor sind schwer voneinander zu unterscheiden. Weitere weniger verbreitete Formen der laryngealen Dystonie sind die laryngeale Atmungsdystonie (12) und die Sängerdystonie (13).

Kraniale Dystonie

Die kraniale Dystonie ist durch abnorme und unwillkürliche Bewegungen verschiedener Gesichtsmuskeln, Kaumuskeln, Zungenmuskeln sowie der Kehlkopfmuskeln und Rachenmuskeln



Abbildung 1.1 Beispielhafte Darstellung von Blepharospasmus. a) Blepharospasmus vom Lidöffnungsinhibitionstyp (Frontalis-Muskel Kontraktionen). Die Patientin kann die Augenöffnung nicht kontrollieren. b) Klassischer Blepharospasmus mit kräftigem Liedschluss. Aus Ceballos-Baumann A, 2018, Abb.1 (15).

gekennzeichnet. Die häufigste fokale Form der kranialen Dystonie ist der Blepharospasmus

(Abbildung 1.1). Diese Dystonieform ist durch unwillkürliche Spasmen der Augenringmuskeln (Musculus orbicularis oculi) gekennzeichnet. Der Muskel umgreift ringförmig das Auge und dient dem Schluss der Lidspalte. Der Blepharospasmus verursacht eine unkontrollierte Schließung der Augen (Lidkrampf) und in schwerwiegenden Fällen eine funktionelle Blindheit. (14).

Ein weiteres Beispiel der kranialen Form ist die oromandibuläre Dystonie. Dabei treten Kontraktionen der Muskeln des Kiefers und des Mundes auf. Die Spasmen der unteren kranialen Muskeln verursachen unwillkürliches schließen und öffnen des Kiefers oder eine Kombination beider Bewegungen. Es kann auch zu einer seitlichen Kieferabweichung, Grimassen und Zungenverkrümmung kommen. Sprechen, Kauen und Schlucken sind oft gestört. Die Symptome entstehen weitgehend ohne eine spezifische Überlastung der Muskeln. Die oromandibuläre Dystonie kann aber auch aufgabenspezifisch sein, zum Beispiel durch Spielen auf einer Flöte (14, 15).

Dystonie der oberen Extremitäten

Die Dystonie der oberen Extremitäten ist durch den Verlust der feinmotorischen Steuerung der Muskeln im Arm, Unterarm und in der Hand charakterisiert (16, 17). Zusätzlich kann ein Tremor vorliegen. Die Dystonie der oberen Extremitäten kann aufgabenspezifisch sein, d.h. wenn die Dystonie nur eine Aufgabe betrifft, wie z.B. das Schreiben oder das Spielen eines Musikinstruments. Die Aufgabenspezifität kann jedoch mit der Zeit nachlassen und während unspezifischer Handlungen oder spontan auftreten. Die anatomischen Grundlagen dieser Dystonieform sind bisher kaum bekannt. Im Allgemeinen wird davon ausgegangen, dass die Dystonie der oberen Extremitäten auf eine Störung der Basalganglien zurückzuführen ist. Neuere Studien bei der sporadischen Dystonie der oberen Extremitäten haben jedoch gezeigt, dass auch der sensomotorische Kortex, der Thalamus und das Kleinhirn Veränderungen aufweisen (16).

Aufgabenspezifische Dystonie

Die aufgabenspezifische Dystonie gehört zu der fokalen Form der isolierten Dystonie und stellt den Verlust der feinmotorischen Steuerung dar, die auf die Ausführung einer spezifischen feinmotorischen Tätigkeit (z.B. schreiben, Instrument spielen) beschränkt ist. Der Tremor und die abnorme Körperhaltung können begleitende Symptome sein. Die Symptome treten typischerweise zwischen dem dritten und dem sechsten Jahrzehnt des Erwachsenenalters auf (17). Im Gegensatz zu den anderen Dystonieformen sind Männer von dieser Form häufiger betroffen (17). Zu Beginn der Erkrankung können die Symptome unauffällig und schmerzlos sein. Im Laufe der Zeit breiten sich die Symptome auf andere Aufgaben aus oder greifen sogar auf zuvor nicht betroffene Körperbereiche über und können zu einer anormalen Körperhaltung führen (17, 18). Eine Fallstudie

beschreibt den Verlust der feinmotorischen Steuerung beim wiederholten Schreiben nur eines einzigen Buchstabens oder einer Zahl in einer Stresssituation (19). Weitere Studien berichten über einen professionellen Flamenco-Tänzer, der von der aufgabenspezifischen Dystonie der unteren Extremitäten betroffen war (20). Die aufgabenspezifische Dystonie wurde bei Musikern, Schriftstellern, Sportlern, Frisören, Malern, Schustern und Schneidern beschrieben. Die Musikerdystonie und der Schreibkrampf sind die häufigsten Formen der aufgabenspezifischen Dystonie (17, 18).

1.1.2.2 Segmentale Dystonie

Die segmentale Form der isolierten Dystonie ist durch übermäßige und unwillkürliche Aktivität der Muskulatur von zwei oder mehr benachbarten Bereichen des Körpers charakterisiert. Ein Beispiel für die segmentale Form der kranialen Dystonie ist das Meige-Syndrom (2). Beim Meige-Syndrom handelt es sich um eine Kombination von Spasmen des Augenringmuskels (Blepharospasmus) und der unteren kranialen Muskeln (oromandibuläre Dystonie). Möglicherweise hielt Pieter Brueghel schon im 16. Jahrhundert anhand eines Porträts (Abbildung 1.2) das Krankheitsbild fest (14, 15). Ein weiteres Beispiel der segmentalen Dystonie ist eine Kombination von Spasmen der Armmuskulatur zusammen mit der Rumpfmuskulatur.



Abbildung 1.2 Beispielhafte Darstellung von Meige-Syndrom. Das Meige-Syndrom ist eine Kombination des Blepharospasmus und der oromandibulärer Dystonie. Das Gemälde „Le bailleur“ (der Gähner) aus dem Musée des Beaux Arts, Brüssel stellt vielleicht die erste Darstellung der Meige-Dystonie dar. Aus Ceballos-Baumann A, 2018, Abb.2 (15).

1.1.2.3 Multifokale Dystonie

Die multifokale Form der isolierten Dystonie ist durch übermäßige und unwillkürliche Aktivität der Muskulatur in zwei oder mehr nicht benachbarten Körperbereichen charakterisiert, zum Beispiel die unteren kranialen Muskeln zusammen mit den Muskeln der unteren Extremität (2).

1.1.2.4 Generalisierte Dystonie

Die generalisierte Form der isolierten Dystonie ist durch die übermäßige und unwillkürliche Aktivität der Muskulatur des Rumpfes und der Muskulatur zweier weiterer Körperbereiche charakterisiert. Hierbei wird zwischen den generalisierten Formen mit und ohne Beteiligung der Beine unterschieden (2).

1.2 Genetische Aspekte der Dystonie

Eine familiäre Häufung der Dystonie nach klassischen Vererbungsmustern haben schon frühzeitig eine genetische Ursache zumindest bei einigen Dystonieformen nahegelegt (21). Die technologischen Fortschritte der Genomforschung in den letzten zwei Jahrzehnten haben zu einem Anstieg der Anzahl von Genen geführt, die als assoziiert mit Dystonien beschrieben wurden (22). Dies betrifft sowohl die Gene der monogenen und polygenen isolierten Dystonie als auch von Dystonieformen, die Kombination mit anderen Bewegungsstörungen (Parkinsonismus oder Myoklonus) auftreten (23, 24). Insbesondere bei den genetischen Ursachen der isolierten Dystonie gibt es in der Literatur vergleichsweise wenige Studien, die sich in ihren Ergebnissen teilweise widersprechen (25).

1.2.1 Einführung in die Methoden der Genomforschung

Als Genom wird die Gesamtheit der Erbinformationen einer Zelle bezeichnet. Beim Menschen sind die genetischen Informationen in stabilen doppelsträngigen Desoxyribonukleinsäure (DNA)-Molekülen auf 23 Chromosom-Paaren gespeichert. Die DNA eines Menschen ist etwa 3 Mrd. Nukleotide lang (26). Einzelne oder mehrere Basen der DNA können aufgrund von Umwelteinflüssen, wie radioaktiver Strahlung, oder infolge von Fehlern der DNA-Replikation, dauerhaft verändert, d.h. mutiert werden. Eine Neumutation ist bei einer üblichen Spontanmutationsrate von 1 pro 10^5 - 10^6 Basenpaaren zwar sehr unwahrscheinlich, dennoch hat sich über die Jahrtausende eine Vielzahl solcher DNA-Einzelbasenaustausche im Genom des Menschen manifestiert. In der Keimbahn manifestierte Mutationen werden innerhalb von Familien nach den Mendelschen Regeln vererbt. Falls die Mutationen in der untersuchten Population eine Frequenz von mehr als 1 % aufweisen, werden diese als *Polymorphismus* bezeichnet. Bei vielen Polymorphismen handelt es sich um DNA-Einzelbasenaustausche (engl. *single nucleotide*

polymorphism, [SNP]). In Abbildung 1.4A werden zwei SNPs auf Chromosom 9 von drei Personen bildlich dargestellt. Jede Person besitzt zwei Chromosomen. Dabei wurde ein Chromosom von der Mutter und das andere vom Vater vererbt. Die Angabe von einer Base auf einem Strang ist ausreichend, um ein Allel darzustellen. Der erste SNP hat zwei mögliche Allele G und T. Die Genotypen dieser SNPs können dann G/G, G/T, T/T sein. Das Verständnis des Zusammenhangs von DNA-Veränderungen auf bestimmte Krankheitsbilder ist eine der wichtigsten Aufgaben der Humangenetik, sowohl in der klinischen Patientenversorgung als auch in der Grundlagenforschung (26).

Bei monogenen Erkrankungen verursacht eine Mutation innerhalb eines Gens oder anderen Stelle im Genom die Krankheit. Solche Mutationen sind typischerweise sehr selten (<< 1% Frequenz in der Normalbevölkerung) und werden innerhalb der betroffenen Familien nach dominanten oder rezessiven Mustern vererbt. Bei der polygenen Vererbung reicht das Vorliegen einer singulären Mutation (i. d. R. eines Polymorphismus) in einem einzelnen Gen nicht aus, die Krankheit zu verursachen. Erst wenn viele genetische Varianten typischerweise an unterschiedlichen Positionen im Genom in Kombination ihre Wirkung entfalten, kann das Krankheitsrisiko beeinflusst werden. Dabei repräsentiert ein einzelner DNA-Einzelbasenaustausch nur einen kleinen Teil des Gesamtrisikos für die Ausbildung eines Merkmals.

1.2.2 Monogene Formen der Dystonie

Die monogenen Formen der Dystonie folgen den klassischen Regeln der Mendelschen Vererbung (d.h. dominant oder rezessiv) innerhalb betroffener Familien. Dabei ist die Krankheitsursache typischerweise eine einzige Mutation, die häufig in einem funktionellen Abschnitt eines Gens (einem sog. Genlocus) lokalisiert ist. Bisher wurden in der einschlägigen Literatur 28 genetische Loci mit „DYT“ Bezeichnungen in Verbindung mit monogenen Formen der Dystonie beschrieben (27). Allerdings warten viele von diesen Loci noch auf eine Validierung in unabhängigen Datensätzen. Kürzlich wurde eine neue Nomenklatur für genetisch-bedingte Bewegungsstörungen mit insgesamt 32 Genen für die isolierten, kombinierten und komplexen Dystonien von einer Arbeitsgruppe der International Parkinson and Movement Disorder Society (MDS) vorgeschlagen (23). Die in der Literatur bestätigten und etablierten Gene der isolierten Dystonie sind *DYT-TOR1A* (28), *DYT-THAP1* (29) und *DYT-GNAL* (30). Diese monogenen Formen werden autosomal dominant vererbt. Darüber hinaus wurden auch Mutationen in *DYT-ANO3* bei der kraniozervikalen Dystonie berichtet (31). Die Krankheitsrelevanz der DNA-Mutationen in *DYT-ANO3* bleibt allerdings schwierig zu interpretieren, weil nicht alle *ANO3*-Varianten zur Erkrankung beitragen und auch neue DNA-Mutationen innerhalb dieses Gens bei der generalisierten Dystonie identifiziert wurden (32).

Das erste in der Dystonie berichtete Gen, *DYT-TOR1A*, befindet sich auf Chromosom 9q34. Hier handelt es sich nicht um einen Einzelbasenaustausch, sondern um eine Deletion von drei Nukleotiden (c.904_906delGAG; p.302delGlu) innerhalb von *TOR1A* (28). Diese Veränderung verursacht in den meisten Fällen eine generalisierte Dystonie und beginnt im frühen Alter (< 20 Jahre), allerdings können die Dystonieformen von früher generalisierter Dystonie bis spät-einsetzender fokaler Dystonie variieren. Nur 30% aller Mutationsträger sind tatsächlich an der Dystonie erkrankt, d.h. diese genetische Variante zeigt eine reduzierte Penetranz (24,27).

Viele weitere Mutationen, z. B. in *DYT-THAP1*, *DYT-GNAL*, *DYT-ANO3*, *DYT-HPCA*, *DYT-KCTD17*, *DYT-KMT2B*, *DYT-PRKRA* und *DYT-SGCE*, sind als monogene Ursachen der isolierten Dystonie in der Mutationsdatenbank der Gesellschaft für Bewegungsstörungen (MDSGene.org) zu finden, die unter Federführung von Mitgliedern unserer Arbeitsgruppe LIGA ins Leben gerufen wurde (33). Die meisten dieser Mutationen sind allerdings sehr selten bis einmalig und deswegen schwer in unabhängigen Datensätzen replizierbar. Insofern muss die Evidenz für viele der in den o.g. Genen beschriebenen Mutationen derzeit noch als vorläufig angesehen werden.

1.2.3 Polygene Formen der Dystonie

Die Theorie der polygenen Entstehung sogenannter komplex-genetischer Erkrankungen besagt, dass viele genetische Loci für sich genommen nur einen kleinen Risikoeffekt haben und erst in Kombination miteinander zu einer nennenswerten Beeinflussung des Krankheitsrisikos führen (34). Häufig spielen hierbei auch nicht-genetische Faktoren, z. B. Epigenetik, Umwelteinflüsse und Lebensgewohnheiten, eine wichtige Rolle. Die bisher publizierten Studien zu polygenen Formen der Dystonie sind hauptsächlich Kandidatengen-Studien. Bei diesem Ansatz konzentrieren sich die Studien auf die Gene, welche einen (vermeintlichen) Bezug zur Erkrankung haben, also z. B. Gene, welche die monogene Dystonie verursachen oder auf die funktionellen Arbeiten in Bezug auf die Dystonie hinweisen. Z. B. wird vermutet, dass eine Störung im Dopaminstoffwechsel eine Ursache der komplex-genetischen Dystonie sein kann, daher wurden die Gene *GCH1*, *DR1*, *DRD2* und *DRD3* auf Assoziation mit dem Risiko für eine Dystonie getestet (22).

Andere Dystonie Kandidatengene sind *TOR1A* (35,36) und *BDNF* (37,38), für die in der Literatur widersprüchliche Ergebnisse veröffentlicht wurden. Eine Übersichtsarbeit zu *TOR1A* berichtete über einen signifikanten Gesamteffekt der SNPs rs1182 und rs1801968 auf die Entwicklung der fokalen Dystonie und des Schreibkrampfs (36). Eine andere Studie zu *TOR1A* fand dagegen keine Assoziation zwischen diesen beiden SNPs und isolierter Dystonie im Erwachsenenalter (35). Ähnlich wie bei den *TOR1A*-Genvarianten zeigte auch beim *BDNF*-Gen eine Studie eine DNA-Variante als Risikofaktor bei der isolierten Dystonie (38), während eine andere Studie keinen solchen Effekt fand

(37). Ein weitaus aussagekräftigeres Verfahren zur Genidentifizierung bei komplex-genetischen Erkrankungen ist die Durchführung sog. genomweiter Assoziationsstudien (GWAS; im nächsten Abschnitt genauer beschrieben). Zum Zeitpunkt der Durchführung dieser Arbeit gab es in der Dystonie-Literatur nur drei solcher GWAS: eine zur Musikerdystonie (39) und zwei zur zervikalen Dystonie (40,41). Diese werden in Abschnitt 1.3.2. näher beschrieben.

1.3 Methoden zur Identifizierung von genetischen Risikofaktoren

Wie oben bereits ausgeführt basieren komplex-genetische Erkrankungen auf den kombinierten Effekten von Umwelteinflüssen, Lebensgewohnheiten und genetischen Risikofaktoren. In Bezug auf Letztere ergibt sich somit ein individuelles Risikoprofil, das die Summe der Effekte aller individuell vorliegenden Allele an relevanten Genloci darstellt (42). Zur Identifizierung von genetischen Risikofaktoren kommen – analog zum Ansatz der klassischen Epidemiologie zur Identifizierung nicht-genetischer Risikofaktoren – häufig Assoziationsanalysen zum Einsatz. Hier wird der statistische Zusammenhang zwischen der Prävalenz eines vermuteten Risikofaktors in Bezug auf den Krankheitsstatus untersucht. In der Humangenetik gibt es hierfür verschiedene Ansätze. Nachfolgend soll auf die zwei am häufigsten verwendeten, d.h. der Kandidatengen-Ansatz und der genomweite Ansatz, näher eingegangen werden.

1.3.1 Genetische Assoziationsanalysen

Als genetische Assoziation wird allgemein die Beziehung zwischen einem Phänotyp und einem DNA-Polymorphismus bezeichnet. Als DNA-Sequenzvariationen werden hierbei oft SNPs verwendet (43). Bei genetischen Assoziationsanalysen wird die Häufigkeit der an einem SNP-Locus auftretenden Allele auf Assoziation mit einem Merkmal getestet. In seiner einfachsten Form ist das Merkmal dichotom ausgeprägt, z. B. in Kontext von Fall/Kontroll-Studien, so dass die Verteilung der Allelfrequenzen eines SNPs in gesunden und erkrankten Probanden getestet wird. Dabei wird als Effektstärke häufig das Quotenverhältnis (engl. *odds ratio*, [OR]) eines Allels oder Genotyps in Bezug auf die Erkrankung ermittelt.

Die Beziehung zwischen Allel- oder Genotypfrequenzen und der Effektstärke eines SNPs (OR) in Kontext verschiedener genetischer Assoziationsanalysen wird in der Abbildung 1.3 dargestellt. Genetische Effekte in dem oberen rechten Bereich innerhalb der Abbildung erfordern aufgrund ihrer großen Effektstärke nur eine Genotypisierung von relativ kleinen Stichprobengrößen, sind aber insgesamt eher selten. Effekte in dem unteren rechten Bereich der Abbildung sind typisch für die Befunde von polygenen Erkrankungen, die große Stichprobengrößen und eine Vielzahl von Tests an genetischen Markern erfordern. Effekte im unteren linken Bereich der Abbildung sind die große Herausforderung, da sie für einen Nachweis einer Assoziation sehr große Stichprobenzahlen ($>>1$

Million) erfordern (43). Genloci im oberen linken Areal von Abbildung 1.3 bilden den generellen Bereich der monogenen Krankheitsgene ab, die z. B. durch Kopplungsanalysen in familienbasierten Studien identifiziert werden können.

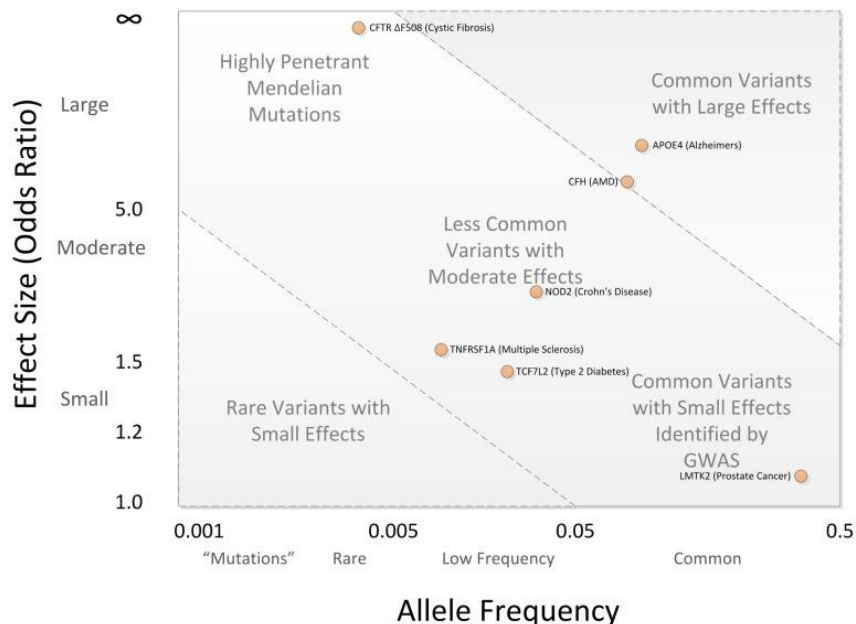


Abbildung 1.3 Eine Beziehung zwischen den Allelfrequenzen und dem Krankheitseffekt. Auf der Abszisse werden die Allelfrequenzen (eine Mutation, sehr selten, selten und häufig) und auf der Ordinate die Effektgröße (klein, moderat und groß) abgebildet. Oben links werden sehr seltene Allele mit großen Effektstärken dargestellt. Diese werden bei monogenetischen Erkrankungen mit der Hilfe von Kopplungsanalysen identifiziert. Unten rechts werden die meisten GWAS-Befunde mit häufigen SNPs, die kleine Effektstärken haben, dargestellt. Der Großteil der entdeckten genetischen Assoziationen liegt auf der Diagonale, welche durch die gestrichelten Linien angezeigt wird. Aus Bush WS und Moore JH, 2012, Abb. 1 (43)

1.3.1.1 Kandidatengen-basierte Assoziationsstudien

Kandidatengene werden typischerweise auf der Grundlage ihrer (vermeintlichen bzw. erwiesenen) funktionellen Relevanz für die Pathogenese der Krankheit ausgewählt. Ein weiteres Auswahlkriterium für ein Kandidatengen ist eine spezifische Region auf dem Chromosom, die in den Kopplungsstudien bereits beschrieben werden (44) oder Varianten in monogenen Loci (s.o).

Seit dem 1 August 2001 wurden mehr als Vierzig solcher Kandidatengen-Studien in der Dystonie mit z. T. widersprüchlichen Ergebnissen (siehe Abschnitt 1.2.3.) veröffentlicht. Einer von vielen möglichen Gründen für die beobachteten widersprüchlichen Ergebnisse ist die vergleichsweise kleine Stichprobengröße und/oder ein (zu) gering angesetztes Alpha-Niveau ($p < \alpha$) was beides zu höheren falsch-positiven Ergebnissen führen kann. Eine weitere Limitation ist die Auswahl nur weniger Kandidatengene, die von der Hypothese und der Datenlagen zu Beginn der Studie abhängen. Typischerweise wird in Kandidatengen-Studien nur eine sehr geringe Markerdichte in

den Kandidatengen untersucht. Viele der o.g. Limitationen des Kandidatengen-Ansatzes können mit Hilfe des GWAS-Ansatzes gelöst werden.

1.3.1.2 Genomweite Assoziationsstudien (GWAS)

Im Gegensatz zu Kandidatengen-Studien wird bei einer genomweiten Assoziationsstudie (GWAS) keine Vorauswahl der getesteten DNA-Varianten (oder Gene) getroffen, sondern eine große Anzahl an Markern (typischerweise SNPs, deren Anzahl je nach verwendeter Technologie zwischen ~200K und ~5M liegen kann) getestet, die mehr oder weniger gleichmäßig über das gesamte Genom verteilt sind (daher „genomweit“) (40, 41, 43). Wegen der fehlenden Vorauswahl den Phänotyp betreffend wird der GWAS-Ansatz häufig auch als „hypothesenfrei“ bezeichnet. Außer der eigentlichen, hypothesenfreien Assoziationsanalyse erlaubt der GWAS-Ansatz viel weitreichendere Möglichkeiten in Bezug auf die Qualitätskontrolle (QC) der verwendeten genetischen Marker und DNA-Proben, auf die in diesem Abschnitt noch weiter eingegangen wird. Für viele Fragestellungen der Genidentifikation bei komplex-genetischen Phänotypen ist der GWAS-Ansatz mittlerweile die Methode der Wahl (43).

Abbildung 1.4 stellt den generellen Ablauf einer GWAS dar. Zuerst werden die SNPs bei jedem Studienprobanden genomweit auf den Chromosomen 1-22 und X genotypisiert, typischerweise mithilfe eines kommerziellen Mikroarrays. Nach ausführlicher QC auf SNP- und Sample-Ebene (s.u.) erfolgt die eigentliche Testung auf Assoziation zwischen den einzelnen SNPs und der Krankheit / Phänotyp auf der Grundlage der Prävalenz der einzelnen SNPs in den verfügbaren Proben. Im Kontext von Fall-Kontroll-Studien (Abbildung 1.4B) wird die Stärke der Assoziation häufig anhand der OR gemessen und die statistische Evidenz dieser Assoziation mittels Teststatistik und p-Wert quantifiziert. Die logarithmierten (zur Basis 10) p-Werte der getesteten SNPs werden danach häufig in Form eines sog. Manhattan-Plots grafisch dargestellt (Abbildung 1.4C). Dabei gilt: je höher die Position eines Punktes auf der Ordinate in der Grafik lokalisiert ist, desto niedriger ist der dazugehörige p-Wert und desto stärker ist die statistische Evidenz für die Assoziation.

Bei der Interpretation von GWAS-Resultaten muss beachtet werden, dass die Ergebnisse lediglich die Assoziationsevidenz quantifizieren. Sie können generell aber keine direkten Aussagen zu möglichen kausalen Zusammenhängen treffen. Der mit der Erkrankung assoziierte SNP kann zwar eine funktionelle, krankheitsverursachende Variante sein (in diesem Fall spricht man von „direkter Assoziation“), er kann aber auch eine nicht-funktionelle Variante darstellen („indirekte Assoziation“). Die indirekte Assoziation kann z. B. durch Kopplungsungleichgewicht (engl. *linkage disequilibrium*, [LD]) mit dem krankheitsverursachenden SNP verursacht sein. Das LD beschreibt welcher Anteil eines Allels eines SNPs mit einem Allel eines anderen SNPs innerhalb einer

Population zusammen vererbt wird (43). Das Vorhandensein von LD führt zu mehreren möglichen positiven Assoziationsergebnissen innerhalb desselben genetischen Locus (siehe Abbildung 1.4C). Ein Modell der indirekten Assoziation ist in Abbildung 1.5 dargestellt. Ob eine direkte oder indirekte Assoziation vorliegt, kann im Kontext einer GWAS allein nicht abschließend geklärt werden. Hierfür bedarf es weiterführender Testungen („Post-GWAS-Analysen“), z. B. unter Zuhilfenahme weiterer genomweiter Datensätze oder funktioneller Vorhersagen (*in silico*) oder mittels Durchführung

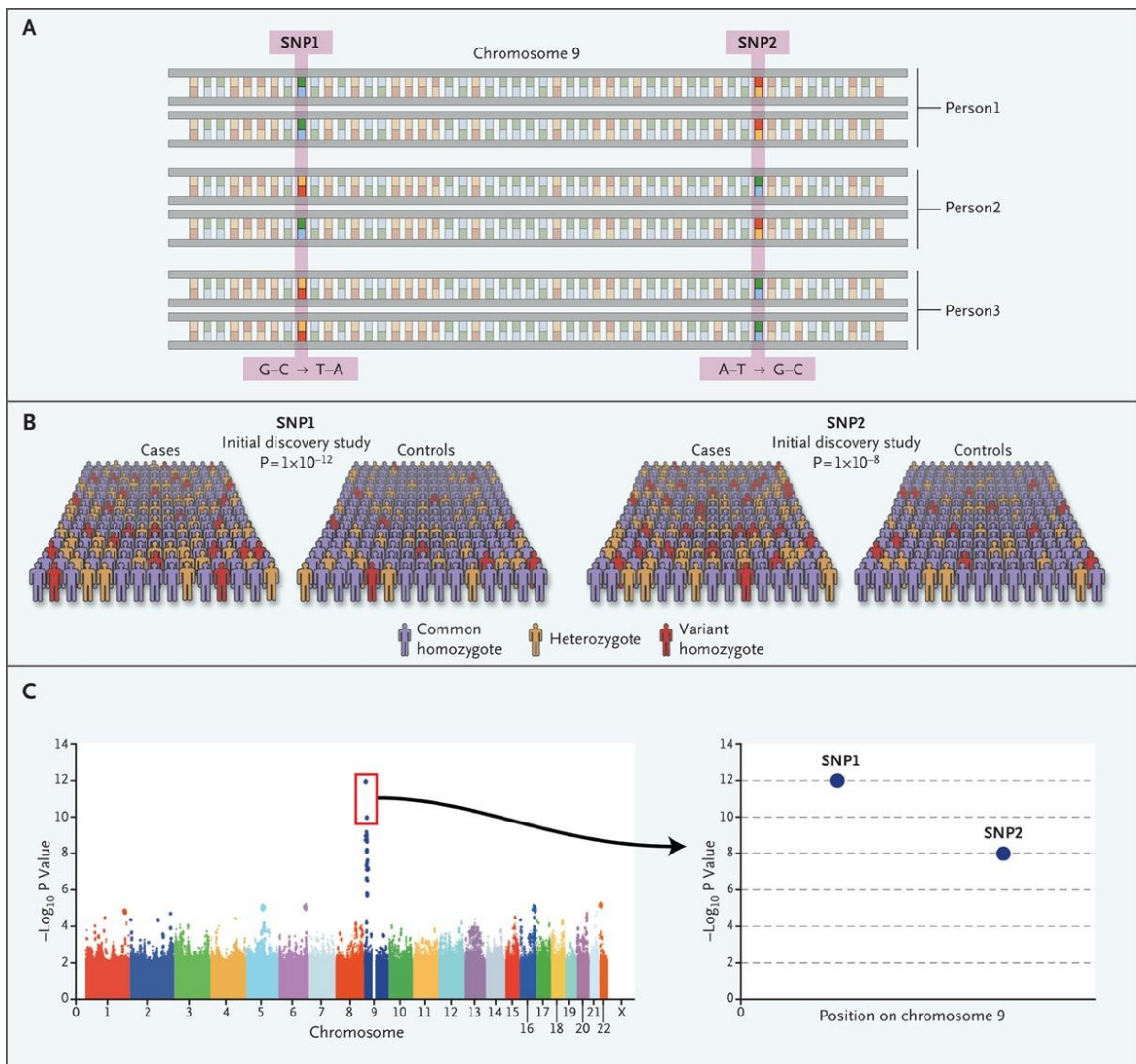


Abbildung 1.4 Beispiel einer genomweiten Assoziationsstudie mit einem Fall-Kontroll-Design. Die genomweite Assoziationsstudie basiert typischerweise auf einem Fall-Kontroll-Design, bei dem Einzel-Nukleotid-Polymorphismen (SNPs) im gesamten menschlichen Genom untersucht werden. A) zeigt einen kleinen Genlocus auf Chromosom 9 und damit ein sehr kleines Fragment des Genoms. B) Hier wird die Assoziation zwischen den einzelnen SNPs und der Krankheit auf der Grundlage der Prävalenz der einzelnen SNPs in Fall- und Kontrollgruppe getestet. In diesem Beispiel sind die SNPs 1 und 2 mit p-Werten von 10^{-12} bzw. 10^{-8} mit der Krankheit assoziiert. C) Die Punkte in einem Manhattan-Plot stehen für die $-\log_{10}(P)$ -Werte für alle auf eine Assoziation getesteten SNPs. Jedes Chromosom wird in einer anderen Farbe dargestellt. Aus Manolio TA, 2010 Abb. 1 (115).

gezielter funktioneller Experimente (*in vitro* oder *in vivo*) mit dem Ziel, die molekularen Mechanismen und funktionellen Ursachen der statistischen Assoziation zwischen DNA-Variante und Phänotyp aufzuklären.

Indirect Association

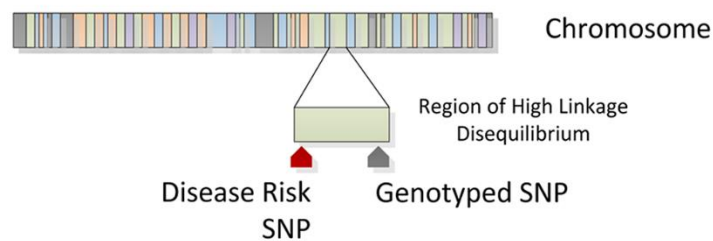


Abbildung 1.5 Beispielhafte Darstellung einer indirekten Assoziation. Der genotypisierte SNP verursacht die Krankheit nicht, sondern ein anderer SNP, der in einem hohen Kopplungsungleichgewicht mit dem genotypisierten SNP liegt. Somit ist der genotypisierte SNP indirekt mit der Erkrankung statistisch assoziiert. Aus Bush WS und Moore JH, 2012, Abb. 3 (43).

Bei der GWAS werden üblicherweise Hunderttausende bis mehrere Millionen SNPs auf eine Assoziation mit einem Merkmal hin getestet. Die kumulative Wahrscheinlichkeit ein falsch positives Ergebnis bei der Assoziationsanalyse zu finden ist damit sehr viel höher als bei der Durchführung nur eines Tests, z. B. im Kontext einer Kandidatengen-Studie. Um die Anzahl an falsch positiven Ergebnissen bei der GWAS zu begrenzen, muss ein α -Niveau (Typ-I-Fehlerrate) definiert werden, das diesem Problem des multiplen Testens gerecht wird und die studienspezifische Irrtumswahrscheinlichkeit bei $\sim 5\%$ hält. Hierfür hat sich in der Fachwelt ein α von 5×10^{-8} als genomweite Signifikanzgrenze etabliert (43). Diese Zahl basiert auf der Anzahl unabhängiger genomischer Loci im human Genom und repräsentiert damit die effektive Anzahl statistischer Tests, die im Kontext einer GWAS durchgeführt und dann mittels der Bonferroni-Methode korrigiert werden (siehe Abschnitt 2.1.3). Für die europäische Bevölkerung gibt es ca. 1.000.000 solcher unabhängiger genomischer Loci (43) für SNPs mit $> 5\%$ Allelhäufigkeit, d.h. die Bonferroni-Korrektur für ein studienspezifisches $\alpha = 0,05 / 1.000.000 = 5 \times 10^{-8}$.

Das letztendliche Ziel einer GWAS besteht darin, das gewonnene Wissen zu genetischen Risikofaktoren zu nutzen, um Krankheitsvorhersagen zu treffen, die biologischen Grundlagen der Krankheitsanfälligkeit zu identifizieren und um neue Präventions- und Behandlungsstrategien zu entwickeln (43).

1.3.2 Genomweite Assoziationsstudien (GWAS) zur Dystonie

Im Dystonie-Feld sind zum Zeitpunkt der Durchführung dieser Arbeit nur drei GWAS vorbeschrieben: eine zur Musikerdystonie (39) und zwei zur zervikalen Dystonie (40, 41). Bei der GWAS zur Musikerdystonie wurden 243 Patienten mit 1.097 Kontrollen metaanalysiert. Genomweite Assoziationsbefunde wurden hierin nur mit der Variante rs11655081 in dem Arylsulfatase G (*ARSG*) Gen auf Chromosom 17 festgestellt. Weitere Analysen mit anderen

klinischen Dystonie-Gruppen ergaben keine Assoziationen mit diesem SNP (25). Bis zur Schriftlegung dieser Arbeit (Juni 2021) gibt es keine unabhängige Replikation der GWAS-Ergebnisse (außer einer Studie mit sehr kleiner Fallzahl [n=111] (45)) und nach meinem Kenntnisstand auch keine unabhängige Studie, die die einen funktionellen Zusammenhang des *ARSG*-Gens und Dystonie hergestellt hätte. In der zweiten Dystonie-GWAS wurden 212 Patienten mit zervikaler Dystonie und 5.173 Kontrollen vom „*Wellcome Trust Case Control Consortium*“ (WTCCC) untersucht. Die stärksten Assoziationen wurden hier mit den Varianten rs9416795 in einem nicht codierenden DNA-Bereich auf Chromosom 10 und dem rs1338041 SNP innerhalb des *NALCN* Gens festgestellt. *NALCN* kodiert Proteine für den *Sodium Leak Channel, non-selective* und wird allgemein im Gehirn und Rückenmark exprimiert (40). Bei einem ersten unabhängigen Replikationsversuch der Ergebnisse zur zervikalen Dystonie GWAS konnten allerdings keine Assoziationen mit den SNPs in *NALCN* nachgewiesen werden (46). In der zweiten zervikalen Dystonie-GWAS wurden 919 Patienten und 1.491 Kontrollen untersucht (41). Genomweite Assoziationsbefunde wurden hierin mit der Variante rs2219975 aufwärts vom Kollagen Typ VIII Alpha 1-Kette (*COL8A1*) Gen auf dem Chromosom 3 berichtet. Die Assoziation wurde in einer unabhängigen Stichprobe mit 473 Patienten mit zervikaler Dystonie und 481 Kontrollen nicht repliziert (41). Eine zusätzliche genbasierte Analyse in dieser Studie ergab, dass das *DENND1A*-Gen (DENN domain containing 1A) signifikant mit zervikaler Dystonie assoziiert ist (41).

1.4 Wissenschaftliche Fragestellung / Hypothesen

Wie oben ausgeführt stellt die Dystonie eine komplex-genetische Erkrankung dar, für die derzeit nur unzureichende molekulargenetische Grundlagen ermittelt wurden. Das Hauptziel dieser Arbeit war die Durchführung verschiedener GWAS-Analysen zur Identifizierung von genetischen Varianten oder Genen, die das Risiko an einer Dystonie zu erkranken beeinflussen. In der Literatur gibt es zwar einige Kandidatengenstudien zu dieser Fragestellung, allerdings waren diese *qua* Design auf sehr wenige genomische Regionen beschränkt und testeten vergleichsweise kleine Fallzahlen. Nicht zuletzt deswegen haben die bisher durchgeführten genetischen Assoziationsstudien zur Dystonie nur sehr widersprüchliche und sehr schwer interpretierbare Ergebnisse hervorgebracht.

Das Hauptziel dieser Arbeit war die Identifizierung und Charakterisierung von (neuen) DNA-Varianten und Genloci, die eine genetische Assoziation mit dem Risiko für eine isolierte Dystonie zeigen. Hierfür wurden folgende drei übergeordnete Fragestellungen bearbeitet:

1. Durchführung einer systematischen Zusammenstellung und Metaanalysen aller bis dato publizierten genetischen Assoziationsstudien zur isolierten Dystonie.

2. Durchführung von GWAS-Analysen zur Identifizierung neuer genetischer Varianten, die mit dem Risiko einer isolierten Dystonie und deren Subformen assoziiert sind.
3. Abgleich der neuen GWAS-Ergebnisse aus #2 mit den Befunden der systematischen Metaanalysen bereits publizierter Daten aus #1.

Herzstück der Arbeit ist die Durchführung neuer GWAS-Analysen (Fragestellung #2). Dafür wurde der weltweit aktuell größte Dystonie-Datensatz mit 2.838 Fällen und 5.212 Kontrollen mittels Assoziationsanalysen von ca. acht Millionen SNPs ausgewertet. Aus diesem sehr umfassenden Datensatz ergaben sich darüber hinaus weitere Analysemöglichkeiten, z. B. Durchführung der weltweit ersten Dystonie-Assoziationsanalyse von Markern auf dem X-Chromosom, SNP-basierte Heritabilitätsanalysen sowie die Berechnung von genetischen Korrelationen zwischen den Dystonie-GWAS-Ergebnissen und den GWAS-Ergebnissen anderer komplex-genetischer Erkrankungen. Zum Zeitpunkt der Schriftlegung dieser Arbeit sind die Ergebnisse der 1. Fragestellung bereits in der Fachzeitschrift *Parkinsonism and Related Disorders* nach Peer-Review publiziert (25). Die Publikation der Ergebnisse der anderen Fragestellungen ist derzeit in Vorbereitung.

2 Material und Methoden

Im folgenden Kapitel werden die Methodik der Erstellung einer Field-Synopsis zum Stand genetischer Assoziationsstudien der Dystonie, die Durchführung der genomweiten Assoziationsanalysen und die dazugehörigen Metaanalysen beschrieben. Abschnitt 2.1 behandelt die Methoden der Literaturrecherche und der Metaanalyse, der in der Literatur berichteten Daten. Die Methodik zur Durchführung der genomweiten Assoziationsanalysen in den hier verfügbaren Dystonie-Datensätzen finden sich im Abschnitt 2.2. Im Abschnitt 2.3 beschreibe ich die Methodik des Vergleichs der neuen GWAS-Ergebnisse mit denen aus der Field-Synopsis.

2.1 Erstellung einer Field-Synopsis zum Stand genetischer Assoziationsstudien der Dystonie

Um den Stand der Forschung für die genetischen Assoziationsstudien in isolierter Dystonie zu ermitteln, wurde eine systematische Erhebung der im Feld publizierten Daten (engl. *field synopsis*; nachfolgend hier auch so bezeichnet) erstellt (Ohlei et al., 2018, (25)). Die Methoden zur Durchführung dieser Studie basieren auf dem zuvor in unserer Gruppe entwickelten Ansatz der Erstellung von Field-Synopsen der Alzheimer Erkrankung (47) und der Parkinson Erkrankung (48). Im Rahmen der Dystonie Field-Synopsis habe ich eine systematische Literaturrecherche durchgeführt, die sich inhaltlich weitestgehend nach den Empfehlungen der PRISMA („Preferred reporting items for systematic reviews and meta-analyses“) Arbeitsgruppe zur Durchführung und Beschreibung systematischer Reviews orientiert (49). Außerdem habe ich eine Datenbank mit Microsoft Excel erstellt und die statistische Analyse der in der Literatur berichteten genetischen Daten in PLINK-Software (v1.9) (50, 51) durchgeführt. Eine Zusammenfassung des Ablaufs und der verschiedenen Schritte der Field-Synopsis finden sich in Abbildung 3.1 (reproduziert aus Ohlei et al, 2018).

2.1.1 Literatursuche, Ein- und Ausschlusskriterien

Die systematische Literaturrecherche basiert auf einer breit angelegten Suche in der NCBI „PubMed“ Datenbank nach allen bis zum 23. Januar 2018 zu genetischen Assoziationsstudien der isolierten Dystonie veröffentlichten Publikationen. Für die Literatursuchewurde im ersten Schritt nach den folgenden Schlüsselwörtern gesucht:

„(dystoni* OR (writer* AND cramp) OR graphospasm OR blepharospasm OR torticollis OR meige OR spasmodic OR dysphonia OR retrocollis OR antecollis OR laterocollis OR (musician* AND cramp) OR (occupation* AND cramp) OR (golf* AND cramp) OR yips) AND associat* AND gene*“

In der zweiten Phase der Literatursuche wurden getrennt nach jedem in Phase 1 identifizierten Gen und nach insgesamt 28 weiteren Genloci mit einer „DYT“ Bezeichnung gesucht (d.h. DYT1-21 und DYT23-28, DYT22 wurde nicht inkludiert, weil es noch nie mit einem Locus, Gen oder klinischen Syndrom in Verbindung gebracht worden ist) (23). Die DYT-Bezeichnungen wurden aus der Literatur übernommen (23, 52) und repräsentieren Genloci, die als mögliche Ursache der monogenen Dystonie in der Literatur beschrieben werden (siehe Einleitung Abschnitt 1.2.2.). Eine vollständige Übersicht über die Suchbegriffe wird in der Tabelle S1 in der Dystonie Field-Synopsis (Ohlei et al, 2018) dargestellt.

Die in PubMed aufgelisteten Publikationen wurden unter der Verwendung des Titels, der Zusammenfassung oder der Volltexte gelesen und nach den Ein- und Ausschlusskriterien beurteilt. Die Publikationen mussten drei Einschlusskriterien erfüllen, um in meine Field-Synopsis aufgenommen zu werden:

- (i) Die Publikationen wurden nur dann eingeschlossen, wenn diese eine Assoziation zwischen bi-allelicen genetischen Polymorphismen und einer Form der isolierten Dystonie untersuchen. Als Polymorphismus wurde ein genetischer Marker mit der Häufigkeit grösser als 0,01 des selteneren Allels in der Allgemeinbevölkerung definiert. Als Referenz für die Allel-Frequenzinformation wurden Daten des „1000 Genomes Project“ (53) genutzt.
- (ii) Veröffentlichung der Publikation in einer wissenschaftlichen Zeitschrift nach „Peer-Review“ und in englischer Sprache.
- (iii) Publikation schloss mindestens 10 Dystonie-Patienten und 10 nicht-erkrankte Probanden in die Assoziationsanalysen ein.

Ausgeschlossen wurden Studien mit familienbasierten Designs, Studien die nur Polymorphismen mit mitochondrialer DNA untersuchten, Studien mit Kontrollgenotypdaten aus öffentlichen Datenbanken, Studien über sekundäre Dystonien oder Studien an Dystonie-Patienten mit einer bekannten Dystonie-Mutation.

In vielen Publikationen auf PubMed wird die Dystonie als „primär“ klassifiziert. Primäre Dystonie ist ein Begriff, der phänomenologische und ätiologische Merkmale vermischt und dessen Verwendung seit 2013 von der Movement Disorder Society nicht mehr empfohlen wird (2). Weil der klinische Teil der Definition der "primären" Dystonie allerdings mit der Definition der "isolierten" Dystonie identisch ist, wird für alle in dieser Arbeit erwähnten Studien ausschließlich der neuere Begriff "isolierte" Dystonie verwendet, obwohl in der Primärstudie u. U. noch der alte Begriff Verwendung fand.

2.1.2 Eingabe der Daten und die Qualitätskontrolle

Aus jedem Artikel, der die Einschlusskriterien erfüllte, wurden die folgenden Informationen in eine Microsoft Excel-Datenbank extrahiert:

- Genbezeichnung
- RSID-Nummer (SNPs Identifizierungskennzeichen, das von der NCBI einer Gruppe von SNPs zugewiesen wird, die einem identischen Standort zugeordnet sind)
- Erstautor
- Publikationsjahr
- PubMed-ID (Identifikationsnummer der Publikation auf „PubMed“)
- Ethnische Abstammung der untersuchten Datensätze
- Anzahl der Fälle und Kontrollen
- Alter bei Krankheitsbeginn
- Durchschnittsalter
- Geschlechtsverteilung
- Anzahl der Allele eines Polymorphismus in den Fällen und Kontrollen
- Art der Dystonie Diagnose

Die anfängliche Dateneingabe führte ich selbst durch. Anschließend prüfte eine fachkundige Person der Arbeitsgruppe LIGA (Dr. Valerija Dobricic) alle Varianten-IDs und die Anzahl der Allele eines Polymorphismus mit den Daten der Originalpublikationen mit dem Ziel evtl. falsche Einträge zu identifizieren und zu korrigieren. Die gesamte Excel-Datenbank mit allen in die Field-Synopsis eingeschlossenen genetischen Assoziationsstudien ist auf der beigefügten *CD-ROM* abgelegt.

2.1.3 Metaanalyse eingeschlossener Daten

Für alle inkludierten Datensätze wurden die OR und die entsprechenden 95% Konfidenzintervalle (KIs) aus den Publikationen für jeden individuellen SNP wie in der Publikation berichtet übernommen. Bei den Fällen, in denen der OR in den Publikationen nicht angegeben wurde, habe ich diesen unter der Annahme eines additiven genetischen Modells anhand der Anzahl der Allele in einer Kontingenztabelle (Tabelle 2.1) berechnet (54).

Tabelle 2.1 Allelverteilung unter der Annahme eines additiven genetischen Modells. A) Anzahl der Genotypen (a/a, a/A und A/A) eines SNPs in Fall- und Kontrollgruppe. Ein Genotyp kann Homozygot (a/a und A/A) oder Heterozygot (a/A) sein. B) Anzahl der Allele eines SNP in Fall- und Kontrollgruppen. Dabei wird die Verteilung der Allele unter Annahme eines additiven genetischen Modells berechnet.

A

Allel	a/a	a/A	A/A	B
Fälle	n_{11}	n_{12}	n_{13}	
Kontrollen	n_{21}	n_{22}	n_{23}	

B

Allel	a	A
Fälle	$2*n_{11}+n_{12} = m_{11}$	$2*n_{13}+n_{12} = m_{12}$
Kontrolle	$2*n_{21}+n_{22} = m_{21}$	$2*n_{23}+n_{22} = m_{22}$

Der OR berechnet sich als:

$$OR = \frac{m_{11} \times m_{22}}{m_{21} \times m_{12}}$$

Bei den Datensätzen, in denen nur die Informationen zu den seltenen Allelenhäufigkeiten eines SNPs in Fall- und Kontrollgruppen vorlagen, wurde der entsprechende Genotyp für jedes einzelne Allel unter der Annahme eines Hardy-Weinberg-Gleichgewichts (HWE) berechnet und entsprechend in der Excel-Datenbank ergänzt, im Einklang mit dem von unserer Arbeitsgruppe bereits publiziertem Vorgehen (47, 48). HWE ist ein populationsgenetisches Modell, das eine Beziehung zwischen den beobachteten Allel- und den Genotypfrequenzen in einer Population wie folgt definiert:

$$(p + q)^2 = p^2 + 2pq + q^2 = 1$$

Bei gegebener Allelhäufigkeit des seltenen Allels p ergeben sich die Genotyp-Häufigkeiten für a/a, a/A, A/A in einer Population als p^2 , $2p(1-p)$, $(1-p)^2$. Diese Gleichung gilt unter der Annahme, dass eine große, sich zufällig paarende (d.h. keine Selektion bestimmter Genotypen), homogene Population ohne neue Mutationen oder Migrationen vorliegt (d.h. Bedingungen, von denen ausgegangen werden kann, dass sie für alle inkludierten Studien erfüllt werden).

In der Literatur, die der Field-Synopsis zugrunde lag, verwendeten vier separate Publikationen dieselben Daten von weitgehend überlappenden Datensätzen für vier SNPs (29, 55–57). In diesen Fällen wurde in dieser Arbeit nur die Anzahl der Allele/Genotypen für den größeren der beiden Datensätze berücksichtigt, d.h. die Daten aus der Referenz Lohmann et al., 2012 (55) und Newman et al., 2014 (57).

Für jede einzelne DNA-Variante mit mindestens zwei sich nicht überlappenden Datensätzen wurde eine Metaanalyse unter Verwendung des „Random-Effects“ Modells (RE-Modell) berechnet. Diese Art der Metaanalyse basiert auf dem DerSimonian- und Laird-Modell (58), das Variablen-Gewichtungen verwendet, die sowohl die Varianz innerhalb als auch zwischen den Studien berücksichtigen. Die Berechnungen wurden unter Verwendung der PLINK-Software (v1.9) (50, 51) durchgeführt.

Wann immer ausreichend viele Datensätze zur Verfügung standen, wurden die Daten für relevante SNPs nach den ethnischen Abstammungsgruppen aufgeteilt. Die Klassifizierung der ethnischen Abstammung wurde analog zu den Vorläufer-Publikationen unserer Arbeitsgruppe zur Alzheimer- und Parkinson-Krankheit (47,48) verwendet. Die Ethnizität wurde hierfür in eine von fünf Hauptgruppen eingeteilt: „afrikanische Herkunft“, „asiatische Herkunft“, „europäische Herkunft“, „hispanische Herkunft“ und „sonstige Herkunft / Mischformen“. Zusätzlich zur Berechnung von Metaanalysen über alle Formen der isolierten Dystonie wurden in der Field-Synopsis außerdem noch folgende Subformen unterschieden: fokale Dystonie (dies inkludierte die zervikale Dystonie, Blepharospasmus, Musikerdystonie, Schreibkrampf und spasmodische Dysphonie), segmentale Dystonie und generalisierte Dystonie. Wenn genügend (≥ 2) unabhängige Datensätze vorlagen, wurden die verschiedenen Formen der Dystonie weiter in Gruppen unterschiedlicher ethnischer Abstammung aufgeteilt. Insgesamt führte diese Analysestrategie zu 134 Metaanalysen von 45 Polymorphismen aus 17 Genen, die 42 unterschiedlichen Publikationen entstammten (25).

Die 45 Polymorphismen sind auf 17 unabhängige Genloci verteilt und wurden auf Assoziation mit der Dystonie innerhalb jeder Metaanalyse getestet. Dadurch fanden innerhalb der meisten jedes Genloci mehrere Tests auf statistische Signifikanz statt. In der Statistik ist dies als Problem des multiplen Testens beschrieben (59). Durch das mehrmalige Testen akkumuliert sich die Wahrscheinlichkeit ein falschpositives Ergebnis zu erhalten. Um dem entgegenzuwirken kann z. B. das studienspezifische Signifikanzniveau angehoben werden. Eine der möglichen Lösungen für das Problem des multiplen Testens ist die Korrektur des Signifikanzniveaus nach Bonferroni (Bonferroni-Korrektur) (59).

Die neue Signifikanzschwelle berechnet sich nach Bonferroni als

$$\alpha = \frac{0.05}{\text{Anzahl der unabhängigen Tests}}$$

Um die Anzahl der unabhängigen Tests (=Genloci) zu bestimmen und basierend darauf die Analysen für das multiple Testen zu korrigieren, wurde das paarweise Kopplungsungleichgewicht (LD, Erklärung siehe Einleitung Abschnitt 1.3.1.2.) der SNPs anhand eines Referenzdatensatzes ermittelt. Die Schätzungen des Kopplungsungleichgewichts über die SNP-Paare wurden von mir innerhalb desselben $\pm 1\text{Mb}$ -Intervalls im „1000 Genomes Project“ (53) mit Software PLINK-Software (v1.9) berechnet. In der Tabelle S2 der Dystonie Field-Synopsis (25) werden hierfür der LD-Korrelationskoeffizient (r^2) und der „Lewontin's“ LD-Koeffizient (D') zwischen den Paaren der relevanten Varianten dargestellt. Zur Beurteilung des LD wurde in dieser Arbeit die r^2 -Metrik verwendet, da sie bei niedrigen Allel-Häufigkeiten weniger anfällig für Verzerrungen (60) und im Gegensatz zu D' nicht von fehlenden Genotypen betroffen ist. Unter der Verwendung von $r^2 < 0,3$ als Schwellenwert wurden 32 unabhängige SNPs identifiziert; dementsprechend wurde die studienweite Signifikanzschwelle nach Bonferroni auf α von $0,05 / 32 = 1,56 \times 10^{-3}$ festgelegt. Die Heterogenität der Studien aus den Metaanalysen wurde anhand der I^2 -Metrik bewertet. Die I^2 -Werte größer als 75% weisen auf eine übermäßige Heterogenität hin (48).

Da die meisten Metaanalysen nur eine kleine Anzahl unabhängiger Datensätze kombinieren (d.h. $n \leq 3$), wird das Vorliegen eines Publikationsbias und anderer mögliche Verzerrungen nicht explizit modelliert oder bewertet (61).

2.2 Durchführung von Genomweiten Assoziationsanalysen in Dystonie-Datensätzen

Der erste Teil dieser Arbeit befasste sich nur mit der Zusammenführung bereits publizierter Befunde aus der Dystonie-Genetikliteratur. Um neue genetische Varianten und Genloci zu identifizieren, die möglicherweise die Entstehung einer Dystonie sowie deren Verlauf beeinflussen, führte ich im zweiten Teil meiner Arbeit verschiedene GWAS-Analysen unter Verwendung eines Fall-Kontroll-Ansatzes durch. Die im Rahmen dieses Teils der Arbeit zusammengeführten deutschen und amerikanischen Dystonie- und Kontrolldatensätze wurden in dieser Form bis dato noch nicht im Kontext von GWAS zur Dystonie verwendet. Analog zu der im ersten Teil dieser Arbeit durchgeführten phänotypischen Klassifikation wurden Dystonie-Fälle aus verschiedenen klinisch-diagnostischen Gruppen sowohl gemeinsam (d.h. alle Fälle der isolierten Dystonie) als auch getrennt voneinander analysiert. Die Möglichkeit der Subgruppenanalyse war von der Verfügbarkeit der Daten bestimmt und teilweise begrenzt. Es gab eine ausreichende Anzahl an

Dystonie Probanden (hier definiert als $n > 50$) für GWAS-Analysen bei den folgenden Dystonie-Gruppen (s. auch Tabelle 3.4):

- fokale Dystonie (welche sich unterteilt in zervikale Dystonie, laryngeale Dystonie [nur in US-amerikanischen Probanden], kraniale Dystonie, Dystonie der oberen Extremitäten und aufgabenspezifische Dystonie [nur in deutschen Probanden])
- segmentale und multi-fokale Dystonie
- generalisierte Dystonie (nur in US-amerikanischen Probanden)

Für alle verfügbaren klinisch-diagnostischen Gruppen führte ich die GWAS-Analysen für die 22 Autosomen und das X-Chromosom getrennt durch (s.u.).

2.2.1 Verwendete Datensätze und Einbindung in die DFG-Forschungsgruppe „ProtectMove“

Die genomweiten und klinischen Daten der Dystonie-Patienten („Fälle“) und die der nicht-erkrankten Probanden („Kontrollen“) der Analysen dieses Teils meiner Arbeit sind im Rahmen einer DFG-Forschungsgruppe gesammelt bzw. generiert worden (engl. *„Reduced penetrance in hereditary movement disorders: Elucidating mechanisms of endogenous disease protection“*; Akronym: „ProtectMove“; Förderkennzeichen: FOR2488). Die ProtectMove-Forschungsgruppe konzentriert sich auf neurologische Bewegungsstörungen einschließlich der Parkinson-Krankheit (PD), Dystonien und der X-Chromosom-gekoppelten Dystonie mit Parkinsonismus (engl. *X-linked dystonia-parkinsonism*; [XDP]). Die ProtectMove-Forschungsgruppe umfasst acht rein wissenschaftliche Projekte (P1-P8) und drei Zentralprojekte (Z1, Z2, INF). Diese Dissertation entstand im Rahmen des Teilprojekts P6 („Identifizierung von Penetranz- und Risikomodifizierenden Varianten in Dystoniegenen mittels genomweiter Technologien“; Projektleiter: Prof. L. Bertram und Prof. J. Erdmann).

Die DNA-Proben von Patienten und nicht-erkrankter Probanden setzen sich aus 11 verschiedenen Stichproben zusammen, die aus 11 Standorten in Deutschland und in den USA stammen. Die für die GWAS-Analysen verwendeten relevanten Phänotypen (z. B. Erkrankungsstatus, Dystonieform, Erkrankungsalter, Geschlecht) wurden vom Teilprojekt „Z2“ der ProtectMove-Forschungsgruppe (Projektleiter: Prof. M. Kasten, Prof. N. Brüggemann) zentral gesammelt, qualitätskontrolliert und verwaltet und schließlich an das P6-Teilprojekt zur Verwendung in den GWAS-Analysen weitergeleitet. Außer den o.g. Personen haben auch Frau Prof. Klein und Frau Dr. Eva-Juliane Vollstedt substantielle und direkte Beiträge bei der Zusammenstellung der Phänotypdaten geleistet.

Beschreibung der einzelnen Datensätze, die in die GWAS-Analysen einbezogen wurden:

DYSTRACT: Die DYSTRACT (engl. *Dystonia Translational Research and Therapy Consortium*)-Studie ist ein Netzwerk deutscher Zentren für die translationale Forschung und Behandlung von Dystonie-Erkrankungen. Das Netzwerk kombiniert klinische Studien am Menschen und experimentelle Tierstudien, um Krankheitsmerkmale auf molekularer Ebene zu charakterisieren (<http://www.dystract.cio-marburg.de/de/>). Für GWAS-Analyse steuerte DYSTRACT n=155 DNA-Proben von Dystonie-Patienten bei.

Tübingen: Die Klinik für Neurologie des Universitätsklinikums Tübingen verfügt über eine Spezialabteilung für die Diagnose und Behandlung von Patienten mit Bewegungsstörungen (<https://www.medizin.uni-tuebingen.de/de/das-klinikum/einrichtungen/kliniken/neurologie/neurodegenerative-erkrankungen>). Für die GWAS-Analysen steuerte der Standort Tübingen n=648 DNA-Proben von Dystonie-Patienten bei.

Lübeck: Die Dystonie-Fälle und -Kontrollen des Standorts Lübeck wurden in der Abteilung für Neurologie des Universitätsklinikums Schleswig-Holstein (UKSH) gesammelt und im Institut für Neurogenetik (ING) in einer Biobank gespeichert. Für die GWAS-Analysen dieses Projekts steuerte der Lübecker Standort DNA-Proben von n=480 Dystonie-Patienten und n=41 gesunden Probanden bei.

Dystonia-Coalition: Die Aufgabe der Dystonia-Coalition besteht darin, die klinische und translationale Forschung zu Dystonie-Erkrankungen zu beschleunigen, um bessere Behandlungsmöglichkeiten zu finden. Derzeit (September 2021) verwaltet die Dystonia-Coalition die weltweit größte Dystonie-Biobank, in der Blutproben und andere biologische Materialien für die Forschung gelagert werden (<https://www.rarediseasesnetwork.org/cms/dystonia/About-Us>). Für die GWAS-Analysen dieses Projekts steuerte die Dystonia-Coalition n=1.121 DNA-Proben von Dystonie-Patienten aus den USA und Deutschland bei.

DeNoPa: DeNoPa (De Novo Parkinson) Studie ist eine Langzeitstudie, welche seit 2009 in der Paracelsus-Elena-Klinik in Kassel durchgeführt wird (<https://www.denopa.de>). Ziel der DeNoPa-Studie ist es, die Parkinson-Krankheit von den ersten Symptomen in einer Langzeitperspektive zu verfolgen. Details zur Probenerhebung und Biomarker-Sammlung sind in der Literatur beschrieben (62). Für GWAS-Analyse steuerte DeNoPa n=97 DNA-Proben von gesunden deutschen Probanden bei.

SFB134: Der Sonderforschungsbereich (SFB) 134 („Essverhalten: Homöostase und Belohnungssysteme“) ist eine Studie zu Nahrungsaufnahmeverhalten in Abhängigkeit vom aktuellen Energiestatus. Für die GWAS-Analysen dieses Projekts steuerte der SFB134 n=562 DNA-Proben von neurologisch gesunden Probanden aus Lübeck bei.

PopGen: PopGen ist eine umfangreiche Biomaterial- und Datenbank für populationsbasierte und krankheitsspezifische Fragestellungen. PopGen dient der Erforschung von molekularen und nicht-molekularen Risikofaktoren für zahlreiche Zivilisationserkrankungen (z. B. Erkrankungen des Herzkreislaufsystems, entzündliche Erkrankungen, neurologische Erkrankungen oder Krebs) (63). Für die GWAS-Analysen dieser Arbeit steuerte PopGen n=3.147 DNA-Proben von gesunden Probanden aus Norddeutschland bei.

EPIPARK: Die EPIPARK (*Epidemiologie nicht-motorischer Symptome beim Parkinson-Syndrom*)-Studie untersucht die Spezifität nicht-motorischer Symptome für die Parkinson-Krankheit. Die Originaldaten der EPIPARK-Studie sind bereits publiziert (64). Für die GWAS-Analysen dieser Arbeit steuerte EPIPARK n=652 DNA-Proben von gesunden deutschen Probanden bei.

Massachusetts General Hospital, Boston, MA, USA: Für GWAS-Analyse steuerte Massachusetts General Hospital n=1.146 DNA-Proben von Dystonie-Patienten aus der USA bei.

GHC: Der GHC (engl. *Group Health Cooperative*)-Datensatz wurde zwischen 1992 und 2000 in den neurologischen und allgemeinmedizinischen Kliniken der GHC im Westen des US-Bundesstaates Washington sowie in der neurologischen Klinik der Universität Washington erstellt. Für die GWAS-Analysen dieses Projekts steuerte der GHC-Datensatz n=261 DNA-Proben von gesunden amerikanischen Probanden bei.

PEG: Probanden für die PEG (engl. *Parkinson's Environment and Gene*)-Studie wurden zwischen Januar 2001 und Dezember 2013 aus drei Orten in den USA rekrutiert. Einzelheiten zur Rekrutierung von Probanden wurden bereits in einer Publikation (65) beschrieben. Für die GWAS-Analysen dieser Arbeit steuerte die PEG-Studie n=1.054 DNA-Proben von gesunden amerikanischen Probanden bei.

Die schriftliche Zustimmung aller Studienteilnehmer wurden direkt von den o.g. Rekrutierungs-Standorten gemäß lokaler Ethikrat-Bestimmungen eingeholt und innerhalb des Z2-Projekts verwaltet. Die Ethikkommission der Universität zu Lübeck erteilte darüber hinaus die Genehmigung für die Verwendung der DNA-Proben im Kontext der ProtectMove-Forschungsgruppe (AZ 16-039; eine Kopie des Ethikvotums findet sich in Anhang, Abbildung A 33).

2.2.2 Genomweite Genotypisierungsexperimente mittels „Global Screening Array“

Zu Beginn der Genotypisierung hatte die Arbeitsgruppe LIGA Zugang zu insgesamt 3.670 DNA-Proben von Dystonie Patienten und 2.667 gesunden Probanden. Für diese Proben wurde die DNA lokal an den beteiligten Studienzentren oder am Institut für Neurogenetik extrahiert und an unser Labor gesendet. Dort führte eine fachkundige Person der Arbeitsgruppe LIGA (Dr. Valerija Dobricic) die Qualitätskontrolle (QC) mit der *PicoGreen* Quantifizierung durch. Die SNP-Genotypisierung

wurde anschließend mittels des „Global Screening Array (with shared custom-content)“ (GSA; Illumina, Inc., USA) durch das Institut für klinische Molekularbiologie (IKMB, UKSH, Campus-Kiel) auf einem iScan-Instrument (Illumina, Inc., USA) gemäß den Herstellerempfehlungen durchgeführt. Zusätzliche 3.147 DNA-Proben aus dem PopGen-Datensatz wurden direkt zur genomweiten SNP-Genotypisierung mittels des gleichen GSA-Mikroarrays an das IKMB in Kiel gesendet.

2.2.3 Analysen der autosomalen Chromosomen

2.2.3.1 *Prozessierung und Imputation der Rohgenotypdaten*

Nach Abschluss der laborexperimentellen Arbeiten der Hochdurchsatz-Genotypisierung am IKMB in Kiel wurden die Genotyprohdaten in Form sog .idat-Dateien an die Arbeitsgruppe LIGA übersendet. Alle nachfolgenden Datenverarbeitungsschritte wurden zentral im Rahmen des FOR2488-Teilprojekts INF-GDAC („Genome Data Analysis Core Unit“; Projektleiter: Prof. L. Bertram und Prof. M. Krawczak) unter Federführung des für dieses Projekt angestellten Wissenschaftlers, Dr. Fabian Kilpert, durchgeführt. Dieser Prozess umfasste die Genotyp-Bestimmung, QC sowie die Imputation nicht direkt bestimmter Genotypen mittels spezieller Referenzsamples und wurde automatisiert in einem extra hierfür entwickelten bioinformatischen Workflow durchgeführt. Die Genotyp-Bestimmung aus den Intensität-Rohdaten wurde hierbei mit der Software GenomeStudio (Illumina, Inc., Version 2.0.2) durchgeführt, die QC mittels des PLINK-Programms (v1.9) (50, 51). Die Genotypimputation wurden schließlich mit der Software MiniMac3 (66) anhand des „Haplotype Reference Consortium“ (HRC) Referenzpanels (67) durchgeführt. Schlussendlich lagen somit Genotypwahrscheinlichkeiten von ca. 39 Millionen SNPs pro Probanden für die eigentlichen GWAS-Analysen vor. Alle Details der Rohdatenprozessierung, QC und Imputation des GDAC-Workflows sind bereits von der Arbeitsgruppe LIGA im Kontext anderer GWAS publiziert (68).

2.2.3.2 *Vorbereitende Analysen und Datenfilterungen*

Die prozessierten und imputierten genomweiten Datensätze die nachfolgend in meinen GWAS-Analysen verwendet wurden, wurden vom FOR2488-Teilprojekt INF-GDAC an mich weitergeleitet. Diese Datensätze beinhalteten nur Probanden und Kontrollpersonen europäischer Abstammung (engl. *non-Hispanic Whites*, [NHW]). Für die Datenfilterung und die anschließenden Assoziationsanalysen habe ich alle Dystonie-Teilnehmer nach dem Herkunftsland der Proben (d.h. Deutschland oder USA) in zwei unabhängige Datensätze aufgeteilt. Vor den eigentlichen GWAS-Analysen habe ich an diesen beiden Datensätzen noch eine detaillierte, weiterführende QC sowohl auf SNP- als auch auf Marker-Ebene entsprechend der folgenden Kriterien durchgeführt:

a) ausgeschlossen wurden Marker mit

- Imputations-Qualitäts-Score $r^2 < 0,7$
- Frequenz des seltenen Allels (MAF) unter 0,5 %
- Genotypisierungsrate unter 98 %
- Signifikanten Abweichungen vom HWE in Kontroll-Individuen ($p < 5 \times 10^{-6}$)

b) ausgeschlossen wurden DNA-Proben mit

- Genotypisierungseffizienz <98%
- Diskrepanzen zwischen genetischem und klinisch-erfasstem Geschlecht
- Nachweise auf Duplikation oder Monozygotie
- Verwandtschaftsverhältnissen zu anderen Probanden bis zum vierten Grad ($\hat{\pi} \geq 0,05$)
- Proben mit implausibler Heterozygotie (Mittelwert $\pm 6 \times SD$)

Ein wichtiger Aspekt bei allen genetischen Assoziationsanalysen ist das Vorliegen der sog. Populationsstratifikation. Hierbei handelt es sich um eine systematische Differenz in den Allelfrequenzen zwischen Subpopulationen eines Datensatzes, die durch subtile Unterschiede in der ethnischen Abstammung (z. B. Nordeuropäer vs. Südeuropäer) entsteht. Dieser Effekt ist besonders relevant bei Fall-Kontroll-Studien, wenn die Allelfrequenz-/Abstammungsunterschiede in Fällen anders sind als in Kontrollprobanden. Das Vorliegen unerkannter Populationsstratifikation kann im schlimmsten Fall zu Verzerrungen in den Assoziationsergebnissen führen. Zur Korrektur der Populationsstratifikation werden Ähnlichkeiten und Unterschiede in der genetischen Abstammung mittels Hauptkomponentenanalyse (engl. *principal component analysis*, [PCA]) berechnet und die Hauptkomponenten als Kontrollvariablen in die nachfolgenden Analysen eingeschlossen (69, 70). Des Weiteren werden vorher Ausreißer der ethnischen Abstammung auf der Grundlage der Falschfindungsrate (engl. *False Discovery Rate* [FDR], Grenzwert hier $< 0,05$) erkannt und entfernt. Die PCA wurde hierbei mit einer Teilmenge von 88.510 unabhängigen SNPs mittels der PLINK-Software (--pca) durchgeführt. Die unabhängigen SNPs wurden ebenfalls in PLINK mittels des sog. LD prunings (Grenzwert: $r^2 < 0,2$; Kommando: --indep-paarweise 1500 150 0,2) ermittelt.

2.2.3.3 Durchführung der GWAS und GWAS-Metaanalysen

Um auf die mögliche Assoziation zwischen den imputierten SNP-Genotypwahrscheinlichkeiten und dem Dystonie Fall-/Kontrollstatus zu testen, wurden logistische Regressionsmodelle unter Annahme eines additiven Modells mittels der PLINK-Software (v2) (51) berechnet. In den hier verwendeten Regressionsmodellen wurden die Variablen Geschlecht sowie die ersten zehn Hauptkomponenten der Abstammungs-PCA als Kontrollvariablen berücksichtigt. Formal lässt sich das Regressionsmodell wie folgt zusammenfassen:

$$\text{logit}(P(Y = 1|x)) = b_0 + b_1 PC_1 + \dots + b_{10} PC_{10} + b_{11} \text{sex} + b_{12} x_1 + \dots + b_{12+n} x_n$$

- x_n ist die unabhängige Variable, die die Anzahl an Allelen an einem SNP-Locus darstellt. Zum Beispiel kodiert ein SNP unter Annahme eines additiven Modells als 0, 1, 2 für a/a, a/A und A/A Genotypen.
- b_0 ist der Schnittpunkt auf der Ordinate (Y-Achse)
- b_n ist der Regressionskoeffizient und schätzt wie stark ein Risikofaktor mit der Dystonie zusammenhängt (kann als $\log(\text{OR})$ interpretiert werden).
- PC_1 - PC_{10} sind die ersten zehn Hauptkomponente der PCA.
- sex ist das biologische Geschlecht (Frau/Mann kodiert 0/1).
- Y ist die abhängige Variable, d.h. Fall- (1) oder Kontrollstatus (0) für jedes in die Analysen eingeschlossene Individuum
- $P(Y=1|x)$ ist die Eintrittswahrscheinlichkeit, dass ein Fallstatus einem gegebenen Genotyp zugeordnet wird.
- $\text{logit}()$ ist die logistische Funktion.

Nach der Schätzung der Stärke von einem Risikofaktor (b_n) wird diese Variable auf die statistische Signifikanz überprüft. Dabei werden zwei Hypothesen aufgestellt:

$$H_0: b_n = 0$$

$$H_1: b_n \neq 0$$

Liegt bei dem statistischen Testen der p-Wert unter dem Signifikanzniveau wird die Nullhypothese H_0 verworfen und die Alternativhypothese H_1 angenommen. Wie oben ausgeführt (s. Einleitung, Abschnitt 1.3.1.2), wurde hier – wie im GWAS-Feld üblich – eine Signifikanzgrenze von $\alpha = 5 \times 10^{-8}$ für die Identifizierung von genomweit signifikanten SNPs verwendet.

Alle GWAS-Analysen führte ich unter Verwendung des o.g. Modells gemäß dem Ursprung der DNA-Proben (Deutschland oder USA) getrennt durch. Die Mindestzahl der Studienteilnehmer für eine Assoziationsanalyse wurde hierfür auf 50 Patienten in einer Dystonie-Gruppen festgelegt. Neben der Analyse aller Fälle der isolierten Dystonie erfolgte eine weitere Unterteilung und GWAS-Analyse nach den folgenden klinischen Dystonie-Gruppen, wie sie vom Teilprojekt „Z2“ der ProtectMove-Forschungsgruppe (Projektleiter: Prof. M. Kasten, Prof. N. Brüggemann; Mitarbeit: Dr. E-J Vollstedt und Prof. Dr. C. Klein) an das P6-Teilprojekt zur Verwendung in den GWAS-Analysen weitergeleitet wurden:

- fokale Dystonie (unterteilt in zervikale Dystonie, laryngeale Dystonie [nur in der US-amerikanischen Stichprobe], kraniale Dystonie, Dystonie der oberen Extremitäten und aufgabenspezifische Dystonie [nur in der deutschen Stichprobe])
- segmentale und Multi-fokale Dystonie
- generalisierte Dystonie (nur in der US-amerikanischen Stichprobe)

Die Ergebnisse der Assoziationsanalysen der gleichen Dystonie-Gruppen bei den DNA-Datensätzen der deutschen und der amerikanischen Ursprungs habe ich basierend auf dem Modell der inversen Varianzgewichtung mit der Hilfe des METAL-Software (71) metaanalysiert. Für jede GWAS-Analyse sowie jede GWAS-Metaanalyse habe ich außerdem den genomischen Inflationsfaktor Λ geschätzt, um den Grad der Abweichung der erwarteten Teststatistikverteilung abzuschätzen. Je stärker der Inflationsfaktor von 1 abweicht, desto grösser ist die Inflation (>1) oder Deflation (<1) der Teststatistikverteilung. Außerdem wurde eine visuelle Inspektion aller GWAS-Ergebnisse auf Vorliegen einer systematischen Verzerrung der Teststatistikverteilung anhand des Quantil-Quantil-Plot (engl. *quantile-quantile-plot*, QQ-Plots) durchgeführt.

2.2.4 Analyse des X-Chromosoms

Für die X-chromosomalen Assoziationsanalysen wurden nur die Probanden eingeschlossen, die nach QC in den GWAS-Analysen der Autosomen verblieben (siehe Abschnitt 2.2.3.2). Das X-Chromosom unterscheidet sich von den Autosomen hinsichtlich der seiner Verteilung bei Männern vs. Frauen. Männliche Probanden haben nur eine Kopie des X-Chromosoms und damit auch nur ein mögliches Allel pro SNP (a, A), während Frauen zwei Kopien des X-Chromosoms und damit auch zwei mögliche Allele besitzen, wie bei den Autosomen (a/a, a/A und A/A). Dieser Unterschied muss in X-Chromosom-weiten Assoziationsanalysen (XWAS) durch ein angepasstes Analysemodell berücksichtigt werden (72). Hierbei wurden zuerst die Ergebnisse der XWAS-Analysen der männlichen und weiblichen Probanden aus Deutschland und den USA getrennt metaanalysiert. Im letzten Schritt fasste ich die Ergebnisse der weiblichen und männlichen Metaanalysen nach der Methode nach Stouffer zusammen (73). Insgesamt basierte das hier verwendete XWAS-Analyse-Protokoll weitgehend den Empfehlungen in König et al., 2014 (74) und Smith et al., 2021 (72) .

2.2.4.1 Prozessierung und Imputation der Rohgenotypdaten auf dem X-Chromosom

Die Bestimmungen der Genotypen auf dem X-Chromosom wurden von der GDAC-Pipeline, wie in Abschnitt 2.2.3.1. beschrieben, durchgeführt. Die QC der X-chromosomalen Daten und die Imputation der fehlenden Genotypen wurde von mir selbst erstellt. Zu Beginn der QC wurden die pseudoautosomalen Regionen (PAR1, PAR2) auf dem X-Chromosom ausgeschlossen. Die MAF der X-chromosomalen SNPs wurde in der größten verfügbaren weiblichen Kontrollgruppe, d.h. anhand

des PopGen-Datensatzes, berechnet und als MAF-Referenz für allen anderen Datensätze verwendet.

Ausgeschlossen wurden Marker mit

- MAF unter 1% in der weiblichen PopGen-Kontrollgruppe
- Genotypisierungsrate unter 98%, getrennt bei Frauen und Männern
- Signifikanten Abweichungen vom HWE in Kontroll-Individuen ($p < 1,0 \times 10^{-4}$)
- Differentieller Genotypisierungseffizienz zwischen Frauen und Männern ($p < 1,0 \times 10^{-4}$)
- Differentieller Allelfrequenz zwischen Frauen und Männern ($p < 1,0 \times 10^{-4}$)
- Mehrdeutigen Allel-Kombinationen (A/T und C/G)

Des Weiteren wurden SNPs ausgeschlossen die nur im „shared custom content“ des GSA vorkamen und SNPs mit inserierten oder deletierten Allelen (sog. „indel-SNPs“). Um die Imputationsgenauigkeit zu erhöhen habe ich die Allele der verbleibenden SNPs vor der Imputation am Referenzgenom „GRCh37/hg19“ mit der „+fixref“-Software (75, 76) gleichgerichtet. Für die Ausrichtung der Allele verwendete ich die vom Sanger Imputation Server empfohlenen „FASTER“-Dateien (<https://imputation.sanger.ac.uk/?resources=1>). Die Imputation weiblicher und männlicher Teilnehmer wurden nachfolgend auf dem Imputations-Server des Sanger-Instituts (<https://imputation.sanger.ac.uk/>) unter Verwendung eines erweiterten HRC-Referenzpanels (67) durchgeführt. Dieses Panel umfasst insgesamt 1.268.718 SNPs auf dem X-Chromosom von 31.500 Personen (männlich und weiblich).

2.2.4.2 Vorbereitende Analysen und Datenfilterungen für das X-Chromosom

Nach der Imputation führte ich, analog zu den Analysen der Autosomen, eine erneute QC der SNPs durch, getrennt an weiblichen und männlichen Datensätzen.

Ausgeschlossen wurden Marker mit

- Imputations-Qualitäts-Score $r^2 < 0,7$
- Frequenz des seltenen Allels (MAF) unter 1%
- SNPs mit einer Genotypisierungsrate unter 98%
- Signifikanten Abweichungen vom HWE in Kontroll-Individuen ($p < 1,0 \times 10^{-4}$)

Auf Sample-Ebene wurden alle Proben mit einer Genotypisierungsrate unter 98% ausgeschlossen.

2.2.4.3 Durchführung von Assoziationsanalysen und Metaanalysen am X-Chromosom

Männliche Probanden haben nur eine Kopie des X-Chromosoms und damit auch nur ein mögliches Allel pro SNP. Die Kodierung der Allele bei männlichen Probanden ist daher 0 oder 1. Bei weiblichen Probanden sind, wie bei den Autosomen, entweder 0, 1 oder 2 Kopien eines Allels vorhanden. Tatsächlich liegt bei Frauen wegen der Inaktivierung eines der beiden X-Chromosomen aber nur ein funktionelles Allel vor. Damit hat das einzelne Risikoallel des Mannes dieselbe Wirkung wie zwei Risikoallele bei der Frau. Um diesem Umstand Rechnung zu tragen, wird daher bei Männern häufig eine 0 vs. 2 anstatt einer 0 vs. 1 Kodierung für X-chromosomale Marker verwendet, um so die Effektgewichtung der Männer an die der Frauen anzupassen (77). Folgende Schritte wurden bei den XWAS-Analysen dieser Arbeit verfolgt:

Schritt 1. Zuerst wurden die XWAS-Analysen getrennt bei weiblichen und männlichen Studienteilnehmern für alle in Abschnitt 2.2.3.3. aufgeführten klinischen Dystonie-Gruppen berechnet. Im Rahmen der Analyse wurde dieselben logistischen Regressionsmodelle wie in Kapitel 2.2.3.3 beschrieben angewendet.

Schritt 2. Im zweiten Schritt wurden getrennt die weiblichen und männlichen Studienteilnehmer aus Deutschland und den USA mit Hilfe der METAL-Software mit dem Modell der inversen Varianzgewichtung metaanalysiert.

Schritt 3. Abschließend wurden die p-Werte aus der Metaanalyse der Frauen und die p-Werte aus der Metaanalyse der Männer mit der Methode nach Stouffer zusammengefasst.

Für jede XWAS sowie für jede Metaanalyse wurde der genomische Inflationsfaktor berechnet und die Testergebnisse visuell anhand des QQ-Plots auf Abweichung der erwarteten Teststatistikverteilung inspiziert.

2.2.5 Post-GWAS-Analysen

Wie am Ende der Einleitung bereits ausgeführt, erlauben die eigentlichen GWAS/XWAS-Analysen nur Aussagen in Bezug auf die statistische Assoziation zwischen genetischen Varianten und den abhängigen Phänotypen, aber keine Aussagen zu möglichen kausalen Zusammenhängen. Die weitere Interpretation der GWAS-Befunde, z. B. zur Eruierung möglicher molekularer Mechanismen und funktioneller Ursachen, ist die Aufgabe sog. „Post-GWAS-Analysen“. Hierbei werden z. B. unter Zuhilfenahme einer großen Anzahl weiterer genomweiter Datensätze und Analyseverfahren weitere Vorhersagen und Implikationen von den vorliegenden GWAS-Ergebnissen gemacht (z. B. der Frage nachgegangen, ob ein Dystonie-assoziiertes SNP auch mit der Expression eines Gens assoziiert ist). Ein weiterer Schritt ist die Durchführung von Gen-basierten Assoziationsanalysen, die

anhand der SNP-basierten GWAS-Ergebnisse berechnet werden können. In diesen Analysen werden die Ergebnisse mehrerer SNPs, die alle im selben Genlocus liegen, zu einer Gen-basierten Teststatistik zusammengefasst. Dies erleichtert z. B. die Vergleichbarkeit von GWAS-Ergebnissen an unterschiedlichen Phänotypen. Diese und andere im Rahmen dieser Arbeit durchgeführten Post-GWAS-Analysen sind nachfolgend näher beschrieben.

2.2.5.1 Bestimmung, funktionelle Zuordnung und Annotation genetischer Loci

Für diesen Teil der Post-GWAS-Analysen habe ich die integrative, web-basierte Plattform FUMA (engl. *Functional Mapping and Annotation*, Version 1.3.6b) (78) verwendet. Als Genlocus wurde ein Bereich definiert, in dem sich die unabhängigen signifikanten SNPs ($P \leq 5 \times 10^{-6}$, $r^2 < 0.6$) basierend auf der physischen Lage auf einem Chromosom innerhalb eines Blocks (< 250 Kilo Basenpaare, kb) überlappten. Es wurden weitere SNPs zur Annotation einbezogen, falls diese in einem LD ($r^2 > 0.6$) mit einem unabhängigen signifikanten SNP stehen. Die LD-Werte der SNPs wurden mit FUMA auf der Grundlage des „1000 Genomes Project“ (53) bei der europäischen Bevölkerung vorberechnet und für die Zuordnung des Genlocus verwendet. Innerhalb von Genloci bestimmt das FUMA-Tool die führenden-SNPs (engl. *lead-SNPs*, nachfolgend hier auch so bezeichnet), die ein $r^2 < 0.1$ mit anderen SNPs innerhalb eines Genlocus aufweisen.

Mittels FUMA wurden die SNPs zu den Genen basierend auf der ANOVAR-Software (79) basierend auf der physische Position auf den Chromosomen annotiert. Die Information über die physische Position von Genen stammten aus den „Ensembl Genome Browser“ Archiv 92 (<https://www.ensembl.org/index.html>). Die CADD-Werte und die RegulomeDB-Werte wurden zu den SNPs hinzugefügt, basierend auf der Position auf dem Chromosom und der Allel-Übereinstimmung. Der CADD-Wert basiert auf einer aufwendigen Vorhersage zur möglichen Pathogenität einer DNA-Variante basierend auf experimentellen und theoretischen Daten. Liegt der CADD-Wert über 12,37, so wird ein SNP als potentiell pathogen angesehen (80). RegulomeDB-Werte geben das regulatorische Potential eines SNPs wieder mit dem Ziel der Vorhersage, ob und auf welcher Ebene eine DNA-Variante an der Genexpressionsregulation beteiligt ist. Ein SNP kann mit Werten von 1 bis 7 eingestuft werden. Je niedriger der Wert ist, desto wahrscheinlicher ist eine regulatorische Funktion eines SNPs (81).

2.2.5.2 Gen-basierte GWAS-Analysen

Die Gen-basierten GWAS-Analysen wurde in FUMA unter der Verwendung des MAGMA-Tools (Version 1.08) (82) durchgeführt. Für die Berechnung der Gen-basierten GWAS habe ich nur die Ergebnisse (d.h. p-Werte) aus der autosomalen GWAS verwendet. Die SNPs werden zu einem Gen, basierend auf ihrer physischen Position annotiert, wie in Kapitel 2.2.5.1 beschrieben. Innerhalb

eines Gens kombiniert die MAGMA-Software die p-Werte der eingeschlossenen SNPs aus der GWAS zu einer Gentest-Statistik mittels eines SNP-weiten Mittelwertmodells (82).

Um die studienweite Signifikanzschwelle für die Gen-basierten GWAS-Analysen zu berechnen, habe ich, wie auf der FUMA-Website empfohlen, die Anzahl der für diese Analysen verwendeten Gene (n=19.516) in FUMA berücksichtigt. Dementsprechend ergibt sich eine Anpassung der studienweiten Signifikanzschwelle nach der Bonferroni Methode mit $\alpha = 0,05/19516 = 2,6 \times 10^{-06}$.

2.2.5.3 Gen-Property Analysen

Gen-Property Analysen wurden mittels des FUMA-Tools unter Verwendung der MAGMA-Software (Version 1.08) (82) berechnet. Die Gen-Property-Analyse testet eine Beziehung zwischen zelltypspezifischer Genexpression und Krankheitsgen-Assoziation mit der Hilfe eines Regressionsmodells (83). In dieser Arbeit werden die Gen-Property-Analysen unter Verwendung der Gen-basierten GWAS p-Werte und der Genexpressionswerte der 54 allgemeinen Gewebetypen des Genotype-Tissue Expression (GTEx)-Projekts (v.8) (84) in FUMA berechnet.

2.2.5.4 Heritabilitätsbestimmungen

Die Heritabilität einer Erkrankung erfasst den Anteil der Variation eines Merkmals, der durch vererbte genetische Varianten erklärt wird (H^2) (85). In dieser Arbeit wurde dieser Ansatz zur Berechnung der SNP-basierten Heritabilität (h^2) angewendet. Die SNP-basierte Heritabilität ist der Anteil der Variation eines Merkmals (bzw., wie in dieser Arbeit: einer Krankheit), der durch die additiven Effekte von SNPs erklärt wird (86). Die SNP-basierte Heritabilität wurde für die verschiedenen Dystonien-Gruppen mit dem „Genome-wide Complex Trait Analysis“ (GCTA) Tool (Version 1.92.1beta5) (86) auf der Grundlage von imputierten Genotypen ($r^2 > 0,7$, MAF $> 0,1$) geschätzt. In der Analyse berücksichtigte ich als Kovariablen das biologische Geschlecht (engl. sex), die ersten 10 PCs und die bevölkerungsbezogene Prävalenz der Dystonie. Die Prävalenzdaten wurden dabei aus den Studien von Castelon-Konkiewitz et al., 2002 (6) für die deutsche Bevölkerung und aus Butler et al., 2004 und Nutt et al., 1988 (7, 87) für die US-Bevölkerung entnommen.

In der Literatur fehlen derzeit Prävalenzschätzungen für die kraniale Dystonie in der deutschen Bevölkerung. In der amerikanischen Bevölkerung fehlen die Angaben zur multifokalen und segmentalen Dystonie. Die fehlenden Prävalenzen schätzte ich unter Einbeziehung der Studie von Ortiz et al, 2018 (88). In dieser Publikation werden die gesuchten Prävalenzen für die finnische Bevölkerung angegeben. Bei der Berechnung wird zuerst der Anteil der diagnostischen Untergruppe bezogen auf die isolierte Dystonie aus Ortiz et al, 2018 (88) bestimmt. Dieser Anteil

wurde mit der Prävalenz der isolierten Dystonie bei Deutschen und bei Amerikanern multipliziert. Die verwendete Prävalenz wird in der Tabelle 3.9 (in Ergebnissteil) dargestellt.

2.2.5.5 Korrelationsanalysen mit anderen komplex-genetischen Phänotypen

Die genetische Korrelation zwischen den drei größten im Rahmen dieser Arbeit durchgeführten Dystonie-Metaanalysen und ausgewählten nicht-Dystonie Phänotypen (Erkrankungen und Umweltfaktoren) untersuchte ich mit Hilfe des „LD-Score-Regression“ (LDSC) Tools (89). Die drei größten Dystonie-Metaanalysen waren die GWAS zur isolierten Dystonie, der fokalen Dystonie und der zervikalen Dystonie. Als nicht-Dystonie Phänotypen habe ich die 12 neuesten und größten GWAS zu anderen neurologischen Erkrankungen, psychiatrischen Erkrankungen, der kognitiven Leistungsfähigkeit, sowie zu lebensstil-bezogenen Daten und dem Körpermasseindex (BMI) ausgewählt (siehe Tabelle 3.12 in Ergebnissteil). Für die Berechnung der genetischen Korrelation werden die LD-Scores aus der europäischen Population anhand der Genotypdaten des „1000 Genomes Project“ (53) verwendet. Die QC der SNPs aus den hier durchgeführten Dystonie GWAS-Analysen wurde basierend auf den Empfehlungen der Publikation von Bulik-Sullivan et al., 2015 (89) durchgeführt.

Die genetischen Korrelationswerte werden mehrfach (d.h. 36-fach) auf statistische Signifikanz getestet und das Signifikanzniveau über alle Analysen hinweg nach der Bonferroni Methode angepasst. Somit ergibt sich für diese Analysen eine statistische Signifikanzgrenze α von $0,05 / 36 = 1,39 \times 10^{-03}$.

2.3 Vergleich der neuen GWAS-Ergebnisse mit denen der Field-Synopsis bereits publizierter Daten

Aus meiner Dystonie Field-Synopsis (25) wurden alle Assoziationsergebnisse mit nominaler Signifikanz ausgewählt, d.h. die mit SNPs rs4532, rs13283584, rs35153737, rs7342975, rs9972951, rs1801968 und rs11655081. Diese vorherigen Assoziationsbefunde verglich ich dann mit der Allelhäufigkeit, der Effektstärke und dem p-Wert der Ergebnisse der einzelnen GWAS und GWAS-Metaanalysen aus dem zweiten Teil dieser Arbeit. Des Weiteren habe ich die Hauptbefunde der in der Einleitung erwähnten (s. Abschnitt 1.3.2), bereits publizierten Dystonie-GWAS mit den Ergebnissen meiner GWAS-Analysen verglichen.

3 Ergebnisse

In diesem Kapitel werden die Ergebnisse der aus Kapitel 2 umgesetzten Methodik vorgestellt. Im Abschnitt 3.1 werden die Ergebnisse der Field-Synopsis zum Stand der Literatur bezüglich genetischer Assoziationsstudien der Dystonie zusammengefasst. Die Resultate der GWAS zu der isolierten Dystonie werden im Abschnitt 3.2 dargestellt. Außerdem wird auf die SNP-basierte Heritabilität und die Korrelationsanalysen von der Dystonie mit anderen komplex-genetischen Phänotypen eingegangen. Im Abschnitt 3.3 wird ein Vergleich der Ergebnisse zwischen den GWAS-Analysen aus 3.2. und der Field-Synopsis aus 3.1. dargestellt.

3.1 Field-Synopsis zum Stand genetischer Assoziationsstudien der Dystonie

Die Ergebnisse dieses Teils der Arbeit wurden bereits 2018 in dem Journal „*Parkinsonism and Related Disorders*“ veröffentlicht (25). Abbildung 3.1, die aus dieser Publikation übernommen

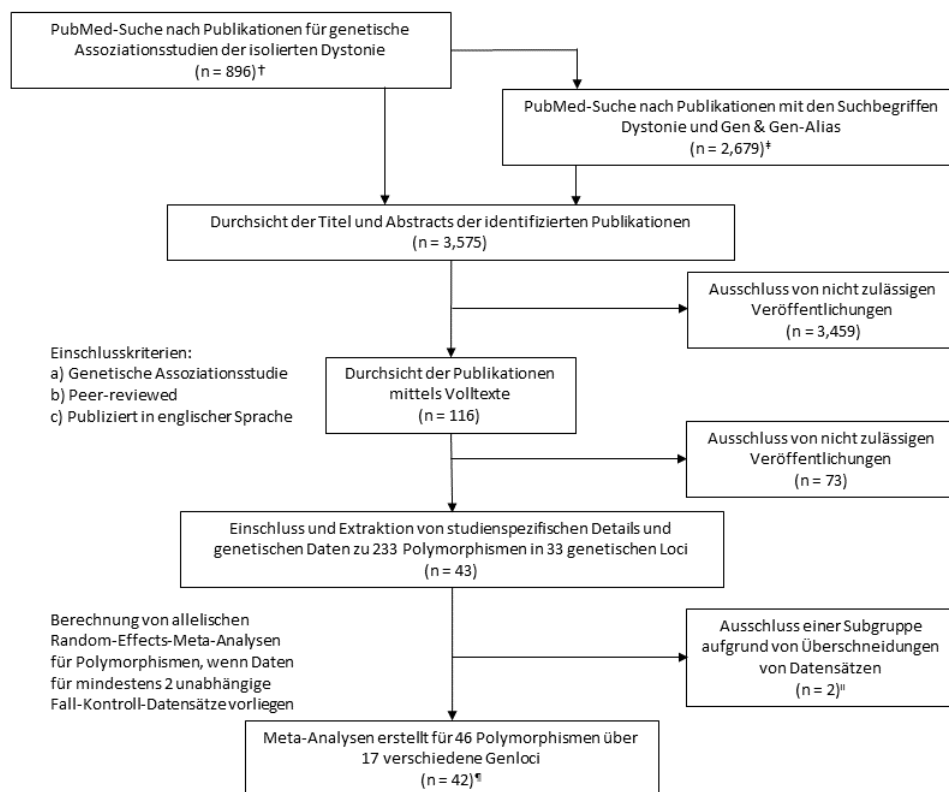


Abbildung 3.1 Flussdiagramm der Literatursuche, Datenextraktion und Analysestrategien, die für die Field-Synopsis angewendet wurden. † Suche nach den Suchbegriffen "(dystoni* OR (writer*AND cramp) OR graphospasm OR blepharospasm OR torticollis OR meige OR spasmodic OR dysphonia OR retrocollis OR antecollis OR laterocollis OR (musician*AND cramp) OR (occupation* AND cramp) OR (golf* AND cramp) OR yips) AND associat* AND gene*". ‡ Suche nach den oben genannten Suchbegriffen und "AND ([GenName] OR [Gen-Alias(e)])" unter Verwendung der HGNC (HUGO Gene Nomenclature Committee at the European Bioinformatics Institute) Nomenklatur (<https://www.genenames.org/>). ¶ Nur 45 von 233 Polymorphismen wurden in mindestens zwei unabhängigen Datensätzen untersucht. Reproduziert von Ohlei et al., 2018 (25).

wurde, stellt eine Übersicht über die Vorgehensweise bei der Literaturrecherche im Rahmen der Field-Synopsis dar.

3.1.1 Literatursuche, Ein- und Ausschlusskriterien

Die systematische Literatursuche wurde in zwei Schritten durchgeführt. Zuerst wurden 896 Artikel bei der Allgemeinen suche nach „([Dystonie Typ] AND associat* AND gene*)“ identifiziert und nach den Einschluss- und Ausschluss-Kriterien, welche im Methodenteil 2.1.1 vorgestellt wurden, untersucht. Die aus dieser Suche hervorgegangenen Artikel wurden zunächst in einer Datenbank (via Microsoft Excel; Ergänzungstabelle 1 auf der beigefügten *CD-ROM*) erfasst. Anschließend wurden die in den einzelnen Publikationen erwähnten Gen-Namen herausgesucht und in Kombination mit den bekannten 28 DYT Bezeichnungen in einer weiteren Suche auf PubMed recherchiert (siehe Abbildung. 3.1 und Kapitel 2.1.1). Die PubMed Datenbankabfragen umfassen die Ergebnisse bis zum 23. Januar 2018. Insgesamt wurden 3.575 potenziell in Frage kommende Artikel aus den Suchanfragen identifiziert. Bei jedem dieser Artikel wurde daraufhin im Titel und dem Abstract geprüft, ob es sich um eine genetische Assoziationsstudie handelt. Bei den Artikeln, bei denen dies im Titel und Abstract nicht erkenntlich war, wurden Volltextversionen zur Prüfung heruntergeladen. Bei diesem Vorgehen wurden insgesamt 116 einzelne Publikationen identifiziert und nach den im Methodenteil (2.1.1.) vorgestellten Einschluss- und Ausschluss Kriterien beurteilt, die von 42 Publikationen erfüllt wurden (25).

3.1.2 Eingabe der Daten und Qualitätskontrolle

Die 42 die in Field-Synopsis eingeschlossenen Publikationen beschreiben Assoziationsdaten von 52 unabhängigen Datensätzen. Die analysierten Datensätze stammen aus 17 verschiedenen Ländern, die über vier Kontinente verteilt sind. Insgesamt wurden in diesen 42 Publikationen die Daten zu 233 Polymorphismen über 33 Genloci ermittelt; diese Daten wurde in die Datenbank für weitergehende Analysen eingegeben und von einer zweiten Person der Arbeitsgruppe (Dr. Dobricic) auf Korrektheit überprüft. Die meisten Publikationen verwendeten den Ansatz der Kandidatengenanalysen (n = 40). Die verbleibenden zwei Publikationen waren GWAS. Die gesamte Datenbasis der Field-Synopsis ist in der Ergänzungstabelle 1 auf der beigefügten *CD-ROM* enthalten.

3.1.3 Metaanalyse eingeschlossener Daten

Für die Metaanalyse wurden nur die DNA-Varianten berücksichtigt, welche in mindestens zwei unabhängigen Datensätzen vorkamen, in denen ebenfalls die Allelhäufigkeiten der Polymorphismen bekannt waren. Eine der zwei ermittelten GWAS-Publikationen zur Musikerdystonie (39), hat die genomweiten Teststatistiken für 557.621 SNPs verfügbar gemacht.

Diese Daten wurden auf eine Überlappung zu den 233 in der Literatur identifizierten Polymorphismen geprüft und ggf. in die Metaanalysen eingeschlossen. Dieses Verfahren ergab insgesamt 45 DNA-Varianten über 17 genetische Loci (Median der Stichprobengröße pro Datensatz: 458, Spannweite: 88-5.385) bei denen die GWAS-Ergebnisse aus Lohman et al., 2014 (39) eingeschlossen wurden. Nach Stratifizierung in Bezug auf die ethnische Abstammung ergaben sich Metaanalysen für 42 Varianten in Datensätzen europäischer Abstammung (in den Primärpublikationen wurden diese häufig als „kaukasisch“ bezeichnet) und 6 DNA-Varianten in Datensätzen asiatischer Abstammung (25).

Tabelle 3.1 Ergebnisse von Random-Effekt Metaanalysen auf der Basis eines additiven Modells bei allen Arten von isolierter (auch bekannt als "primärer") Dystonie. Reproduziert von Ohlei et al., 2018 (25).

Chr-Position	Gene	SNP	A1/A2	AF*	E. †	Fälle vs. Kontrollen (#Datensätze /# Publikationen)	OR (95%KI)	P	I ²
1:236885200	<i>MRT</i>	rs1805087	G/A	0,19	C	227 vs 1072 (2/2)	0,93 (0,66-1,29)	0,5852	0,00
2:178433097	<i>PRKRA</i>	rs6739142	G/A	0,21	C	357 vs 1200 (2/2)	0,99 (0,73-1,36)	0,9910	39,77
2:178449632	<i>PRKRA</i>	rs1967327	C/G	0,18	C	357 vs 1200 (2/2)	1,10 (0,86-1,42)	0,4392	0,00
2: 218326846	<i>PNKD</i>	rs10932774	A/G	0,37	C	357 vs 1200 (2/2)	1,02 (0,84-1,24)	0,8736	0,00
3:114171968	<i>DRD3</i>	rs6280	C/T	0,32	C	188 vs 200 (2/2)	0,82 (0,61-1,11)	0,1979	0,00
5:175441837	<i>DRD1</i>	rs155417	T/C	0,02	C	188 vs 200 (2/2)	0,73 (0,31-1,74)	0,4792	0,00
5:175443147	<i>DRD1</i>	rs4532	C/T	0,37	C	315 vs 1172 (3/3)	1,37 (1,13-1,67)	0,0015	0,00
7:94635805	<i>SGCE</i>	rs10235385	C/A	0,17	C	357 vs 1200 (2/2)	0,90 (0,69-1,16)	0,4028	0,00
8:42838423	<i>THAP1</i>	rs111989331	C/T	0,11	All	499 vs 360 (2/2)	0,69 (0,31-1,54)	0,3686	74,48
8:42842898	<i>THAP1</i>	rs71521601	G/A	0,10	C	864 vs 510 (3/3)	1,20 (0,66-2,21)	0,5484	73,02
8:42843330-1	<i>THAP1</i>	rs370983900 & rs184497763	AA/ CT	0,03	C	3679 vs 2315 (7/7)	1,07 (0,73-1,56)	0,7431	40,61
9:129812583	<i>TOR1A</i>	rs13283584	T/C	0,23	C	473 vs 749 (2/2)	1,32 (1,08-1,61)	0,0058	0,00
9:129813148	<i>TOR1A</i>	rs3842225	- /C	0,16	All	2287 vs 2503 (11/10)	1,05 (0,91-1,21)	0,4862	43,18
9:129813148	<i>TOR1A</i>	rs3842225	-/C	0,26	C	1835 vs 2083 (9/8)	1,01 (0,86-1,18)	0,9450	44,67
9:129813148	<i>TOR1A</i>	rs3842225	-/C	0,19	A	452 vs 420 (2/2)	1,28 (0,99-0,65)	0,0505	0,00
9:129813558	<i>TOR1A</i>	rs35153737	- /C	0,13	A	452 vs 420 (2/2)	1,46 (1,14-1,88)	0,0031	0,00
9:129813781	<i>TOR1A</i>	rs1182	A/C	0,17	All	2213 vs 2526 (11/10)	1,05 (0,85-1,30)	0,6742	72,36
9:129813781	<i>TOR1A</i>	rs1182	A/C	0,22	C	1572 vs 2010 (8/7)	1,01 (0,77-1,32)	0,9607	79,00
9:129813781	<i>TOR1A</i>	rs1182	A/C	0,16	A	641 vs 516 (3/3)	1,15 (0,90-1,45)	0,2647	0,00
9:129816005	<i>TOR1A</i>	rs11787741	G/A	0,22	C	473 vs 749 (2/2)	0,95 (0,77-1,17)	0,6260	0,00
9:129818407	<i>TOR1A</i>	rs13297609	C/G	0,15	All	351 vs 428 (2/2)	1,16 (0,89-1,52)	0,2742	0,00
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,12	All	2075 vs 2699 (11/11)	0,95 (0,78-1,16)	0,6332	43,00
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,14	C	1428 vs 1925 (6/6)	1,02 (0,87-1,20)	0,7870	5,33
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,13	A	626 vs 494 (4/4)	0,74 (0,44-1,26)	0,2641	62,34
9:129822779	<i>TOR1A</i>	rs2296793	A/G	0,22	All	1692 vs 2277 (10/9)	1,02 (0,87-1,18)	0,8381	42,61
9:129822779	<i>TOR1A</i>	rs2296793	A/G	0,24	C	1454 vs 1946 (8/7)	1,02 (0,86-1,21)	0,8221	45,59
9:129822779	<i>TOR1A</i>	rs2296793	A/G	0,20	A	238 vs 331 (2/2)	0,99 (0,60-1,64)	0,9738	64,39
9:133636819	<i>DBH</i>	rs2797849	C/G	0,34	C	490 vs 3054 (2/2)	1,09 (0,89-1,32)	0,4079	38,64
9:133639458	<i>DBH</i>	rs2797851	A/G	0,33	C	490 vs 3054 (2/2)	1,07 (0,87-1,34)	0,5395	42,38
9:133657065	<i>DBH</i>	rs1611131	G/A	0,29	C	490 vs 3054 (2/2)	1,12 (0,83-1,51)	0,4528	69,85
9:133657152	<i>DBH</i>	rs6271	T/C	0,08	C	490 vs 3054 (2/2)	2,97 (0,37-24,02)	0,3002	96,73
9:133663599	<i>DBH</i>	rs129886	T/C	0,19	C	490 vs 3054 (2/2)	1,01 (0,86-1,20)	0,8704	0,00
10:28431147	Interg.	rs1249277	G/C	0,15	C	464 vs 5515 (2/2)	0,77 (0,41-1,45)	0,4215	89,34
11:27658369	<i>BDNF</i>	rs6265	T/C	0,19	All	1591 vs 2780 (7/7)	1,09 (0,98-1,21)	0,1165	0,00
11:27658369	<i>BDNF</i>	rs6265	T/C	0,19	C	1146 vs 2350 (5/5)	1,05 (0,91-1,21)	0,4936	7,52
11:27658369	<i>BDNF</i>	rs6265	T/C	0,49	A	445 vs 430 (2/2)	1,19 (0,98-1,43)	0,0759	0,00
11:48246304	<i>OR4X2</i>	rs67863238	C/G	0,06	C	464 vs 5515 (2/2)	2,17 (0,40-11,70)	0,3482	96,50
11:113400106	<i>DRD2</i>	rs1800497	A/G	0,19	C	222 vs 267 (3/3)	0,89 (0,65-1,22)	0,4758	0,00
11:113425564	<i>DRD2</i>	rs1079597	T/C	0,17	C	188 vs 200 (2/2)	1,00 (0,69-1,46)	0,9913	0,00
11:113475529	<i>DRD2</i>	rs1799732	- /G	0,10	C	188 vs 200 (2/2)	0,74 (0,46-1,19)	0,2064	0,00
13:101406511	<i>NALCN</i>	rs1338041	C/A	0,41	All	565 vs 5804 (3/3)	0,90 (0,59-1,36)	0,6014	89,79
13:101406511	<i>NALCN</i>	rs1338041	C/A	0,36	C	464 vs 5515 (2/2)	0,75 (0,51-1,10)	0,1445	83,78
13:101407929	<i>NALCN</i>	rs9518385	C/A	0,36	C	339 vs 6145 (2/2)	0,73 (0,51-1,05)	0,0901	76,89

Tabelle 3.1 (Fortsetzung)

Chr-Position	Gene	SNP	A1/A2	AF*	E. ‡	Fälle vs. Kontrollen (#Datensätze /# Publikationen)	OR (95%KI)	P	i ²
13:101430922	<i>NALCN</i>	rs61973742	G/A	0,10	All	665 vs 5804 (3/3)	1,38 (0,41-4,64)	0,6001	96,35
13:101430922	<i>NALCN</i>	rs61973742	G/A	0,07	C	464 vs 5515 (2/2)	1,85 (0,26-13,37)	0,5245	97,41
14:54837362	<i>GCH1</i>	rs2145945	A/G	0,36	C	357 vs 1200 (2/2)	0,92 (0,66-1,27)	0,6038	64,12
14:54839739	<i>GCH1</i>	rs10483639	C/G	0,19	C	360 vs 2062 (2/2)	1,09 (0,72-1,64)	0,6990	76,69
14:54841286	<i>GCH1</i>	rs17739146	G/A	0,16	C	357 vs 1200 (2/2)	0,91 (0,66-1,27)	0,5921	22,85
14:54888313	<i>GCH1</i>	rs998259	T/C	0,22	C	357 vs 1200 (2/2)	0,97 (0,74-1,27)	0,7973	36,39
17:68368608	<i>ARSG</i>	rs1558877	T/C	0,48	C	357 vs 5272 (2/2)	0,96 (0,78-1,18)	0,6883	39,26
17:68368663	<i>ARSG</i>	rs1558878	C/T	0,49	C	357 vs 5272 (2/2)	0,88 (0,75-1,02)	0,0846	0,00
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,18	All	2420 vs 3608 (9/1)	1,33 (0,95-1,86)	0,0989	77,20
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,05	C	2243 vs 3409 (8/1)	1,38 (0,92-2,08)	0,1224	79,97
17:68395091	<i>ARSG</i>	rs7342975	G/A	0,01	C	357 vs 5272 (2/2)	2,00 (1,15-3,47)	0,0057	0,00
17:68395135	<i>ARSG</i>	rs9972951	A/G	0,03	C	357 vs 5272 (2/2)	2,16 (1,06-4,40)	0,0342	68,35
22:19950115	<i>COMT</i>	rs5993883	T/G	0,49	C	490 vs 3056 (2/2)	0,90 (0,61-1,31)	0,5671	83,57
22:19964609	<i>COMT</i>	rs4646316	T/C	0,23	C	490 vs 3056 (2/2)	1,08 (0,92-1,27)	0,3332	0,00
22:19968169	<i>COMT</i>	rs9332377	T/C	0,16	C	490 vs 3056 (2/2)	1,11 (0,92-1,34)	0,2678	0,00

Interg. = intergenische (d.h. zwischen annotierten Genen liegende) Region

Fettgedruckte Schrift kennzeichnet signifikante Ergebnisse vor Korrektur auf multiples Testen

‡ E. (Ethnizität). C = Europäische („kaukasische“) Abstammung (d. h. mehr als 90 % der Individuen in einem Datensatz gehörten dieser Abstammungsgruppe an); A = Asiatische Abstammung; All = Alle ethnischen Gruppen zusammen.

* Effekt-AF (Allelfrequenz) für jede Population, wie in der gnomAD-Datenbank (<http://gnomad.broadinstitute.org/>) aufgeführt. A – Asiatische Populationen beziehen sich auf gnomAD AF als "Ostasiatisch", wie in den Originalpublikationen berichtet. Für zwei Varianten, rs35153737 und rs3842225, waren die AF in gnomAD nicht verfügbar, so dass die AF stattdessen von dbSNP (<https://www.ncbi.nlm.nih.gov/SNP/>) übernommen wurde.

Fünf Varianten in *DRD1*, *TOR1A* und *ARSG* zeigten eine nominal signifikante Assoziation, d.h. vor Korrektur für multiples Testen: rs4532 in *DRD1* (OR: 1,37; 95% KI: 1,13-1,67; p = 0,00153), rs35153737 in *TOR1A* (OR: 1,46; 95% KI: 1,14-1,88; p = 0,00315), rs13283584 in *TOR1A* (OR: 1,32; 95% KI: 1,08-1,61; p = 0,00584), rs7342975 in *ARSG* (OR: 2,00; 95% KI: 1,15-3,47; p = 0,00572), rs9972951 in *ARSG* (OR: 2,16; 95% KI: 1,06-4,40; p = 0,0342; Tabelle 3.1, reproduziert von Ohlei et al., 2018 (25)). Nur die Variante rs4532 in *DRD1* zeigte auch nach Korrektur für multiples Testen noch einen signifikanten p-Wert. Weitere Metaanalysen mit ethnischer Abstammung stratifizierten Untergruppen ergaben keine zusätzlichen signifikanten Assoziationen (siehe Tabelle 3.1).

In weiteren Analysen wurde die Dystonie nach diagnostischen Untergruppen unterteilt. Die Ergebnisse der dazugehörigen Metaanalysen sind in der Tabelle 3.2 zusammengefasst, reproduziert von Ohlei et al., 2018 (25). Die Analyse nach diagnostischen Untergruppen ergab fünf signifikante SNPs vor Korrektur für multiples Testen, von denen zwei in den Analysen ohne diagnostische Stratifizierung nicht beobachtet wurden: SNP rs1801968 im *TOR1A* Gen (OR: 3,10; 95 % KI: 1,25 - 7,68; p = 0,0142) ist hierbei nominal signifikant mit dem Risiko für Schreibkrampf assoziiert, und SNP rs11655081 im *ARSG* Gen zeigte eine genomweit signifikante Assoziation (OR: 4,42; 95 % KI: 2,72-7,19; p = 2,11 × 10⁻⁰⁹) mit dem Risiko, an einer Musikerdystonie zu erkranken (siehe Tabelle 4.2). Nur die letzte Variante, rs11655081, behielt nach der Korrektur für multiples Testen einen signifikanten p-Wert. Allerdings basierte diese letztgenannte Metaanalyse auf zwei Datensätzen, die beide aus derselben Arbeit stammten (39). Dieselbe Variante wurde auch bei unabhängigen Datensätzen mit der fokalen (d.h. zervikale Dystonie, Blepharospasmus; Tabelle 3.2C, 3.2D) und der

segmentalen (Tabelle 1.3G) Formen von Dystonien untersucht. Die Ergebnisse dieser Untersuchungen zeigten allerdings keine Hinweise auf eine Assoziation mit diesen Dystonie-Subtypen (25).

Tabelle 3.2 Ergebnisse von Random-Effekt Metaanalysen auf der Basis eines additiven Modells in diagnostischen Subgruppen der isolierten Dystonie: a) alle Formen der fokalen Dystonie, b) Musikerdystonie, c) Blepharospasmus, d) zervikale Dystonie, e) spasmodische Dysphonie, f) Schreibdystonie, g) segmentale Dystonie, h) generalisierte Dystonie. Reproduziert von Ohlei et al., 2018 (25).

Chr-Position	Gene	SNP	A1/A2	AF*	E. †	Fälle vs. Kontrollen (#Datensätze /# Publikationen)	OR (95%KI)	P	I ²
A. Alle Formen der fokalen Dystonie									
1:236885200	<i>MRT</i>	rs1805087	G/A	0,19	C	227 vs 1072 (2/2)	0,93 (0,66-1,29)	0,5852	0,00
3:114171968	<i>DRD3</i>	rs6280	C/T	0,32	C	188 vs 200 (2/2)	0,82 (0,61-1,11)	0,1979	0,00
5:175441837	<i>DRD1</i>	rs155417	T/C	0,02	C	188 vs 200 (2/2)	0,73 (0,31-1,74)	0,4792	0,00
5:175443147	<i>DRD1</i>	rs4532	C/T	0,37	C	315 vs 1172 (3/3)	1,37 (1,13-1,67)	0,0015	0,00
8:42843330-1	<i>THAP1</i>	rs370983900 & rs184497763	AA/CT	0,03	C	1838 vs 1885 (2/2)	1,13 (0,82-1,56)	0,4483	0,00
9:129813148	<i>TOR1A</i>	rs3842225	-/C	0,26	C	747 vs 906 (4/3)	1,09 (0,92-1,29)	0,3423	0,00
9:129813781	<i>TOR1A</i>	rs1182	A/C	0,17	All	895 vs 1200 (5/4)	1,07 (0,93-1,27)	0,3105	0,00
9:129813781	<i>TOR1A</i>	rs1182	A/C	0,22	C	747 vs 906 (4/3)	1,07 (0,90-1,27)	0,4348	0,00
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,12	All	1116 vs 1905 (6/6)	1,06 (0,89-1,27)	0,5114	5,33
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,14	C	889 vs 1674 (4/4)	1,07 (0,87-1,31)	0,5170	13,98
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,13	A	227 vs 231 (2/2)	0,99 (0,55-1,80)	0,9872	41,70
9:129822779	<i>TOR1A</i>	rs2296793	A/G	0,22	All	729 vs 973 (5/4)	0,97 (0,81-1,16)	0,7446	14,35
9:129822779	<i>TOR1A</i>	rs2296793	A/G	0,24	C	612 vs 842 (4/3)	1,01 (0,83-1,23)	0,9066	9,29
9:133636819	<i>DBH</i>	rs2797849	C/G	0,34	C	490 vs 3054 (2/2)	1,09 (0,89-1,32)	0,4079	38,64
9:133639458	<i>DBH</i>	rs2797851	A/G	0,33	C	490 vs 3054 (2/2)	1,07 (0,87-1,34)	0,5395	42,38
9:133657065	<i>DBH</i>	rs1611131	G/A	0,29	C	490 vs 3054 (2/2)	1,12 (0,83-1,51)	0,4528	69,85
9:133657152	<i>DBH</i>	rs6271	T/C	0,08	C	490 vs 3054 (2/2)	2,97 (0,37-24,02)	0,3002	96,73
9:133663599	<i>DBH</i>	rs129886	T/C	0,19	C	490 vs 3054 (2/2)	1,01 (0,86-1,20)	0,8704	0,00
10:28431147	Interg.	rs1249277	G/C	0,15	C	464 vs 5515 (2/2)	0,77 (0,41-1,45)	0,4215	89,34
11:27658369	<i>BDNF</i>	rs6265	T/C	0,19	All	739 vs 2030 (4/4)	1,13 (0,91-1,41)	0,2678	43,72
11:27658369	<i>BDNF</i>	rs6265	T/C	0,19	C	589 vs 1813 (3/3)	1,07 (0,81-1,41)	0,1689	43,77
11:48246304	<i>OR4X2</i>	rs67863238	C/G	0,06	C	464 vs 5515 (2/2)	2,17 (0,40-11,70)	0,3482	96,50
11:113400106	<i>DRD2</i>	rs1800497	A/G	0,19	C	222 vs 267 (3/3)	0,89 (0,65-1,22)	0,4758	0,00
11:113425564	<i>DRD2</i>	rs1079597	T/C	0,17	C	188 vs 200 (2/2)	1,00 (0,69-1,46)	0,9913	0,00
11:113475529	<i>DRD2</i>	rs1799732	-/G	0,10	C	188 vs 200 (2/2)	0,74 (0,46-1,19)	0,2064	0,00
13:101406511	<i>NALCN</i>	rs1338041	C/A	0,41	All	665 vs 5804 (3/3)	0,90 (0,59-1,36)	0,6014	89,79
13:101406511	<i>NALCN</i>	rs1338041	C/A	0,36	C	464 vs 5515 (2/2)	0,75 (0,50-1,11)	0,1445	83,78
13:101407929	<i>NALCN</i>	rs9518385	C/A	0,36	C	339 vs 6145 (2/2)	0,73 (0,51-1,05)	0,0901	76,89
13:101430922	<i>NALCN</i>	rs61973742	G/A	0,1	All	665 vs 5804 (3/3)	1,38 (0,41-4,64)	0,6001	96,35
13:101430922	<i>NALCN</i>	rs61973742	G/A	0,07	C	464 vs 5515 (2/2)	1,85 (0,26-13,37)	0,5245	97,41
17:68368608	<i>ARSG</i>	rs1558877	T/C	0,48	C	357 vs 5272 (2/2)	0,96 (0,78-1,18)	0,6883	39,26
17:68368663	<i>ARSG</i>	rs1558878	C/T	0,49	C	357 vs 5272 (2/2)	0,88 (0,75-1,02)	0,0846	0,00
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,18	All	2007 vs 3608 (9/1)	1,34 (0,93-1,94)	0,1108	78,22
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,05	C	1877 vs 3409 (8/1)	1,41 (0,91-2,20)	0,1226	80,61
17:68395091	<i>ARSG</i>	rs7342975	G/A	0,01	C	357 vs 5272 (2/2)	2,00 (1,15-3,47)	0,0057	0,00
17:68395135	<i>ARSG</i>	rs9972951	A/G	0,03	C	357 vs 5272 (2/2)	2,16 (1,06-4,40)	0,0342	68,35
22:19950115	<i>COMT</i>	rs5993883	T/G	0,49	C	490 vs 3056 (2/2)	0,90 (0,61-1,31)	0,5671	83,57
22:19964609	<i>COMT</i>	rs4646316	T/C	0,23	C	490 vs 3056 (2/2)	1,08 (0,92-1,27)	0,3332	0,00
22:19968169	<i>COMT</i>	rs9332377	T/C	0,16	C	490 vs 3056 (2/2)	1,11 (0,92-1,34)	0,2678	0,00
B. Musikerdystonie									
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,14	C	239 vs 1213 (2/2)	0,79 (0,58-1,08)	0,1522	0,00
17: 68368608	<i>ARSG</i>	rs1558877	T/C	0,48	C	258 vs 5272 (2/2)	0,93 (0,69-1,25)	0,6176	67,05
17: 68368663	<i>ARSG</i>	rs1558878	C/T	0,49	C	258 vs 5272 (2/2)	0,87 (0,73-1,05)	0,1464	12,22
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,05	C	243 vs 1097 (2/1)	4,42 (2,72-7,19)	2,11e-09	0,00
17:68395091	<i>ARSG</i>	rs7342975	G/A	0,01	C	258 vs 5272 (2/2)	2,00 (1,15-3,47)	0,0141	0,00
17:68395135	<i>ARSG</i>	rs9972951	A/G	0,03	C	258 vs 5272 (2/2)	2,16 (1,02-4,55)	0,0425	66,99

Tabelle 3.1 (Fortsetzung)

Chr-Position	Gene	SNP	A1/A2	AF*	E. ‡	Fälle vs. Kontrollen (#Datensätze /# Publikationen)	OR (95%KI)	P	i ²
C. Blepharospasmus									
8:42843331	<i>THAP1</i>	rs370983900 & rs184497763	AA/CT	0,03	C	228 vs 1885 (2/2)	0,80 (0,39-1,65)	0,5470	3,80
9:129813148	<i>TOR1A</i>	rs3842225	-/C	0,26	C	205 vs 381 (2/1)	1,10 (0,70-1,72)	0,6807	52,83
9:129813781	<i>TOR1A</i>	rs1182	A/C	0,17	All	227 vs 675 (3/2)	1,12 (0,84-1,49)	0,4261	0,00
9:129813781	<i>TOR1A</i>	rs1182	A/C	0,22	C	205 vs 381 (2/1)	1,08 (0,73-1,59)	0,7185	38,15
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,12	All	125 vs 472 (3/2)	1,21 (0,60-2,44)	0,5994	53,92
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,11	A	94 vs 231 (2/2)	0,91 (0,38-2,19)	0,8377	44,02
9:129822779	<i>TOR1A</i>	rs2296793	A/G	0,22	All	281 vs 512 (3/2)	0,85 (0,66-1,09)	0,2000	0,00
9:129822779	<i>TOR1A</i>	rs2296793	A/G	0,24	C	205 vs 381 (2/1)	0,87 (0,65-1,16)	0,3309	0,00
11:27658369	<i>BDNF</i>	rs6265	T/C	0,19	All	289 vs 1004 (2/2)	1,44 (0,76-2,71)	0,2634	80,88
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,19	C	303 vs 1020 (3/1)	0,91 (0,63-1,33)	0,6372	0,00
D. Zervikale Dystonie									
8:42843331	<i>THAP1</i>	rs370983900 & rs184497763	AA/CT	0,03	C	712 vs 1885 (2/2)	1,12 (0,69-1,82)	0,6472	22,08
9:129813781	<i>TOR1A</i>	rs1182	A/C	0,17	All	448 vs 655 (2/2)	1,03 (0,82-1,29)	0,787	0,00
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,12	All	570 vs 833 (4/4)	1,19 (0,95-1,49)	0,1276	0,00
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,14	C	478 vs 602 (2/2)	1,27 (0,92-1,75)	0,1502	38,78
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,13	A	92 vs 231 (2/2)	0,93 (0,51-1,71)	0,8158	0,00
9:129822779	<i>TOR1A</i>	rs2296793	A/G	0,22	All	399 vs 492 (2/2)	1,00 (0,79-1,26)	0,9953	0,00
10:28431147	Interg.	rs1249277	G/C	0,15	C	464 vs 5515 (2/2)	0,77 (0,41-1,45)	0,4215	89,34
11:27658369	<i>BDNF</i>	rs6265	T/C	0,19	All	289 vs 1004 (2/2)	1,04 (0,84-1,29)	0,7138	0,00
11:48246304	<i>OR4X2</i>	rs67863238	C/G	0,06	C	464 vs 5515 (2/2)	2,17 (0,40-11,70)	0,3482	96,50
13:101406511	<i>NALCN</i>	rs1338041	C/A	0,41	All	665 vs 5804 (3/3)	0,90 (0,59-1,36)	0,6014	89,79
13:101406511	<i>NALCN</i>	rs1338041	C/A	0,36	C	464 vs 5515 (2/2)	0,75 (0,51-1,10)	0,1445	83,78
13:101430922	<i>NALCN</i>	rs61973742	G/A	0,1	All	665 vs 5804 (3/3)	1,38 (0,41-4,64)	0,6001	96,35
13:101430922	<i>NALCN</i>	rs61973742	G/A	0,07	C	464 vs 5515 (2/2)	1,85 (0,26-13,37)	0,5245	97,41
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,18	All	1051 vs 2233 (6/1)	1,01 (0,82-1,23)	0,9456	0,00
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,05	C	961 vs 2034 (5/1)	0,97 (0,76-1,24)	0,8145	0,00
E. Spasmodische Dysphonie									
8:42843333	<i>THAP1</i>	rs370983900 & rs184497763	AA/TC	0,03	C	482 vs 1885 (2/2)	1,47 (0,92-2,35)	0,104	0,00
F. Schreibkrampf									
9:129818622	<i>TOR1A</i>	rs1801968	G/C	0,14	C	55 vs 341 (2/1)	3,10 (1,25-7,68)	0,0142	65,44
G Segmentale Dystonie in allen Populationen									
8:42843333	<i>THAP1</i>	rs370983900 & rs184497763	AA/TC	0,03	C	267 vs 1885 (2/2)	1,09 (0,59-2,03)	0,7765	0,00
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,18	All	1,10 (0,76–1,59)	1,06 (0,80–1,40)	0,6835	0,00
17:68386068	<i>ARSG</i>	rs11655081	C/T	0,05	C	297 vs 1020 (3/1)	1,10 (0,76–1,59)	0,6078	7,72
H. Generalisierte Dystonie									
9:129818622	<i>THAP1</i>	rs370983900 & rs184497763	AA/CT	0,3	C	99 vs 1885 (2/2)	0,85 (0,30-2,43)	0,7672	0,00

Interg. = intergenische (d.h. zwischen annotierten Genen liegende) Region

Fettgedruckte Schrift kennzeichnet signifikante Ergebnisse vor Korrektur auf multiples Testen

‡ E. (Ethnizität). C = Europäische („kaukasische“) Abstammung (d. h. mehr als 90 % der Individuen in einem Datensatz gehörten dieser Abstammungsgruppe an); A = Asiatische Abstammung; All = Alle ethnischen Gruppen zusammen.

* Effekt-AF (Allelfrequenz) für jede Population, wie in der gnomAD-Datenbank (<http://gnomad.broadinstitute.org/>) aufgeführt. A – Asiatische Populationen beziehen sich auf gnomAD AF als "Ostasiatisch", wie in den Originalpublikationen berichtet. Für zwei Varianten, rs35153737 und rs3842225, waren die AF in gnomAD nicht verfügbar, so dass die AF stattdessen von dbSNP (<https://www.ncbi.nlm.nih.gov/SNP/>) übernommen wurde.

3.1.4 Zusammenfassung der Ergebnisse

Diese Arbeit stellte die erste systematische Field-Synopsis mit Metaanalysen von genetischen Assoziationsstudien bei Formen der isolierten Dystonie dar. Insgesamt wurden hierbei die Ergebnisse von 134 Metaanalysen zu 45 Varianten in 17 Genen vorgestellt. Nur die Varianten rs4532 in *DRD1* bei der isolierten Dystonie und die Variante rs11655081 in *ARSG* bei der

Musikerdystonie lieferten nach Korrektur für multiples Testen statistisch signifikante Assoziationen (25). In Abschnitt 3.3. werde ich auf diese und sämtliche andere Varianten, die in der Field-Synopsis mindestens nominal signifikante Assoziation mit Formen der isolierten Dystonie gezeigt haben, noch einmal eingehen.

3.2 Durchführung von genomweiten Assoziationsstudien in Dystonie-Datensätzen

3.2.1 Vorliegende Datensätze und Einbindung in die DFG-Forschungsgruppe „FOR2488“

In diesem Teil der Arbeit wurden die genomweiten Genotypisierungsdaten von 11 Datensätzen mittels des GWAS-Ansatzes untersucht. Die klinisch-demographischen Charakteristika der verwendeten Studienpopulationen wurden teilweise bereits beschrieben (62–65, 90, 91). Der Krankheitsstatus, die Datensatzbezeichnung, die Anzahl der Probanden, die Altersspanne und der Anteil der Frauen von 3.670 Dystonie-Patienten und 5.814 Kontrollen werden in der Tabelle 3.3 dargestellt. Sämtliche Phänotypdaten, die in den GWAS-Analysen Verwendung fanden, wurden vom Team des Z2-Projekts der FOR2488 sowie anderen klinischen Experten der FOR2488 zusammen- und für meine Analysen zur Verfügung gestellt (s. Abschnitt 2.2.1).

Tabelle 3.3 Demografische Merkmale von Dystonie-Patienten und Kontrollpersonen, die in diese Studie einbezogen wurden.

Dystonie Status	Herkunftsland	Datensatz	# Probanden	Altersspanne	Frauen (%)
Fall	Deutschland	Dystract	155	0-78	102 (65)
Fall	Deutschland	Tuebingen	648	0-80	400 (61)
Fall	Deutschland	Dystonia-Coalition	120	7-77	77 (64)
Fall	Deutschland	Luebeck	480	0-80	215 (44)
Fall	USA	Dystonia-Coalition	1121	0-82	760 (68)
Fall	USA	Boston	1146	0-80	798 (69)
Kontrolle	Deutschland	Luebeck	41	n.v.	6 (14)
Kontrolle	Deutschland	DeNoPa	97	44-84	36 (37)
Kontrolle	Deutschland	SFB134	562	25-52	347 (61)
Kontrolle	Deutschland	PopGen	3147	18-81	1546 (49)
Kontrolle	Deutschland	EPIPARK	652	40-79	320 (49)
Kontrolle	USA	GHC	261	43-85	96 (36)
Kontrolle	USA	PEG	1054	n.v.	558 (53)

n.v. = nicht vorhanden

3.2.2 Genomweite Genotypisierungsexperimente mittels „Global Screening Array“

Wie schon im Abschnitt 2.2.2. beschrieben, wurden für diese Arbeit DNA-Proben am IKMB in Kiel mittels „Global Screening Array“ genotypisiert. Insgesamt wurden hierbei bis zu 696.375 genetische

Marker je Proband gemessen und wurden von dem Genotypisierungslabor am IKMB in Kiel zu unserer Arbeitsgruppe zur weiteren Prozessierung und statistischen Analyse übertragen.

3.2.3 Analysen der autosomalen Chromosomen

3.2.3.1 *Qualitätskontrolle der Genotypisierungsdaten und Genotyp-Imputation*

Die QC der Genotypisierungsdaten und die Genotyp-Imputation wurden mittels eines von Dr. Fabian Kilpert entwickelten bioinformatischen Workflows durchgeführt (siehe Abschnitt 2.2.3.1). Insgesamt passierten 435.244 SNPs die QC-Kriterien. Im zweiten Schritt wurden die nicht genotypisierten genetischen Varianten mit der Minimac3-Software anhand des HRC-Referenzpanels (EGAD00001002729, v1.1) imputiert. Dadurch entstand eine Gesamtanzahl von 39.131.578 DNA-Varianten für weitere Analysen.

3.2.3.2 *Vorbereitende Analysen und Datenfilterungen*

Die Dystonie-Datensätze wurden in zwei Gruppen nach dem Ursprung der DNA-Samples in eine „deutsche“ und „amerikanische“ Unterpopulation aufgeteilt. Nach der Anwendung weiterer Qualitätskriterien (beschrieben in Abschnitt 2.2.3.2) in den Datensätzen der beiden Gruppen, wurden SNPs aufgrund folgender Kriterien ausgeschlossen:

- Imputations-Qualitäts-Score $r^2 < 0,7$: ($n_{GER} = 27.605.813$, $n_{USA} = 27.605.813$)
- Frequenz des seltenen Allels $< 0,5\%$ ($n_{GER} = 3.117.590$, $n_{USA} = 3.103.061$)
- SNPs mit Genotypisierungsrate $< 0,98\%$ ($n_{GER} = 0$, $n_{USA} = 0$)
- SNPs, die das Hardy-Weinberg-Equilibrium verletzten ($n_{GER} = 86$, $n_{USA} = 69$)

Die Gesamtzahl der zu untersuchenden Varianten wurde dadurch auf 8.408.082 SNPs in der deutschen und 8.422.635 SNPs in der amerikanischen Gruppe reduziert. Des Weiteren wurde eine QC auf Ebene der DNA-Samples durchgeführt, die zum Ausschluss folgender Patientenzahlen führte:

- Personen, bei denen das angegebene mit dem genetischen Geschlecht nicht übereinstimmte ($n_{GER} = 100$, $n_{USA} = 34$)
- Duplizierte DNAs in den Proben ($n_{GER} = 48$, $n_{USA} = 62$, führte zum Ausschluss einer Probe)
- Verwandtschaftsverhältnisse bis zum vierten Grad ($n_{GER} = 237$, $n_{USA} = 54$)
- Proben mit implausibler Heterozygotie ($n_{GER} = 8$, $n_{USA} = 8$)

Nach Anwendung der Qualitätsfilter wurde eine PCA zur Bestimmung der genetischen Abstammung in der Gruppe mit den amerikanischen Probanden sowie in der Gruppe mit den deutschen Probanden mit der Hilfe der PLINK-Software (v1.9) durchgeführt. Dabei wurden die Ausreißer auf der Grundlage der FDR ($p < 0,05$) bestimmt und ausgeschlossen. Insgesamt traf dies auf 8

Studienteilnehmer aus den USA und 19 Studienteilnehmer aus Deutschland zu, die damit aus den nachfolgenden Analysen ausgeschlossen wurden.

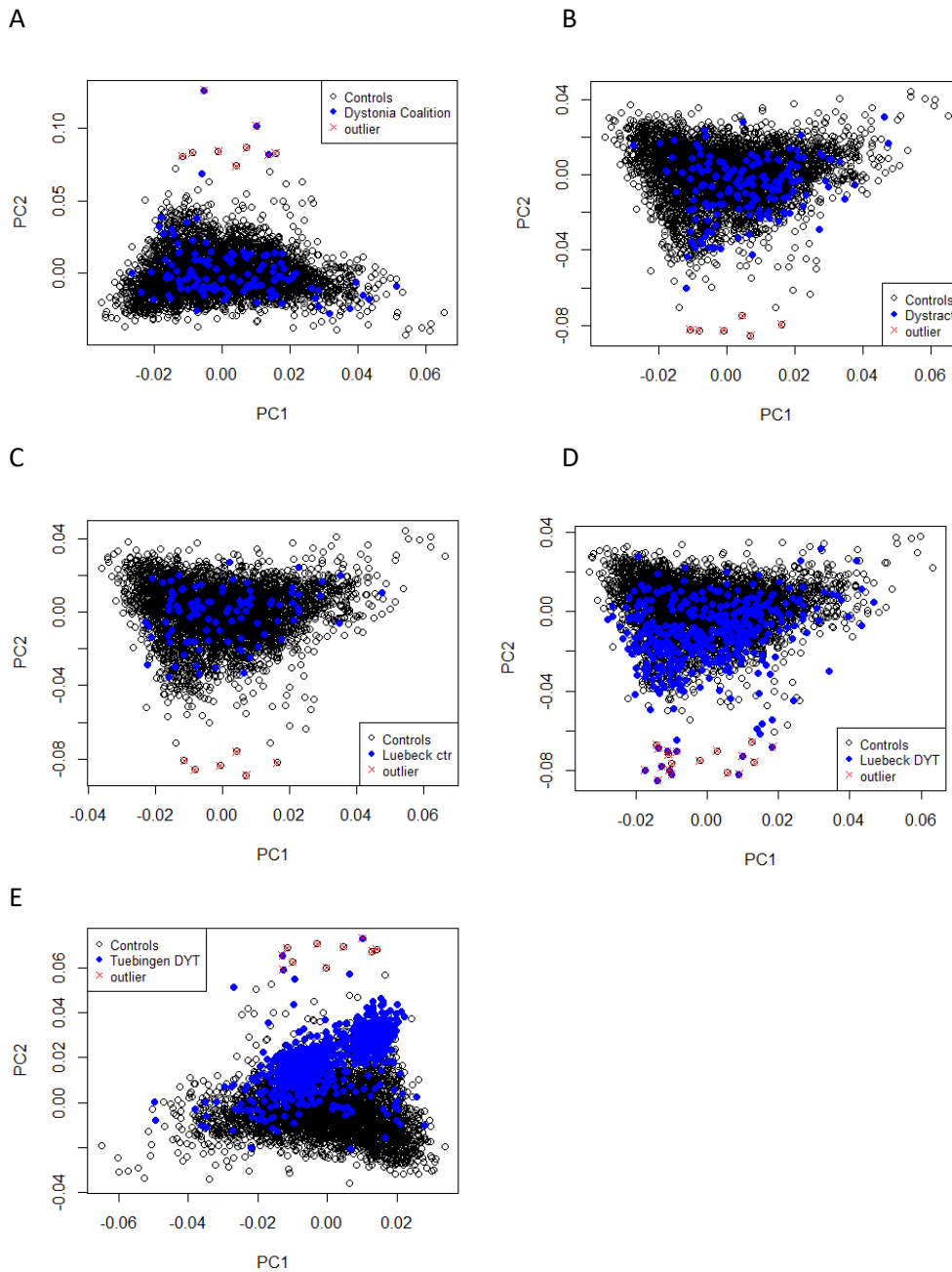


Abbildung 3.2 Plot der Eigenvektoren der Hauptkomponenten 1 und 2 (PC1 und PC2) der PCA-Analysen. Die Ausreißer werden basierend auf der FDR ($p < 0,05$) bestimmt und sind in Rot auf dem Plot markiert. Die Gruppe der gesunden Probanden besteht aus dem PopGen, Denopa, Epipark und SFB134 Datensätzen. A) Gesunde Probanden mit der Dystonie-Coalition Gruppe. B) Gesunde Probanden mit der Dystrakt Gruppe. C) Gesunde Probanden im Vergleich zu den Probanden aus Lübeck ohne Dystonie Erkrankung. D) Gesunde Probanden im Vergleich zu den Probanden aus Lübeck mit der Dystonie Erkrankung. E) Gesunde Probanden mit der Tübinger Gruppe.

Für DNA-Samples deutschen Ursprungs fehlten weitestgehend die Angaben zur Ethnizität. Die geografische Zugehörigkeit der Probanden wurde daher anhand der PCA bestimmt. Diese wurde getrennt mit jedem Dystonie-Datensatz in Verbindung mit den PopGen, Denopa, Epipark und

SFB134 Gruppen als Referenzdatensatz bestimmt. Wie im Kapitel 2.2.3.2 beschrieben, wurden auch hier Ausreißer auf Grundlage der FDR ($p < 0,05$) definiert. Dadurch wurden weitere 25 Studienteilnehmer ausgeschlossen. Die Eigenvektoren der PC1 und PC2 in den Dystonie-Datensätzen deutschen Ursprungs werden in Abbildung 3.2 exemplarisch dargestellt.

Die nach der QC verbleibenden Dystonie-Patienten wurden nachfolgend in 12 klinische Dystonie-Subgruppen eingeteilt. Die Anzahl an Probanden in den diagnostischen Dystonie-Gruppen, die in die GWAS eingeschlossen wurden, sind in Tabelle 3.4 dargestellt.

Tabelle 3.4 Überblick über die deutschen und amerikanischen Datensätze, die in die GWAS-Analysen einbezogen wurden.

Diagnose	# aus Deutschland	# aus USA	# Metaanalyse
<i>Fokale</i>	955	1.157	2.112
Zervikale	508	309	817
Laryngeale	1 (n.v.)	398	n.v.
Kraniale	140	169	309
Dystonie der oberen Extremitäten	233	110	343
Unbekannter Ort und andere ¹	69 (n.v.)	158 (n.v.)	n.v.
Aufgabenspezifische Dystonie ²	190	1 (n.v.)	n.v.
<i>Multifokale + segmentale</i>	173	404	577
<i>Generalisierte</i>	42 (n.v.)	67	n.v.
<i>Unbekannter Ort und Hemidystonie</i>	35 (n.v.)	5 (n.v.)	n.v.
Alle isolierten Dystonien	1 205	1.633	2.838
Gesunde Probanden	4 287	925	5.212

n.v. = nicht verwendet aufgrund zu geringer Stichprobengröße ($n < 50$ in der Fall-Gruppe) oder wegen unbekannter Lokalisation der Dystonie.

¹ Zu dieser Gruppe gehören Fälle mit trunkaler Dystonie und mit einer Dystonie der unteren Gliedmaße.

² Der aufgabenspezifische Dystonie-Datensatz umfasst Fälle mit Dystonien der oberen und unteren Gliedmaße sowie einige Fälle, bei denen eine Dystonie mit unbekannter Lokalisation diagnostiziert wurde. N=126 der aufgabenspezifischen Dystonie-Fälle wurden als Musiker-Dystonie klassifiziert.

3.2.3.3 Durchführung der GWAS und der GWAS-Metaanalysen

Insgesamt standen 2.838 Dystonie-Fälle und 5.212 Kontrollen für die GWAS-Analysen zur Verfügung. Davon stammten 5.492 (68 %; 1.205 Dystonie-Fälle, 4.287 Kontrollen) aus Deutschland und 2.558 (32 %; 1.633 Dystonie-Fälle, 925 Kontrollen) aus den USA. Sieben GWAS wurden in der deutschen Stichprobe (entsprechend der klinischen Informationen zu sieben vorliegenden Dystonieformen) und acht GWAS wurden (entsprechend der hierfür vorliegenden acht Dystonieformen) bei der amerikanischen Stichprobe berechnet (Tabelle 3.4). Für sechs der Dystonieformen waren ausreichend Daten vorhanden, um sie per Metaanalyse zusammenzufassen.

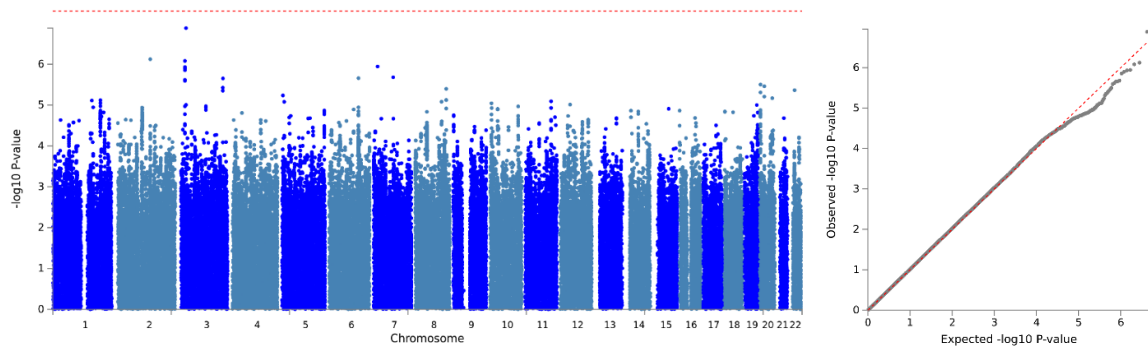
Exemplarisch werden im Hauptteil der Arbeit Manhattan und QQ-Plots der GWAS zu isolierter Dystonie, kranialer Dystonie und der Dystonie der oberen Extremitäten gezeigt (Abbildungen 3.3 – 3.5), sämtliche anderen Manhattan- und QQ-Plots der Ergebnisse der GWAS und GWAS-

Metaanalysen werden in den Ergänzungsabbildungen A 1-18 aufgeführt. Alle „Topvarianten“ (d.h. SNPs mit einem $p < 5 \times 10^{-6}$; nachfolgend als „suggestiv signifikant“ bezeichnet) aus allen Metaanalysen zur isolierten Dystonie sowie zu diagnostischen Unterformen sind in Ergänzungstabelle S2.1-S2.9 auf der beiliegenden *CD-ROM* enthalten. Die genetische Heterogenität der Dystonie-Datensätze wurde mit dem genomischen Inflationsfaktor „*Lambda*“ (λ) geschätzt (Tabelle A 1 im Anhang). Das λ in den Dystonie-Datensätzen variierte zwischen 0,939 und 1,04 und liefert damit keine Evidenz für eine nennenswerte Inflation der genomweiten Teststatistiken. In Tabelle 3.5 sind alle genomweit signifikanten ($p < 5 \times 10^{-8}$) SNPs der Metaanalysen zusammengefasst.

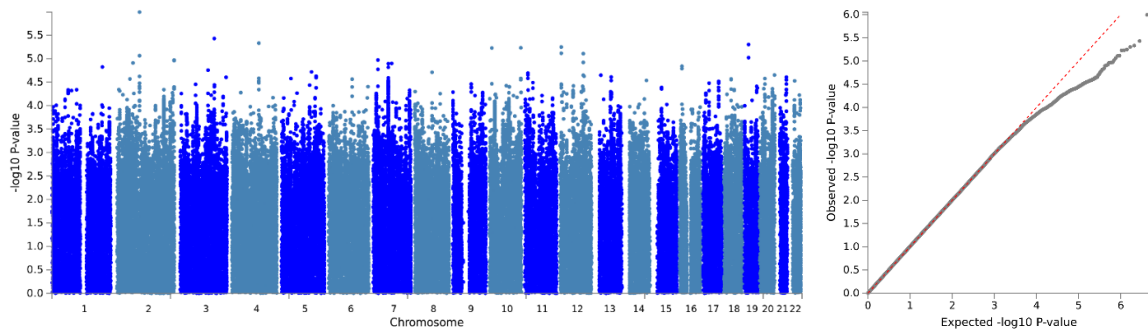
GWAS-Ergebnisse, die alle isolierten Dystonie-Fälle zusammenfassen.

Der Datensatz, der alle isolierten Dystonie-Fälle zusammenfasst, ist naturgemäß der umfangreichste Datensatz, welcher in dieser Arbeit untersucht wird und beinhaltet 2.838 Fälle aus allen klinisch-diagnostischen Gruppen sowie 5.212 Kontrollen. Die Manhattan-Plots zu der GWAS der isolierten Dystonie und der Metaanalyse werden in der Abbildung 3.3 dargestellt. Die Inflation der genomweiten Teststatistiken in der deutschen GWAS ($\lambda = 1,03$), der amerikanischen GWAS ($\lambda = 1,02$) und bei der GWAS-Metaanalyse ($\lambda = 1,03$) ist jeweils sehr gering. Dies zeigt, dass die Ergebnisse wahrscheinlich nicht aufgrund eines systematischen Fehlers in den Daten verzerrt wurden. Wie aus Abbildung 3.3 ersichtlich ist, gab es keine zu dieser Dystonieform durchgeführten GWAS genomweit signifikante Ergebnisse. Bei den SNP-basierten Ergebnissen erreichten allerdings eine Reihe von Markern die Signifikanzschwelle für genomweit suggestiv signifikante Assoziation ($p < 5 \times 10^{-6}$), z. B. rs6899902 auf Chromosom 6 (OR: 5,20; $p = 1,24 \times 10^{-6}$) und rs6590967 auf Chromosom 11 (OR: 0,81; $p = 8,93 \times 10^{-7}$). Der letztgenannte SNP befindet sich in einem Intron von *RP11-31506.2*, das in der genbasierten GWAS ebenfalls genomweit suggestiv signifikante Hinweise auf eine Assoziation erhielt ($z_{\text{meta}} = 4,38$; $p_{\text{meta}} = 5,91 \times 10^{-6}$; Tabelle 3.5; Abb.3.3 und die Abb. A 18 im Anhang). *RP11-31506.2* ist als lange nicht kodierende RNA (engl. *long non-coding RNA*, [lncRNA]) annotiert, die an der Expressionsregulierung des Transmembranproteins 123 (*TMEM123*) beteiligt ist. *RP11-31506.2* kodiert für einen mutmaßlichen Zelloberflächenrezeptor, der mit dem onkotischen Zelltod in Verbindung gebracht wurde (<https://www.uniprot.org/uniprot/Q8N131>) und ubiquitär im menschlichen Gehirn exprimiert wird (GTEx [v8]) (84). In Übereinstimmung mit der potenziellen Regulierung der *TMEM123*-Expression durch die lncRNA *RP11-31506.2* steht die Beobachtung, dass die Hauptvariante in unserer GWAS-Metaanalyse (rs6590967) laut GTEx (v8) als eQTL von *TMEM123* im Schilddrüsengewebe fungiert.

A



B



C

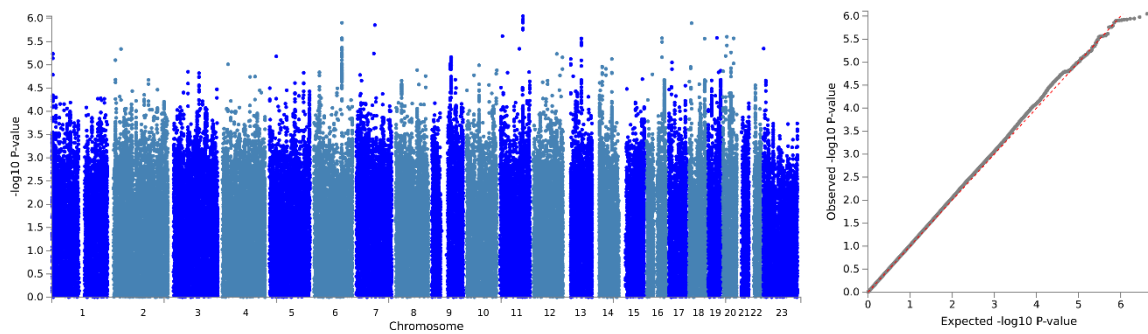


Abbildung 3.3 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der **isolierten Dystonie** A) Deutscher Datensatz B) Amerikanischer Datensatz C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie kennzeichnet das genomweite Signifikanzniveau (5×10^{-08}). NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (siehe Kapitel Methoden).

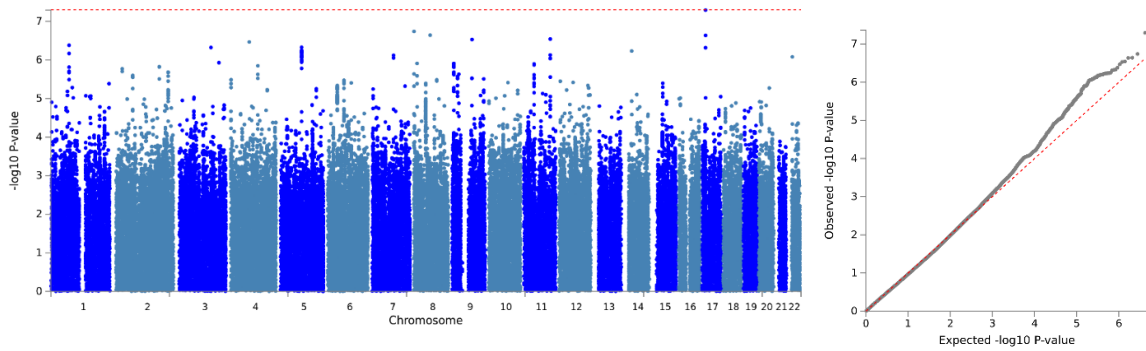
Während der GWAS-Katalog mehrere menschliche Phänotypen / Krankheiten auflistet, die mit SNPs in der allgemeinen *TMEM123*-Region auf Chromosom 11 assoziiert sind, zeigt keines davon eine eindeutige Verbindung zur ZNS-/Gehirnfunktion. Zusammenfassend lässt sich feststellen, dass das suggestiv signifikante Signal auf Chromosom 11q22 in der Nähe des *TMEM123*-Gens liegt. Zukünftige Arbeiten müssen die mögliche Rolle dieses Gens bei der Pathogenese der Dystonie aufklären.

GWAS zu kranialer Dystonie

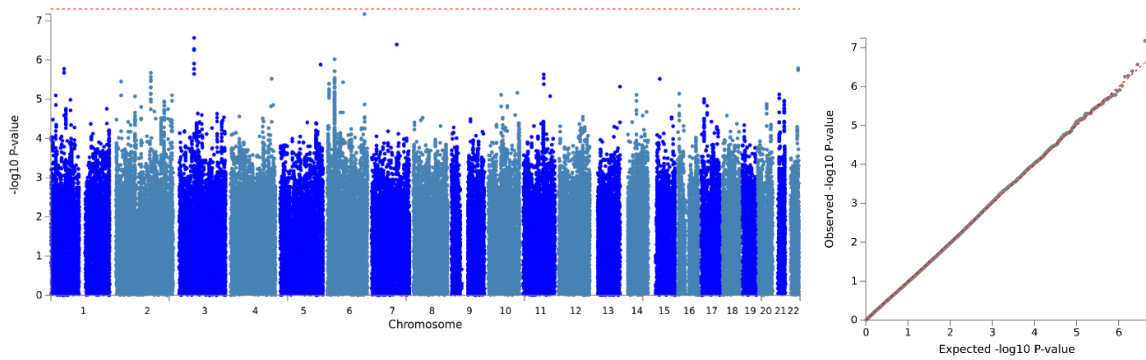
Die DNA-Proben mit deutschem Ursprung umfassten nach der QC 140 Fälle mit kranialer Dystonie und 4.287 Kontrollen. Nach der QC beliefen sich die DNA-Proben amerikanischen Ursprungs auf 169 Fälle mit kranialer Dystonie und 925 Kontrollen. Die Ergebnisse der GWAS der beiden Stichproben wurden anschließend metaanalysiert, d.h. mit einer effektiven Stichprobengröße von 309 Fällen und 5.212 Kontrollen. Die Manhattan-Plots und die QQ-Plots zur kranialen Dystonie GWAS und zu der Metaanalyse werden in Abbildung 3.4 dargestellt. Die Inflation der genomweiten Teststatistiken in der deutschen GWAS ($\lambda = 0,99$), der amerikanischen GWAS ($\lambda = 1,01$) und bei der GWAS-Metaanalyse ($\lambda = 1,01$) ist jeweils sehr gering. Für die GWAS der kranialen Dystonie wurde das stärkste – und genomweit signifikante – SNP-basierte Signal mit rs143591361 auf Chromosom 6q12 beobachtet ($OR_{meta}: 3,82$; $p_{meta} = 5,52 \times 10^{-09}$; Tabelle 3.5, Abbildung 3.4). An diesem Locus erreichten insgesamt 30 SNPs einen p-Wert unter 5×10^{-06} , von denen neun SNPs die Schwelle der genomweiten Signifikanz unterschritten (ergänzende Tabelle S2.5). Trotz der konsistenten genetischen Assoziationsnachweise in dieser Region sind die potenziellen funktionellen Auswirkungen dieses Signals nach wie vor unklar. Der führende SNP (rs143591361) befindet sich zwischen zwei lncRNAs (*RP11-542F9.1* und *RP11-406O16.1*) mit unbekannter Funktion. Beide werden im menschlichen Gehirn nur in geringen Mengen exprimiert (GTEx v8) (84). Während der GWAS-Katalog zwei ZNS-relevante Einträge in der allgemeinen assoziierten Region (<https://www.ebi.ac.uk/gwas/regions/chr6:68671305-68972804>) auflistet, d. h. einen mit der Messung des Volumens des oberen parietalen Kortex und einen mit kognitiver Beeinträchtigung, bleibt die Korrelation zu unserem *lead*-SNP und/oder lncRNAs ungewiss. Das zweitstärkste Signal wurde mit rs3094685 auf Chromosom 6p21 beobachtet ($OR_{meta}: 1,89$; $p_{meta} = 3,68 \times 10^{-08}$). Dieser SNP ist dem Haupthistokompatibilitätskomplex (MHC) zugeordnet und wird in der Nähe des HLA-B-Gens annotiert. Insgesamt erreichten 111 SNPs einen p-Wert von weniger als 5×10^{-06} (siehe ergänzende Tabelle S2.5), und ein SNP (rs3094685) unterschritt die Schwelle zur genomweit signifikanten Assoziation. Darüber hinaus zeigten die genbasierten Analysen suggestive Hinweise auf einen Zusammenhang mit *DDX39B* ($z_{meta} = 4,04$; $p_{meta} = 2,66 \times 10^{-05}$; Tabelle 3.5). Während der *lead*-SNP (rs3094685) nicht auf GTEx aufgeführt wird, zeigt ein Proxy-SNP (rs3132476, r^2 mit

rs3094685 = 0,8) mehrere hoch signifikante eQTL-Effekte im Gehirn auf GTEx (am stärksten mit der ncRNA *HCG27* im Kleinhirn). Die beiden anderen Assoziationssignale mit der kranialen Dystonie waren SNP rs12022195 auf Chromosom 1q44 ($OR_{meta}: 2,48; p_{meta} = 4,88 \times 10^{-08}$) und rs76357386 auf Chromosom 16p13 ($OR_{meta}: 2,58; p_{meta} = 4,76 \times 10^{-08}$). Diese zeigten jedoch nur eine grenzwertige genomweite Signifikanz und korrelierten mit vergleichsweise wenigen zusätzlich assoziierten SNPs, so dass diese GWAS-Signale weit weniger überzeugend wirken als die oben beschriebenen Signale auf Chromosom 6. Zusammenfassend lässt sich feststellen, dass der GWAS-Locus in der MHC-Region für die Analysen der kranialen Dystonie besonders interessant ist, weil er einen ersten überzeugenden Hinweis darauf liefert, dass ein (oder mehrere) Gen(e), die an der Reaktion des erworbenen Immunsystems beteiligt sind, in die Pathogenese der Dystonie involviert sein könnten. Allerdings ist dies nur ein erster Hinweis, der in unabhängigen Datensätzen überprüft und in funktionellen Studien weiter charakterisiert werden muss.

A



B



C

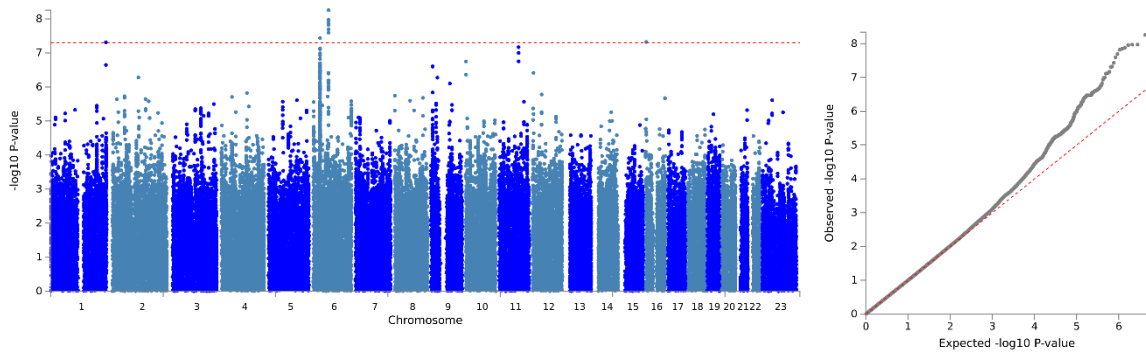
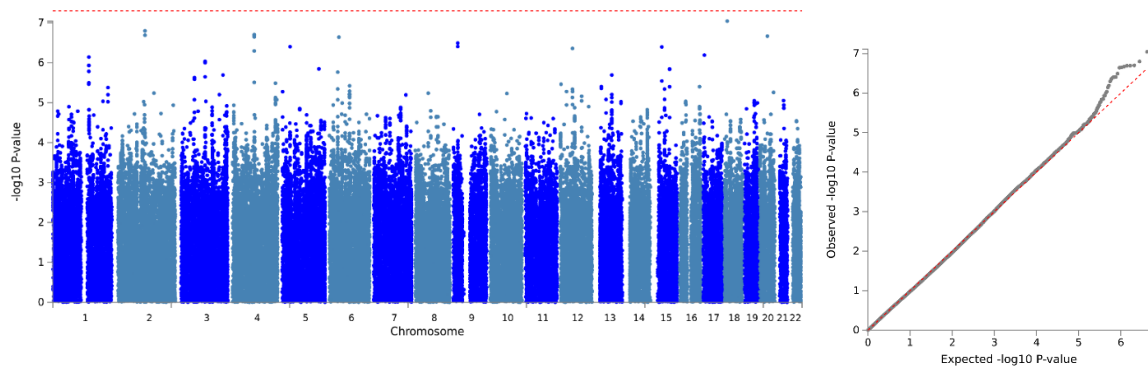


Abbildung 3.4 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der **kranialen Dystonie** A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie kennzeichnet das genomweite Signifikanzniveau (5×10^{-08}). NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (siehe Kapitel Methoden).

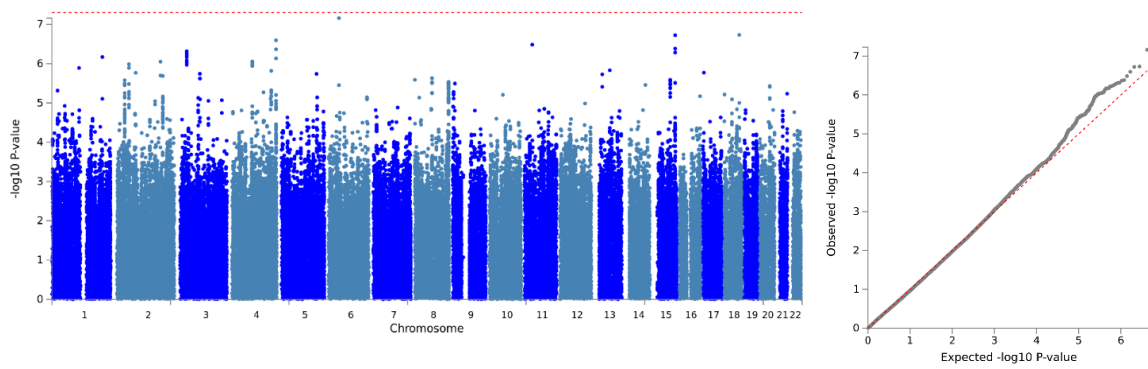
GWAS zu Dystonie der oberen Extremitäten

Die DNA-Proben mit deutschem Ursprung umfassten nach der QC 233 Fälle mit Dystonie der oberen Extremitäten und 4.287 Kontrollen. Nach der QC enthielten die DNA-Proben mit amerikanischem Ursprung 110 Fälle mit Dystonie der oberen Extremitäten und 925 Kontrollen. Die Ergebnisse der GWAS der beiden Stichproben wurden anschließend metaanalysiert, d.h. mit einer effektiven Stichprobengröße von 343 Fällen und 5.212 Kontrollen. Die Manhattan-Plots und QQ-Plots zu den GWAS-Analysen der Dystonie der oberen Extremitäten und der Metaanalyse werden in der Abbildung 3.5 dargestellt. Die Inflation der genomweiten Teststatistiken in der deutschen GWAS ($\lambda = 1,01$), der amerikanischen GWAS ($\lambda = 0,99$) und bei der GWAS-Metaanalyse ($\lambda = 0,99$) ist auch hier jeweils sehr gering. Bei den Analysen erreichten zwei SNPs in der Metaanalyse, die die deutschen und amerikanischen Datensätze kombinierte, genomweite Signifikanz (Tabelle 3.5, Abbildung 3.5). Das stärkste Signal wurde mit rs182982101 auf Chromosom 3q12 beobachtet ($OR_{meta}: 7,48$; $p_{meta} = 2,95 \times 10^{-08}$). Während dieses Signal in Bezug auf die Effektgröße in beiden unabhängigen Datensätzen recht konsistent war ($OR_{Ger}: 7,47$; $p_{Ger} = 2,25 \times 10^{-06}$ und $OR_{USA}: 7,50$; $p_{USA} = 3,80 \times 10^{-03}$; Tabelle 3.5), stellte sich heraus, dass die MAF bei gesunden Probanden für den *lead*-SNP niedrig war ($MAF_D = 0,4\%$, $MAF_{USA} = 0,4\%$; $MAF \sim 0,7\%$ in GnomAD). Hingegen war die MAF bei den Dystone Erkrankten deutlich höher ($MAF_D = 2\%$, $MAF_{USA} = 2\%$). In Kombination mit der vergleichsweise geringen Stichprobengröße für diesen Dystonie-Subtyp ($n=342$ in der Metaanalyse) ist die Wahrscheinlichkeit höher als bei den anderen in unserer Studie entdeckten Loci, dass dieses Signal einen falsch-positiven Befund darstellt. Das zweite Signal wurde durch den SNP rs4081356 auf Chromosom 1p31 hervorgerufen ($OR_{meta}: 1,63$; $p_{meta} = 3,85 \times 10^{-08}$). Auch hier waren die Assoziationsnachweise in den verschiedenen Datensätzen bemerkenswert konsistent ($OR_{Ger}: 1,60$; $p_{Ger} = 1,26 \times 10^{-05}$ und $OR_{USA}: 1,69$; $p_{USA} = 7,99 \times 10^{-04}$; Tabelle 3.5), und es gab sieben weitere SNPs in dieser Region, die zumindest suggestiv signifikante Hinweise auf eine Assoziation zeigten ($p < 5 \times 10^{-06}$; ergänzende Tabelle S2.6). Interessanterweise lagen die *lead*-SNPs für beide GWAS-Signale auch in der Nähe von lncRNA-Sequenzen (d. h. *RP11-393N21.1* und *RP11-221J22.1*; Tabelle 3.5), was auf eine mögliche Rolle bei der Transkriptionsregulation schließen lässt. Ähnlich wie bei den anderen oben hervorgehobenen ncRNAs ist die Expression der beiden lncRNAs, die hier mit der Dystonie der oberen Gliedmaßen assoziiert sind, gering (auch im Gehirn). Schließlich listet der GWAS-Katalog zwar einige bemerkenswerte Assoziationssignale in der allgemeinen Region des neuen Chromosom-1p-Locus (<https://www.ebi.ac.uk/gwas/regions/chr1:67882019-68183518>) auf, einschließlich der Hirnmorphologie und der kognitiven Leistung, doch bleibt seine Relevanz für die Dystonie unklar und erfordert weitere Untersuchungen.

A



B



C

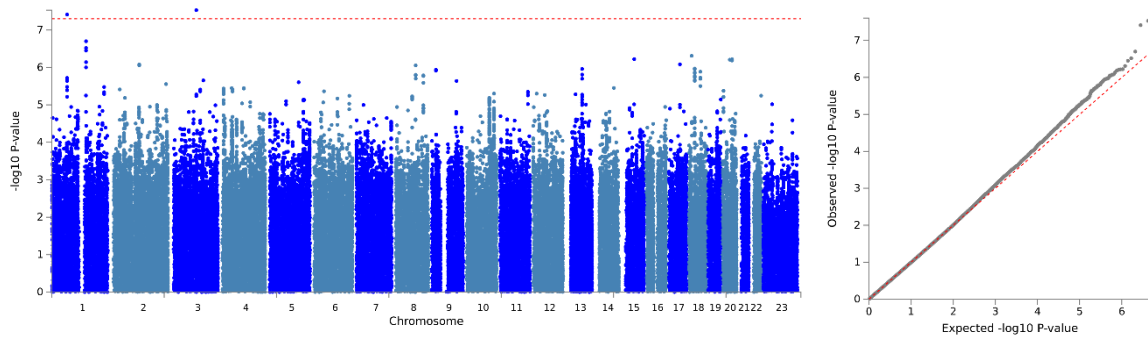


Abbildung 3.5 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der **Dystonie der oberen Extremitäten** A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie kennzeichnet das genomweite Signifikanzniveau (5×10^{-8}). NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (siehe Kapitel Methoden).

Tabelle 3.5 Eine Übersicht der genomweit signifikanten *lead*-SNPs ($p < 5,0 \times 10^{-8}$) und Gene ($p < 2,6 \times 10^{-6}$) aus den GWAS-Metaanalysen zur kranialen Dystonie, der Dystonie der oberen Extremitäten sowie der segmentalen und multifokalen Dystonie.

Best SNP	p/q	Pos	A1 ¹	AF ²	Nächstliegendes Gen	Deutscher Datensatz GWAS		Amerikanische Datensatz GWAS		SNP-basierte Metaanalyse		Gene-basierte Metaanalyse
						OR	P	OR	P	OR	P	P, Z
A. Kraniale Dystonie (Metaanalyse)												
rs12022195	1q44	245289386	C	0,06	EFCAB2	2,61	$1,74 \times 10^{-5}$	2,33	$7,24 \times 10^{-4}$	2,48	$4,88 \times 10^{-8}$	$2,28 \times 10^{-1}$, 0,7
rs3094685	6p21.33	31296409	T	0,18	DDX39B ³	1,65	$2,81 \times 10^{-3}$	2,14	$1,93 \times 10^{-6}$	1,89	$3,68 \times 10^{-8}$	$2,66 \times 10^{-5}$, 4,0
rs143591361	6q12	68822055	T	0,01	RP11-542F9.1	3,99	$3,37 \times 10^{-6}$	3,59	$4,17 \times 10^{-4}$	3,82	$5,52 \times 10^{-9}$	n.v.
rs76357386	16p13.3	2914276	A	0,04	LA16c-325D7.2	2,86	$1,02 \times 10^{-5}$	2,07	$1,05 \times 10^{-3}$	2,58	$4,76 \times 10^{-8}$	n.v.
B. Dystonie der oberen Extremitäten (Metaanalyse)												
rs4081356	1p31.3	68032769	T	0,45	RP11-393N21.1	1,60	$1,26 \times 10^{-5}$	1,69	$7,99 \times 10^{-4}$	1,63	$3,85 \times 10^{-8}$	n.v.
rs182982101	3q12.3	101720936	T	0,0065	RP11-221J22.1:2	7,48	$2,26 \times 10^{-6}$	7,51	$3,80 \times 10^{-3}$	7,48	$2,95 \times 10^{-8}$	n.v.
C. Segmentale und multifokale Dystonie (Metaanalyse)												
rs6601370	8p23.1	9684764	T	0,31	TNKS ⁴	1,68	$3,55 \times 10^{-5}$	1,32	$4,91 \times 10^{-3}$	1,44	$1,91 \times 10^{-6}$	$2,47 \times 10^{-6}$, 3,7

¹A1 ist das getestete Allel und bezeichnet das Minor-Allel eines SNPs bei den gesunden Probanden.

²AF Allelhäufigkeit der nordwesteuropäischen Kontrollen nicht-finnischen Ursprungs aus GnomAD (<https://gnomad.broadinstitute.org/>).

³Das am nächsten gelegene Gen war *XXbac-BPG248L24.10*, aber das signifikanteste Gen in dieser Region ist *DDX39B*.

⁴Das am nächsten gelegene Gen war *snoU13*, aber das signifikanteste Gen in dieser Region ist *TNKS*.

3.2.4 Analyse des X-Chromosoms

3.2.4.1 Qualitätskontrolle der Genotypisierungsdaten und Genotyp-Imputation

Diese Arbeit befasst sich mit den GWAS-Analysen, die das X-Chromosom einbeziehen, welches aufgrund der methodischen Besonderheiten in GWAS zu komplex-genetischen Phänotypen nur vergleichsweise selten berücksichtigt wird. Nach meinem Wissen wurde eine solche „XWAS“ im Feld der Dystonie in dieser Arbeit zum ersten Mal durchgeführt. Vor der QC lagen insgesamt 19.994 genotypisierte Marker in den nicht-PAR Abschnitten des X-Chromosoms vor.

Tabelle 3.6 Ergebnisse der SNP-Level-QC für die genotypisierten SNPs auf dem X-Chromosom. In der Tabelle wird die absolute Anzahl der SNPs dargestellt.

	m/w	# deutsche Gruppe	# amerikan. Gruppe
Anzahl der SNPs nach der Genotypisierung		19.994	19.994
Ausschluss „Custom Content“ SNPs		2.916	2.916
Ausschluss von SNPs mit MAF < 0.01		4.335	4.335
SNPs mit Genotypisierungseffizienz < 0,98 %	w	243	270
	m	967	1.300
Abweichung von HWE in Kontrollen ($p < 1 \times 10^{-4}$)		12	2
Test auf differentielle Genotypisierungseffizienz		335	233
Test auf differentielle Allelfrequenz		1	6
Ausschluss von Allelen Löschung und Einfügung		2	2
Ausschluss von mehrdeutigen SNPs (C/G, A/T)		864	836
Verblieben für die Imputation		10.559	10.360

m = Männer-Gruppe, f = Frauen-Gruppe, # = Anzahl SNPs.

Nach der Anwendung der QC (Abschnitt 2.2.4.1) in den deutschen und amerikanischen Stichproben, wurde die Gesamtzahl der zur Verfügung stehenden X-Chromosom-Varianten auf 10.559 SNPs in der deutschen und 10.360 SNPs in der amerikanischen Stichprobe reduziert. Details zu ausgeschlossenen SNPs werden in der Tabelle 3.6 dargestellt. Ferner wurden auf Sample-Ebene hier nur Studienteilnehmer berücksichtigt, welche die QC der autosomalen Analysen bestanden haben (s.o.). Danach wurden noch weitere 20 Probanden aus der deutschen Stichprobe und 19 Probanden aus der amerikanischen Stichprobe ausgeschlossen, weil ihre Genotypisierungseffizienz auf dem X-Chromosom <98 % war. Die Anzahl der Probanden in jeder der diagnostischen Dystonie-Gruppen ist in Tabelle 3.7 zusammengefasst.

Die Genotypdaten der weiblichen und männlichen Teilnehmer wurden dann getrennt mithilfe des Sanger Imputation Server (<https://imputation.sanger.ac.uk/>) imputiert (wie im Methodenteil Abschnitt 2.2.4.1 beschrieben Abschnitt).

Tabelle 3.7 Anzahl der Proben deutscher und amerikanischer Herkunft, welche die QC nach der Imputation bestanden haben und welche in der XWAS und in den Metaanalysen eingeschlossen sind

Diagnose	# aus Deutschland		# aus USA.		gesamt
	Frauen	Männer	Frauen	Männer	
<i>Fokale</i>	544	409	794	358	2.105
Zervikale	329	178	227	81	815
Laryngeale	1 (n.v.)	0 (n.v.)	294	100	n.v.
Kraniale	103	37	108	61	309
Dystonie der oberen Extremitäten	91	142	54	56	343
Unbekannter Ort und andere ¹	20 (n.v.)	52 (n.v.)	111 (n.v.)	60 (n.v.)	n.v.
Aufgabenspezifische Dystonie ²	48	141	0 (n.v.)	1 (n.v.)	n.v.
<i>Multifokale + segmentale</i>	121	50	275	122	568
<i>Generalisierte</i>	20	21	35	31	107
Unbekannter Ort und Hemidystonie	24 (n.v.)	11 (n.v.)	2 (n.v.)	3 (n.v.)	n.v.
Alle isolierten Dystonien	709	491	1.106	514	2.820
Gesunde Probanden	2.163	2.109	455	464	5.191

n.v. = nicht verwendet aufgrund zu geringer Stichprobengröße ($n < 50$).

¹ Zu dieser Gruppe gehören die trunkale Dystonie und die Dystonie der unteren Gliedmaßen.

² Der aufgabenspezifische Dystonie-Datensatz umfasst Fälle mit Dystonien der oberen und unteren Gliedmaßen sowie einige Fälle, bei denen eine Dystonie mit unbekannter Lokalisation diagnostiziert wurde.

3.2.4.2 Vorbereitende Analysen und Datenfilterungen

Vom Sanger Imputation Server wurden insgesamt 1.228.034 Varianten imputiert und heruntergeladen. Die Schritte der Post-Imputations-QC, die nachfolgend getrennt bei weiblichen und männlichen Teilnehmern durchgeführt wurden, sind in Tabelle 3.8 aufgeführt.

Tabelle 3.8 Ergebnisse der SNP-Level-QC für die imputierten SNPs auf dem X-Chromosom basierend auf den Daten aus den Gruppen der deutschen und amerikanischen Studienteilnehmern. Dargestellt wird die absolute Anzahl der genetischen Marker.

	w/m	# deutsche Gruppe	# amer. Gruppe
Anzahl der SNPs nach der Imputation		1.228.034	1 228.034
Ausschluss von SNPs mit INFO < 0.7		808.514	683.534
Ausschluss von SNPs mit MAF < 0.01	w	211.428	334.618
	m	210.781	332.860
SNPs mit Genotypisierungseffizienz < 0,98 %	w	0	0
	m	0	0
Abweichung von HWE in Kontrollen ($p < 1 \times 10^{-4}$)	w	141	3
	m	140	3
Verblieben für die XWAS Analyse	w	207.951	209.879
	m	208.599	211.637

m = Männer-Gruppe, f = Frauen-Gruppe, # = Anzahl an SNPs

3.2.4.3 Durchführung der XWAS und XWAS-Metaanalysen

Während in den XWAS zwei SNPs Hinweise auf eine suggestiv signifikante Assoziation erreichten ($p < 5 \times 10^{-06}$; ergänzende Tabelle S2.1, S2.5), erfüllte keiner der beiden die Kriterien für genomweite Signifikanz. Bei den Analysen, die beide Geschlechter durch eine Metaanalyse kombinierten (siehe Methoden Abschnitt 2.2.4.3), wurde das stärkste X-Chromosom-Signal mit dem SNP rs371872054 und dem Subtyp der kranialen Dystonie beobachtet (OR_f : 3,25; OR_m : 3,93; $p = 2,45 \times 10^{-06}$). Dieses Signal befindet sich ca. 18.000 bp stromabwärts des Gens, das für die ubiquitinspezifische Peptidase 11 (*USP11*) kodiert. Dieses Gen ist an der Regulierung von Signalwegen beteiligt (<https://www.uniprot.org/uniprot/P51784>) und zeigt eine sehr ausgeprägte Expression im menschlichen Gehirn (laut GTEx [v8]) (84). In Anbetracht des allgemeinen Mangels GWAS, die das X-Chromosom einbezogen haben sind im GWAS-Katalog keine Assoziationen in einem 300kb-Intervall um den assoziierten SNP aufgeführt. Vollständige Angaben zu allen geschlechtsspezifischen SNP-basierten Assoziationen für das X-Chromosom werden in der ergänzenden Tabelle S4.1 sowie in den ergänzenden Abbildungen A 19-27 dargestellt. Zusammenfassend lässt sich feststellen, dass die meisten geschlechtsspezifischen Signale bei Frauen (isolierte Dystonie, fokale Dystonie, kraniale Dystonie und aufgabenspezifische Dystonie) suggestiv signifikant sind, bei Männern jedoch nicht (außer für die segmentale und multifokale Dystonie).

3.2.5 Post-GWAS-Analysen

Die oben beschriebenen GWAS-Ergebnisse liefern nur erste (statistische) Hinweise auf eine mögliche Involvierung der hervorgehobenen SNPs in die Pathogenese der Dystonie. In den nachfolgend beschriebenen sog. Post-GWAS-Analysen habe ich eine Reihe von weiterführenden Berechnungen durchgeführt, die die funktionelle und epidemiologische Einordnung der primären GWAS-Befunde erleichtern soll.

3.2.5.1 Bestimmung, funktionelle Zuordnung und Annotation genetischer Loci

Für alle SNP-basierten GWAS-Ergebnisse mit mindestens suggestiv signifikanter Assoziation ($p < 5 \times 10^{-06}$) in den Metaanalysen wurden zuerst die Genloci mit Hilfe von FUMA annotiert, wie im Methodenteil Abschnitt 2.2.5.1 beschrieben. Ein Genlocus bezeichnet die physikalische Lage einer Assoziationsregion auf einem Chromosom. Für jeden unabhängigen SNP pro Genlocus ermittelte FUMA dann funktionelle Vorhersagen mittels weiterer *in silico* Vorhersageprogramme, z. B. basierend auf den CADD- und RegulomeDB-Scores (Ergänzungstabelle S2.1-2.9 auf der beigefügten CD-ROM enthalten). Des Weiteren wurden für jede Region mit mindestens einer genom-weit signifikanten Assoziation ein sog. Locus-Zoom-Plot erstellt (ebenfalls in FUMA; Abbildung 3.6 zeigt exemplarisch Locus-Zoom Plot der neu entdeckten Region auf Chromosom 6 für die kraniale Dystonie). Locus-Zoom-Plots stellen kleine Ausschnitte („Zooms“) der Manhattan-Plots dar, in denen die Korrelationen zwischen den GWAS-SNPs dargestellt und mit potenziell funktionell relevanten Informationen (hier: Lokalisation der bekannten Gene) verbunden werden. Die Abbildung. 3.6 zeigt den Locus-Zoom Plot für die Region des *lead*-SNPs rs3094685 (violett eingefärbt) aus der Meta-Analyse zur kranialen Dystonie, welcher in der Nähe des *HLA-B* Gens liegt. Am distalen Ende der Assoziationsregion sind zwei weitere Gene (*ATP6V1G2-DDX39B* und *DDX39B*) aufgeführt, die auch in den Gen-basierten GWAS-Analysen genomweit signifikante Assoziationen mit der kranialen Dystonie aufweisen (siehe Abschnitt 3.2.5.2). Alle weiteren Locus-Zoom Plots zu den Topregionen dieser Arbeit werden in den Abbildungen A 28-32 im Anhang dargestellt.

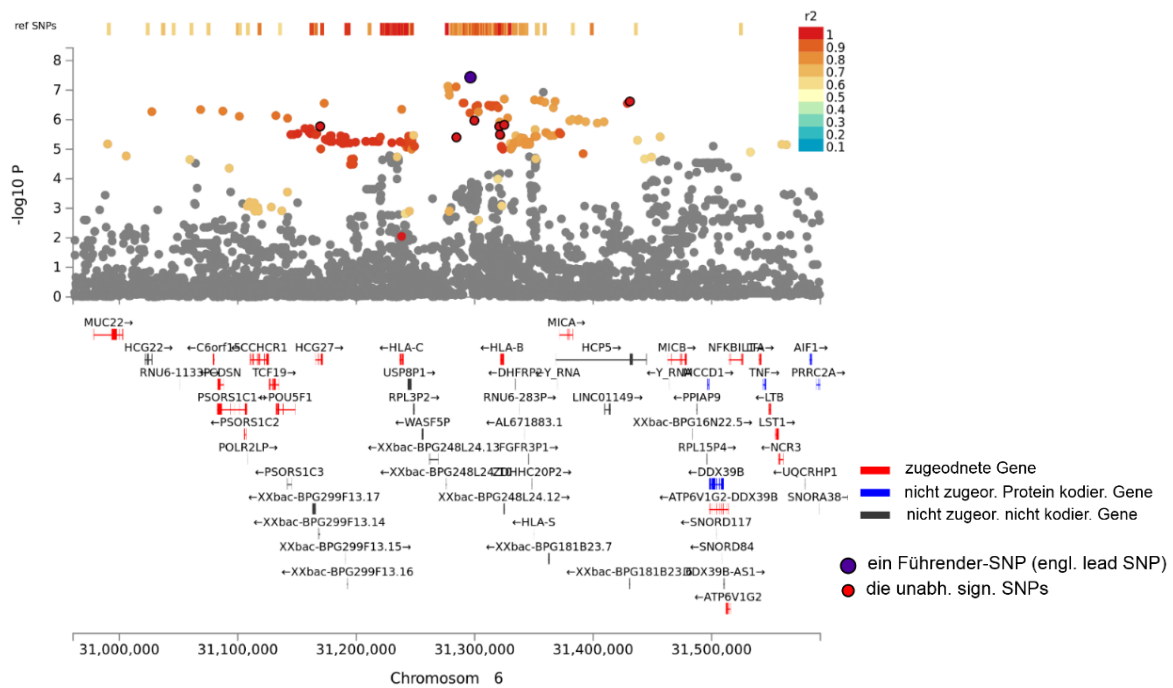


Abbildung 3.6 Locus-Zoom Plot der neu entdeckten Region auf chr. 6p21.33 für die **kraniale Dystonie**. Der *lead*-SNP in dieser Region ist rs3094685. r^2 quantifiziert das Kopplungsungleichgewicht zwischen den *lead*-SNP und anderen SNPs in der Region. Die Genzuordnungen und die Schätzungen des Kopplungsungleichgewichts wurden mit Hilfe des FUMA-Tools annotiert (Methoden).

3.2.5.2 Gen-basierte GWAS-Analysen

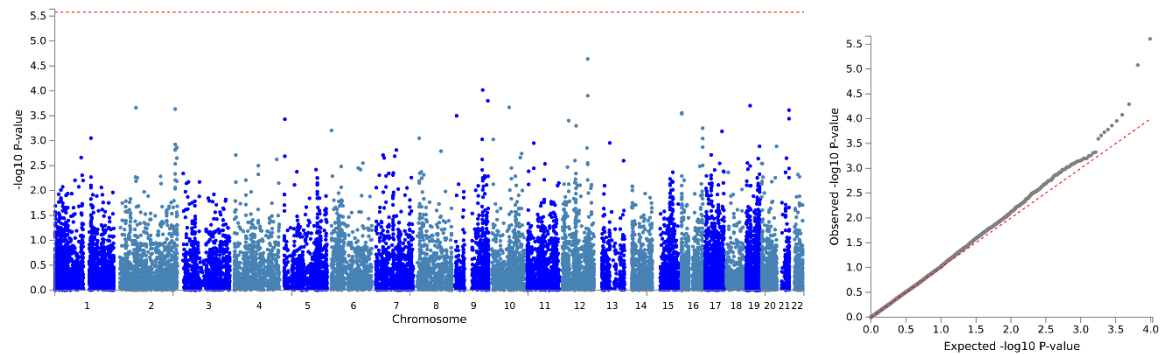
Für die Gen-basierten GWAS-Analysen wurden die GWAS-SNPs zu einem von 19.516 Genen zugeordnet, wenn sie in einem Intervall von ± 10 kb um dieses Gen lagen, wie im Abschnitt 2.2.5.1 und 2.2.5.2 beschrieben. Die Gen-basierten GWAS wurden mit Hilfe der in FUMA implementierten MAGMA-Software (Version 1.08) (82) berechnet. Hierbei werden innerhalb eines Gens die SNP-basierten GWAS p-Werte zu einer Gentest-Statistik mittels eines SNP-weiten Mittelwertmodells kombiniert. Die Ergebnisse dieser Analysen werden in Form von Manhattan- und QQ-Plots in der Abbildung 3.7 sowie in den Abbildungen A 2-18 im Anhang dargestellt. Wie unten noch kurz ausgeführt wird, wurden in diesen Analysen genomweit signifikante Gen-basierte Assoziationen ($2,6 \times 10^{-6}$) mit der segmentalen und multifokalen Dystonie beobachtet. Alle Ergebnisse der Gen-basierten GWAS-Analysen sind in der Ergänzungstabelle S3.1-3.9 auf der beigefügten CD-ROM enthalten.

Gen-basierte GWAS zu segmentaler und multifokaler Dystonie

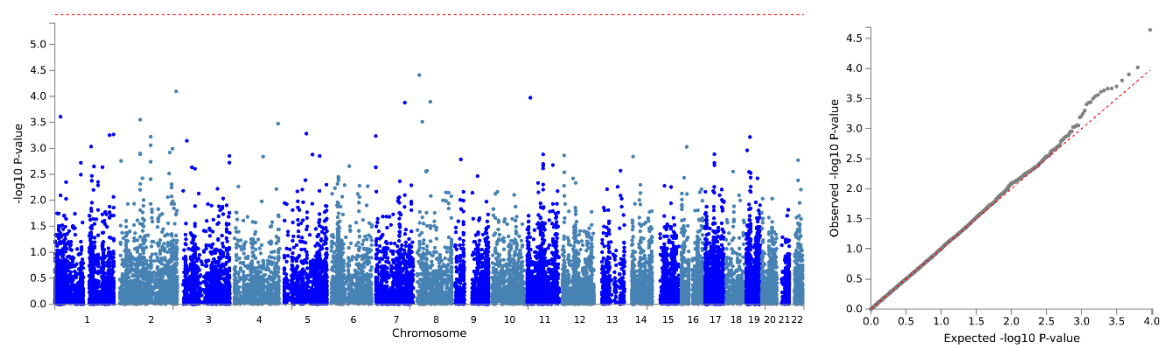
In den GWAS-Metaanalysen zu segmentaler und multifokaler Dystonie ($n = 577$ Fälle und $n = 5\,212$ Kontrollen) wurde ein Gen-basiertes Signal mit genomweiter Signifikanz beobachtet (Tabelle 3.5, Abbildung 3.7). Die Assoziation wurde durch Varianten im Tankyrase-Gen (*TNKS*; $Z_{meta} = 4,56$; $p_{meta} = 2,47 \times 10^{-6}$; Tabelle 3.5) hervorgerufen. In GTEx (v8) wird *TNKS* als ubiquitär exprimiert annotiert,

wobei die stärkste Expression im Kleinhirn vorliegt. Im GWAS-Katalog wurden Varianten dieses Gens mit einer Reihe von menschlichen Merkmalen in Verbindung gebracht, darunter mehrere, die für die Funktion des ZNS relevant sind, z. B. Hyperintensitäten der weißen Substanz und Messungen der kortikalen Oberfläche (<https://www.ebi.ac.uk/gwas/genes/TNKS>).

A



B



C

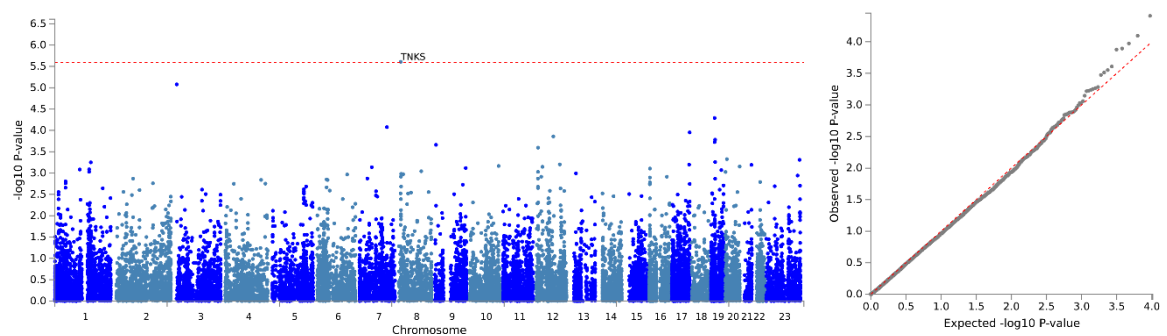


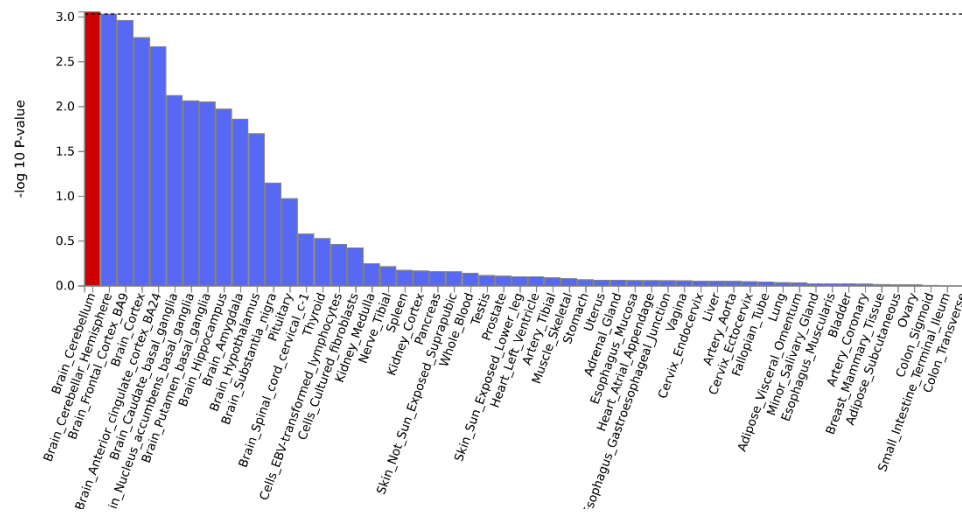
Abbildung 3.7 Manhattan- und QQ-Plots der genomweiten Assoziationsergebnisse auf Gen-Ebene mit der segmentalen und multifokalen Dystonie. A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie zeigt den Schwellenwert für Gen-basierte genomweite Signifikanz, d. h. $\alpha = 5 \times 10^{-6}$. NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (siehe Kapitel Methoden).

Schließlich deuten neuere funktionelle Studien (92, 93) auf eine Rolle von *TNKS* bei der Aggregation und Ablagerung des transaktiven Response-DNA-bindenden Proteins von 43 kDa (*TDP-43*) hin, einem führenden Faktor bei der Prädisposition für die fronto-temporale Demenz und die amyotrophe Lateralsklerose. Das zweitstärkste, aber nur suggestiv signifikante Gen-basierte Signal in dieser Dystonie-Untergruppe wurde mit dem Gen Contactin 6 (*CNTN6*; $z_{meta} = 4,30$; $p_{meta} = 8,35 \times 10^{-06}$; Tabelle 3.5) beobachtet. Wie *TNKS* zeigt auch dieses Gen laut GTEx (v8) (84) die höchste Expression im menschlichen Kleinhirn. Trotz seiner potenziellen Rolle bei der Vermittlung von Zelloberflächeninteraktionen während der Entwicklung des Nervensystems (<https://www.uniprot.org/uniprot/Q9UQ52>) zeigen bisher publizierte GWAS derzeit keine eindeutigen Zusammenhänge in Bezug auf ZNS/Gehirn-Erkrankungen oder Gehirn-Phänotypen (<https://www.ebi.ac.uk/gwas/genes/CNTN6>). Zusammenfassend lässt sich sagen, dass die GWAS-Metaanalyse unter Verwendung des Subtyps der segmentalen und multifokalen Dystonie Gen-basierte Signale mit zwei interessanten Genkandidaten aufgedeckt hat, die beide Hinweise auf eine Beteiligung an der Entwicklung und/oder Funktion des ZNS zeigen.

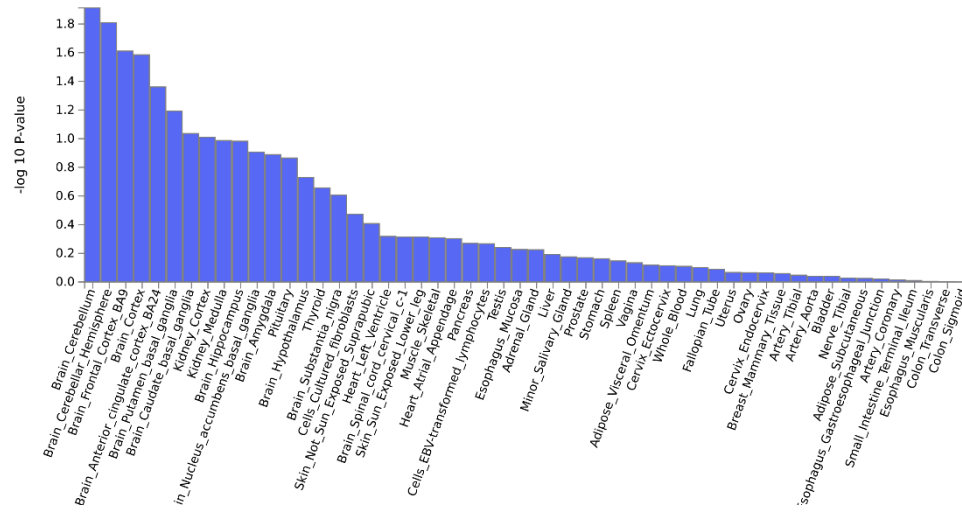
3.2.5.3 Gen-Property-Analysen

Basierend auf den Gen-basierten GWAS-Ergebnissen wurden mithilfe der FUMA-Software außerdem Gen-Property-Analysen durchgeführt. Diese Analysen zielen darauf ab, Beziehungen zwischen zelltypspezifischen Genexpressionen von 54 Gewebearten vom GTEx-Projekt (v8) (84) und Gen-basierten Assoziationen zu testen. Nach der Korrektur für multiples Testen (d. h. unter Verwendung eines Schwellenwerts von $p < 9,2 \times 10^{-04}$ nach Bonferroni-Korrektur, wie sie von FUMA empfohlen wird) erreichte der Gewebetyp Kleinhirn für die isolierte Dystonie die Signifikanzgrenze (Beta: 0,019; $p = 8,75 \times 10^{-04}$). Obwohl keine anderen der analysierten klinischen Dystonie-Untergruppen signifikante Signale mit den Gewebearten von GTEx ergaben, ist es interessant, dass die in GTEx verfügbaren Gehirnregionen fast immer oben in den Gen-Property-Analysen rangieren, also den „richtigen“ Gewebetyp als assoziiert hervorheben (ergänzende Tabellen 5.1-5.9; Abbildung 3.8). Exemplarisch werden im Hauptteil der Arbeit die Ergebnisse der Gen-Property-Analysen der drei größten GWAS-Metaanalysen zu der isolierten Dystonie, der fokalen Dystonie und der zervikalen Dystonie dargestellt (Abbildung 3.8). Zusammenfassend lässt sich sagen, dass in allen Metaanalysen (mit einer Ausnahme: Dystonie der oberen Extremitäten) ein klarer und konsistenter Trend zum Gehirngewebe zu erkennen ist. Dies deutet darauf hin, dass zumindest einige der neuen GWAS-Ergebnisse, die in dieser Studie identifiziert wurden (einschließlich derjenigen mit subgenomweiter Signifikanz), eine Anreicherung für den "richtigen" Gewebetyp aufweisen und somit in der Tat an der Pathophysiologie der Dystonie beteiligt sein könnten.

A



B



C

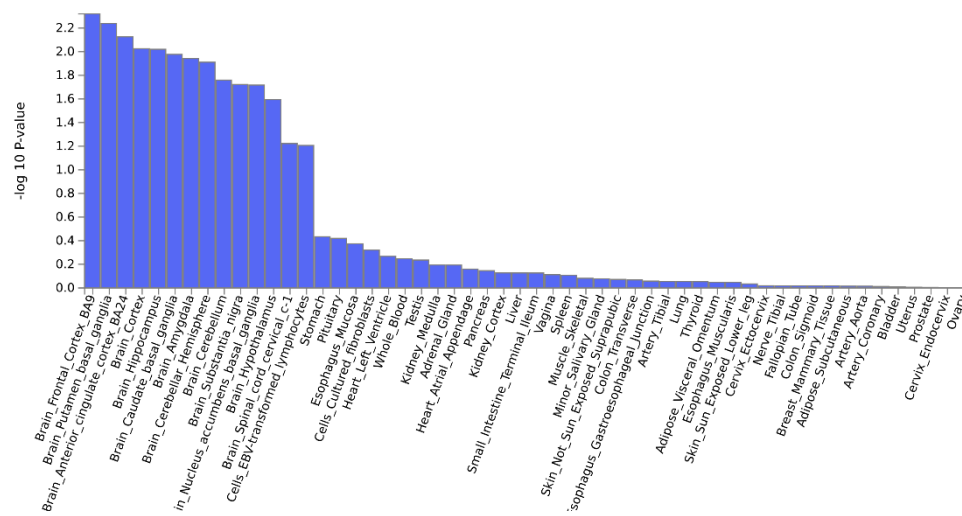


Abbildung 3.8 Gen-Property-Analysen mit 54 Gewebearten aus GTEx-Projekt (v8) (84) und Ergebnissen der Gen-basierten GWAS-Metaanalysen zu A) Isolierter Dystonie; B) Fokaler Dystonie und C) Zervikaler Dystonie. Die Plots und die eigentlichen Analysen wurden mit Hilfe der FUMA Software erstellt.

3.2.5.4 Heritabilitätsbestimmungen

Die Verfügbarkeit von SNP-Genotypisierungsdaten mit hoher Dichte in den für diese Arbeit gesammelten isolierten Dystonie-Fällen ermöglicht eine Schätzung des Anteils der phänotypischen Varianz, der durch die Genetik erklärt wird (die sog. SNP-basierte Heritabilität). Diese Schätzungen wurden mit Hilfe des GREML-Ansatzes (implementiert im GCTA-Programm; (94)) unter der Annahme eines additiven linearen Modells berechnet. Keine der drei bisher publizierten Dystonie-GWAS hat diese Methode nach meinem Kenntnisstand auf ihre Daten angewandt, so dass diese Heritabilitätsschätzungen als neuartig angesehen werden können.

Wie aus Tabelle 3.9 ersichtlich, ergaben die meisten Schätzungen der SNP-basierten Heritabilität für die Dystonieformen dieser Arbeit Werte um ~10% (Spanne: 8,25-14,76%). D.h. dass ca. 10% der phänotypischen Varianz auf den Einfluss von SNPs mit einer MAF > 10% zurückzuführen ist. Diese Schätzungen sind nicht nur gut im Einklang mit denen anderer komplex-genetischer Erkrankungen wie der Alzheimer-Krankheit ($h^2 = 0,101$; SE = 0,022) (95) und der Parkinson-Krankheit ($h^2 = 0,105$; SE = 0,017) (95), sondern auch über die hier analysierten Dystonie-Subtypen und Studienpopulationen (deutsch vs. amerikanisch) bemerkenswert konsistent. Eine Ausnahme stellen die Subtypen der Dystonie der oberen Extremitäten, laryngeale Dystonie und der aufgabenspezifischen Dystonie (Tabelle 3.9) dar, deren SNP-basierte Heritabilität als sehr viel geringer geschätzt wurden und auch statistisch nicht signifikant waren. Dies könnte auch auf der Tatsache beruhen, dass diese Berechnungen außer auf den SNP-Genotypdaten auch auf Schätzungen der Krankheitsprävalenz beruhen. Die verfügbare Literatur zu Prävalenzschätzungen der hier analysierten Dystonieformen ist jedoch vergleichsweise spärlich, weshalb ich mich für diese Arbeit auf relativ alte Studien beziehen musste: Für Deutschland wurden hauptsächlich die Daten von Castelon Konkiewitz et al. (6) verwendet, welche 2002 veröffentlicht wurden, wogegen für die USA sogar noch ältere Daten (Nutt et al. (7)) verwendet werden mussten, welche bereits 1988 veröffentlicht wurden. Außerdem fällt auf, dass die Prävalenzschätzungen für die USA höher sind als für Deutschland/Europa (siehe Tabelle 3.9), was möglicherweise auch die etwas höheren Ergebnisse der SNP-basierten Heritabilitätsschätzungen in diesem Datensatz erklärt. Diese Analysen sollten – wie alle Ergebnisse dieser Arbeit – in unabhängigen Stichproben idealerweise auf Basis aktuellerer Prävalenzzahlen wiederholt werden, bevor sie als „gesichert“ angesehen werden können.

Tabelle 3.9 Ergebnisse der SNP-basierten Heritabilitätsschätzung. Die Ergebnisse wurden mit Hilfe von GCTA (GCTA v1.92.1beta5) berechnet und basieren auf hartkodierten Genotypen ($R^2 > 0,7$, MAF > 10%). Als Kovariablen wurde das Geschlecht und die ersten zehn Hauptkomponente der PCA sowie die Krankheitsprävalenz in dem Modell berücksichtigt.

	Gruppen aus Deutschland				Gruppen aus USA			
	Prävalenz ¹	Heritabilität (SE)	p-Wert	# Fall / # Kontrolle	Prävalenz. ²	Heritabilität (SE)	p-Wert	# Fall / # Kontrolle
Dystonie								
Isolierte	142	0,0866 (0,0180)	$3,9298 \times 10^{-07}$	1205/4287	329	0,1074 (0.0345)	$8,5577 \times 10^{-04}$	1633/925
<i>Fokale</i>	101	0,0861 (0,0184)	$8,6846 \times 10^{-07}$	955/4287	295	0,1129 (0.0382)	$1,2463 \times 10^{-03}$	1157/925
Zervikale	54	0,0825 (0,0187)	$3,4549 \times 10^{-06}$	508/4287	89 ³	0,1476 (0.0538)	$2,4012 \times 10^{-03}$	309/925
Kraniale	31 ⁵	0,0929 (0,0315)	$1,4179 \times 10^{-03}$	140/4287	69 ⁴	0,1073 (0.0715)	$5,8801 \times 10^{-02}$	169/925
Oberen Extremitäten	4 ⁵	0,0003 (0,0006)	$2,4715 \times 10^{-1}$	233/4287	16 ⁵	0,0162 (0.0162)	$1,6442 \times 10^{-01}$	110/925
Laryngeale	10	n.v.	n.v.	n.v.	52	0,0637 (0.0280)	$1,0164 \times 10^{-02}$	398/925
Aufgabenspezifische	6	0,0024 (0,0014)	$4,4453 \times 10^{-02}$	190 /4287	n.v.	n.v.	n.v.	n.v.
<i>Multifokale + segmentale</i>	38	0,1067 (0,0336)	$6,2188 \times 10^{-04}$	173/4287	101 ⁵	0,0725 (0.0501)	$7,4635 \times 10^{-02}$	404/925
<i>Generalisierte</i>	3	n.v.	n.v.	n.v.	34	0,0000 (0.0752)	$5,0000 \times 10^{-01}$	67/925

n.v. = nicht vorhanden

¹ Prävalenz pro Million Personen in Deutschland (6)

² Prävalenz pro Million Personen in den USA (7)

³ Tortikollis Dystonie (7)

⁴ Oromandibuläre Dystonie (7)

⁵ Approximiert aus verfügbaren Prävalenz-Werte in der finnischen Population (88)

3.2.5.5 Korrelationsanalysen mit anderen komplex-genetischen Phänotypen

In den in diesem Abschnitt vorgestellten Analysen wurde untersucht, ob die genetischen Risikofaktoren der Dystonie mit denen anderer neurologischer Erkrankungen, psychiatrischer Erkrankungen, der kognitiven Leistungsfähigkeit, dem Lebensstil und dem Körpermasseindex (BMI) korrelieren. Die Analysen beruhen auf dem kürzlich beschriebenen Ansatz der „LD-Score-Regression“ (LDSC; (89)). Tabelle 3.10 fasst die LDSC-Ergebnisse meiner Arbeit zusammen. Als neurologische Erkrankungen wurden die Korrelationen zu neueren GWAS der Alzheimer- (96) und Parkinson-Krankheit (97) untersucht. Die Analysen ergaben nominal signifikante Korrelationen zwischen der Alzheimer-Krankheit und der isolierten Dystonie ($r = -0,33$; $p = 4,76 \times 10^{-02}$) sowie der fokalen Dystonie ($r = -0,35$; $p = 3,75 \times 10^{-02}$), nicht jedoch mit den GWAS-Ergebnissen der Parkinson-Krankheit, d.h. dass das genetische Risikoprofil zumindest teilweise zwischen der Dystonie und der Alzheimer-Krankheit korreliert. Die negativen Vorzeichen der Korrelationskoeffizienten bedeuten hierbei, dass die Allele, die das Risiko für die Dystonieformen erhöhen, protektiv auf das Risiko der Alzheimer-Krankheit wirken und umgekehrt.

Als Beispiele aus dem Feld der psychiatrischen Erkrankungen wurden Korrelationen der drei Dystonieformen dieser Arbeit mit den fünf größten GWAS aus der Literatur berechnet, d.h. für die Autismus-Spektrum-Störung (98), bipolare Störung (99), schwere depressive Störung (100), Zwangsneurose (101) und Schizophrenie (99). Diese Analysen ergaben signifikante Korrelationen zwischen der isolierten Dystonie und der Zwangsstörung ($r = 0,44$; $p = 1,39 \times 10^{-03}$), sowie zwischen der isolierten Dystonie und der Schizophrenie ($r = 0,26$; $p = 1,28 \times 10^{-04}$) und der fokalen Dystonie und der bipolaren Störung ($r = 0,17$; $p = 3,69 \times 10^{-02}$). Des Weiteren ergaben sich signifikante Korrelationen zwischen der fokalen Dystonie und der Zwangsstörung ($r = 0,49$; $p = 1,00 \times 10^{-03}$) sowie der Schizophrenie ($r = 0,25$; $p = 3,29 \times 10^{-04}$). Von allen analysierten psychiatrischen Erkrankungen zeigte die zervikale Dystonie nur eine genetische Korrelation mit der Genetik der Schizophrenie ($r = 0,18$; $p = 3,20 \times 10^{-03}$). Allen diesen Ergebnissen ist gemeinsam, dass der Korrelationskoeffizient – anders als bei den analogen Analysen zur Alzheimer-Krankheit (s.o.) – positiv ist. Das bedeutet, dass die genetischen Risikofaktoren der Dystonie auch das Risiko der hier beschriebenen psychiatrischen Erkrankungen erhöhen.

Außer den oben beschriebenen Analysen zu neurologisch-psychiatrischen Erkrankungen habe ich auch die Überlappung mit den GWAS-Ergebnissen zur kognitiven Leistungsfähigkeit (102), BMI (103) und zu verschiedenen Lebensstilfaktoren (erhoben als Teil der UK-BioBank-Studie (104) [<http://www.nealelab.is/uk-biobank/>] sowie dem Zigarettenkonsum (105)) untersucht. Nominal signifikante Werte ergeben sich hierbei für die genetische Korrelation zwischen der zervikalen Dystonie und dem Kaffeekonsum ($r = -0,14$; $p = 4,06 \times 10^{-02}$) sowie dem Zigarettenkonsum ($r = -$

0,1034; $p = 4,92 \times 10^{-02}$). Dies könnte darauf hindeuten, dass die genetischen Faktoren, die „Suchtverhalten“ beeinflussen sich protektiv (wegen des negativen Vorzeichens des Korrelationskoeffizienten) auf das Risiko der Dystonie auswirken. Bei dem Vergleich mit dem BMI ergab sich für die isolierte Dystonie ($r = -0,22$; $p = 4,18 \times 10^{-06}$) und die zervikale Dystonie ($r = -0,12$; $p = 3,90 \times 10^{-03}$) ebenfalls negative, nominal signifikante Korrelationen. Dies entspräche einem protektiven Effekt auf die Dystonie von denselben genetischen Faktoren, die sonst eine Körpergewichterhöhung (gemessen als BMI) bewirken. Keine signifikanten Korrelationen wurden zwischen den hier analysierten drei Dystonie-Phänotypen und der kognitiven Leistungsfähigkeit beobachtet.

Insgesamt wurden in diesen Analysen 36 Vergleiche durchgeführt (Tabelle 3.10), so dass die p-Werte zur Interpretation entsprechend für multiples Testen korrigiert werden sollten, dies ergibt einen $\alpha = 0,05/36 = 1,39 \times 10^{-03}$. Dieser Wert, der aufgrund der teilweisen Korrelation der Dystonieformen untereinander wahrscheinlich zu konservativ ist, wird nur von den Korrelationen zwischen der Dystonie und der Zwangsneurose, der Schizophrenie und dem BMI unterschritten. Insbesondere hervorzuheben ist hierbei die negative Korrelation zwischen der isolierten Dystonie mit dem BMI, die mit einem p-Wert von $4,18 \times 10^{-06}$ das mit Abstand ausgeprägteste Signifikanzniveau aller Korrelationsanalysen zeigte.

Zusammenfassend lässt sich feststellen, dass die drei größten in meiner Arbeit zur Verfügung stehenden Dystonie-GWAS-Datensätze einige moderat signifikante Korrelationen mit GWAS-Daten der ausgewählten anderen Phänotypen und Krankheiten zeigte, die auf eine zumindest partiell geteilte genetische Architektur zwischen diesen Phänotypen hinweisen könnte. Einschränkend ist bemerkt, dass nur drei dieser Korrelationen auch nach Korrektur für multiples Testen noch signifikant waren. Interessanterweise wurde in keiner der Analysen eine signifikante Korrelation mit der Genetik der Parkinson-Erkrankung gefunden, was darauf hindeutet, dass diese beiden Erkrankungen trotz einiger klinischer Gemeinsamkeiten keine (bzw. keine starke) Überlappung in ihrer genetischen Architektur aufweisen.

Tabelle 3.10 Ergebnisse der genetischen Korrelationsanalysen mittels LDSC für die isolierte Dystonie, fokale Dystonie und zervikale Dystonie.

Neurologische Krankheiten	Aktuellste GWAS			Isolierte Dystonie		Fokale Dystonie		Zervikale Dystonie	
	PMID	Jahr	# N	r	p	r	p	r	p
Alzheimer-Krankheit	30617256	2019	444006	-0,3292	4,7571 x 10 ⁻⁰²	-0,3547	3,7545 x 10 ⁻⁰²	-0,1741	2,0080 x 10 ⁻⁰¹
Parkinson-Krankheit	31701892	2019	482730	0,1782	6,9467 x 10 ⁻⁰²	0,1683	9,4806 x 10 ⁻⁰²	0,1721	6,8400e x 10 ⁻⁰²
Psychiatrische Erkrankungen									
Autismus-Spektrum-Störung	30804558	2019	46350	-0,0021	9,8285 x 10 ⁻⁰¹	-0,0173	8,5859 x 10 ⁻⁰¹	-0,0631	4,5220 x 10 ⁻⁰¹
Bipolare Störung	29906448	2018	41653	0,1508	7,3470 x 10 ⁻⁰²	0,1715	3,6900 x 10 ⁻⁰²	0,0952	1,8790 x 10 ⁻⁰¹
Schwere depressive Störung	29700475	2018	173005	0,1138	1,7054 x 10 ⁻⁰¹	0,0631	4,3111 x 10 ⁻⁰¹	-0,0297	6,8640 x 10 ⁻⁰¹
Zwangsneurose	28761083	2018	9725	0,4444	1,3915 x 10 ⁻⁰³	0,4911	1,0001 x 10⁻⁰³	0,2383	6,2700 x 10 ⁻⁰²
Schizophrenie	29906448	2018	65967	0,2673	1,2831 x 10⁻⁰⁴	0,2503	3,2969 x 10⁻⁰⁴	0,1820	3,2000 x 10 ⁻⁰³
Kognitive Phänotyp									
Kognitive Leistungsfähigkeit	30038396	2018	257828	0,0512	4,7066 x 10 ⁻⁰¹	-0,0060	9,0972 x 10 ⁻⁰¹	-0,0933	5,9400 x 10 ⁻⁰²
Lebensstil									
Kaffeekonsum	UKBB*	2018	334659	-0,1082	1,4421 x 10 ⁻⁰¹	-0,0786	2,8816 x 10 ⁻⁰¹	-0,1360	4,0600 x 10 ⁻⁰²
Ehemalige Alkoholabhängigkeit	UKBB*	2018	6514	-0,1688	3,2158 x 10 ⁻⁰¹	-0,2601	1,5806 x 10 ⁻⁰¹	-0,0531	7,5150 x 10 ⁻⁰¹
Ehemaliger Zigarettenkonsum	30643258	2019	518633	-0,0945	9,5645 x 10 ⁻⁰²	-0,1088	5,2022 x 10 ⁻⁰²	-0,1034	4,9200 x 10 ⁻⁰²
Zusätzliche Phänotyp									
Körpermasseindex (BMI)	30124842	2018	681275	-0,2255	4,1770 x 10⁻⁰⁶	-0,2553	8,6271 x 10⁻⁰⁷	-0,1205	3,900 x 10 ⁻⁰³

*<http://www.nealelab.is/uk-biobank/>

3.3 Vergleich der neuen GWAS-Ergebnisse mit denen der Field-Synopsis

Die Field-Synopsis zur systematischen Erfassung genetischer Assoziationsstudien in der Dystonie bis 2018 (25) (siehe auch Abschnitt 3.1.) ergab sieben SNPs (rs4532, rs13283584, rs35153737, rs7342975, rs9972951, rs1801968, rs11655081), die eine mindestens nominal signifikante Assoziation mit verschiedenen Dystonie-Typen zeigten. Diese wurden in diesem Teil der Arbeit noch einmal explizit mit den Ergebnissen der von mir durchgeführten GWAS-Analysen verglichen; alle Ergebnisse dieses Vergleichs sind in Tabelle 3.11 dargestellt. Zusammenfassend lässt sich feststellen, dass die Ergebnisse der Field-Synopsis nur sehr eingeschränkt mit denen der GWAS korrespondierten.

Fünf der sieben o.g. SNPs aus der Field-Synopsis zeigten eine nominal signifikante Assoziation in den Metaanalysen über alle Datensätze mit isolierter Dystonie hinweg (Tabelle 3.13). Zwei dieser SNPs, rs4532 (im Gen *DRD1*) und rs9972951 (im Gen *ARSG*), zeigten als einzige in den korrespondierenden GWAS-Analysen mindestens nominal signifikante p-Werte ($p = 0,016$ und $p = 0,046$, respektive) mit denselben Effekt-Richtungen wie in der Field-Synopsis. SNP rs9972951 war signifikant in der GWAS-Metaanalyse, die alle isolierten Dystonie-Datensätze kombinierte, während rs4532 nur in Patienten mit fokaler Dystonie signifikant war. Zwei weitere SNPs der Field-Synopsis (rs13283584 [im Gen *TOR1A*] und rs7342975 [im Gen *ARSG*]) konnten dahingegen in den GWAS keiner der analysierten Stichproben bestätigt werden (kleinste p-Werte: $p = 0,778$ und $p = 0,473$), obwohl bei der *ARSG*-Variante zumindest die Effektrichtungen mit denen der Field-Synopsis korrespondierten. Für SNP rs35153737 (im Gen *TOR1A*) lagen keine GWAS-Genotypen vor, weil diese SNP wurde nicht anhand des HRC-Referenzpanels (67) imputiert, so dass die Daten für diesen SNP nicht verglichen werden konnten.

Die signifikantesten Assoziationen der Field-Synopsis waren rs11655081, rs7342975 und rs9972951, die allesamt im Kontext einer GWAS (39) zur Musikerdystonie identifiziert wurden. Hier konnte diesbezüglich nur die allgemeine Gruppe von Patienten, die unter einer aufgabenspezifischen Dystonie leiden, auf Assoziation getestet werden. Obwohl die Effektrichtung für alle drei Varianten in dieselbe Richtung wie in der Vorläufer-GWAS zeigten, war keines der Resultate statistisch signifikant (Tabelle 3.11). Keiner der oben beschriebenen nominal signifikanten Assoziationen übersteht eine Korrektur für multiples Testen: der entsprechend Bonferroni-korrigierte $\alpha = 0,05/96 = 5,2^{-04}$.

Ein Ergebnisvergleich in umgekehrter Richtung, d.h. Ermittlung der Assoziationsevidenz der neuen GWAS-Befunde hier in den Daten der Field-Synopsis war nicht möglich, da keine der im Kontext meiner GWAS-Analysen identifizierten Varianten vorher getestet wurden. Im Umkehrschluss bedeutet dies, dass die Gene, in deren Nähe die hier identifizierten neuen GWAS-Signale lokalisiert

sind, keine besonderen „Kandidatengene“ der Dystonie darstellten. Letztlich unterstreicht dies die Stärke des GWAS-Ansatzes, der – zumindest was die funktionelle Fragestellung zum Phänotyp betrifft – hypothesenfrei nach Gen-Phänotyp-Verbindungen sucht.

Nach Publikation der Field-Synopsis haben Sun et al. im Juli 2021 (41) eine GWAS zur zervikalen Dystonie veröffentlicht. Obwohl diese nicht Teil meiner Synopsis war, möchte ich abschließend hier kurz auf die Ergebnisse dieser neuesten Studie eingehen. Hauptergebnisse aus Sun et al., 2021 wurden mit SNP rs2219975 in *COL8A1* sowie in Gen-basierten Analysen mit Markern in *DENND1A* berichtet. Zusätzlich war SNP rs147331823 in *GABBR2* mit dem Alter zum Beginn der Krankheit assoziiert (siehe Tabelle A 2 im Anhang). Interessanterweise zeigten weder diese SNPs noch das *DENND1A*-Gen einen Hinweis auf eine signifikante Assoziation in den GWAS-Metaanalysen zur zervikalen Dystonie (rs2219975: $p = 4,73 \times 10^{-1}$, rs147331823: $p = 5,45 \times 10^{-1}$, *DENND1A*: $p = 5,03 \times 10^{-1}$). Auch in den beiden Subgruppen aus Deutschland (rs2219975: $p = 8,01 \times 10^{-1}$, rs147331823: $p = 6,88 \times 10^{-1}$) oder den USA (rs2219975: $p = 6,60 \times 10^{-2}$, rs147331823: $p = 8,22 \times 10^{-2}$) zeigten diese SNPs keine signifikanten Ergebnisse. Das Fehlen eines Assoziationssignals mit den Varianten der neuesten publizierten GWAS zur zervikalen Dystonie ist überraschend, weil die Stichprobengröße – anders als bei den Metaanalysen der Field-Synopsis – mit $n = 919$ Fällen mit der hier verwendeten Anzahl an Fällen der zervikalen Dystonie vergleichbar war. Allerdings war zumindest in der deutschen Stichprobe die Anzahl der verwendeten Kontrollen wesentlich höher als bei Sun et al., 2021, was die Diskrepanz zumindest zum Teil erklären könnte. Unsere Nicht-Replikation der Hauptbefunde aus Sun et al., 2021 deckt sich allerdings mit ihren eigenen Ergebnissen: Sun et al. konnten ihre GWAS-Befunde in einer unabhängigen Stichprobe von $n=473$ Fällen und 481 Kontrollen ebenfalls nicht unabhängig bestätigen (41).

Tabelle 3.11 Vergleich der Ergebnisse der Dystonie Metaanalysen aus der Field-Synopsis mit den Ergebnissen der GWAS, welche in dieser Arbeit erstellt wurde.

SNP	Chr	Pos	A1	A2	Gen	Metaanalyse aus der Literatur (25)		GWAS der deutschen Datensätze		GWAS der amerikanischen Datensätze		GWAS-Metaanalyse	
						OR	p	OR	p	OR	p	OR	p
A. Isolierte Dystonie													
rs4532	5	174870150	C	T	DRD1	1,37	1,5 × 10⁻⁰³	1,11	4,0 × 10 ⁻⁰²	1,09	9,0 × 10 ⁻⁰²	1,10	1,6 × 10⁻⁰²
rs13283584	9	132574862	T	C	TOR1A	1,32	5,8 × 10⁻⁰³	1,05	4,0 × 10 ⁻⁰¹	0,96	5,9 × 10 ⁻⁰¹	1,01	7,8 × 10 ⁻⁰¹
rs35153737	9	132575836	-	C	TOR1A	1,46	3,1 × 10⁻⁰³	n.v.	n.v.	n.v.	n.v.	n.v.	n.v.
rs7342975	17	66391232	G	A	ARSG	2,00	5,7 × 10⁻⁰³	1,16	5,1 × 10 ⁻⁰¹	1,09	7,3 × 10 ⁻⁰¹	1,13	4,7 × 10 ⁻⁰¹
rs9972951	17	66391276	A	G	ARSG	2,16	3,4 × 10⁻⁰²	1,18	3,0 × 10 ⁻⁰¹	1,41	7,0 × 10 ⁻⁰²	1,28	4,6 × 10⁻⁰²
B. Fokale Dystonie													
rs4532	5	174870150	C	T	DRD1	1,37	1,5 × 10⁻⁰³	1,13	3,0 × 10 ⁻⁰²	1,09	2,3e × 10 ⁻⁰¹	1,11	1,4 × 10 ⁻⁰²
rs7342975	17	66391232	G	A	ARSG	2,00	5,7 × 10⁻⁰³	1,36	2,0 × 10 ⁻⁰¹	1,09	7,6 × 10 ⁻⁰¹	1,24	2,4 × 10 ⁻⁰¹
rs9972951	17	66391276	A	G	ARSG	2,16	3,4 × 10⁻⁰³	1,20	2,9 × 10 ⁻⁰¹	1,01	9,6 × 10 ⁻⁰¹	1,12	4,1 × 10 ⁻⁰¹
C. Schreibkrampf (in GWAS: Aufgabenspezifische Dystonie)													
rs1801968	9	132580901	G	C	TOR1A	3,10	1,4 × 10⁻⁰²	0,88	4,9 × 10 ⁻⁰¹	n.v.	n.v.	n.v.	n.v.
D. Musikerdystonie (in GWAS: Aufgabenspezifische Dystonie)													
rs11655081	17	66382209	C	T	ARSG	4,42	2 × 10⁻⁰⁹	1,59	6,0 × 10 ⁻⁰²	n.v.	n.v.	n.v.	n.v.
rs7342975	17	66391232	G	A	ARSG	1,32	1,4 × 10⁻⁰²	1,85	1,2 × 10 ⁻⁰¹	n.v.	n.v.	n.v.	n.v.
rs9972951	17	66391276	A	G	ARSG	2,16	4,2 × 10⁻⁰²	1,40	2,7 × 10 ⁻⁰¹	n.v.	n.v.	n.v.	n.v.

n.v. = nicht vorhanden

4 Diskussion

Die Dystonie ist eine klinisch sowie genetisch komplexe Erkrankung, die isoliert oder in der Kombination mit anderen Erkrankungen vorkommen kann. In dieser Arbeit geht es um die Identifizierung und Charakterisierung von genetischen Faktoren, die zur Krankheitsentstehung der isolierten Dystonie beitragen. Trotz einer recht umfangreichen Literatur zur Genetik der monogenen isolierten Dystonie war zum Zeitpunkt dieser Dissertation vergleichsweise wenig bekannt zur Genetik der isolierten Dystonie (106). In der vorliegenden Arbeit habe ich zur Beantwortung dieser Fragestellung zwei Ansätze verfolgt: 1. Durchführung einer Field-Synopsis zur systematischen Aufarbeitung der publizierten genetischen Datenlage zur isolierten Dystonie und 2. Durchführung verschiedener GWAS an bisher für diese Zweck noch nicht analysierten Datensätzen von über 8.000 Individuen, mit dem Ziel neue Dystonie-Risikogene zu identifizieren. Für die Field-Synopsis habe ich eine systematische Literaturrecherche zu genetischen Assoziationsstudien bei der isolierten Dystonie und bei verschiedenen anderen klinische Dystonie-Gruppen durchgeführt und anschließend die Daten der so erhobenen genetischen Varianten metaanalysiert. Für die GWAS habe ich verschiedene unabhängige Datensätze, die dem Teilprojekt P6 der DFG-Forschungsgruppe „ProtectMove“ (FOR2488) von deutschen und amerikanischen Stichproben für diesen Zweck zur Verfügung standen durchgeführt und anschließend ebenfalls metaanalysiert. Diese GWAS-Analysen schlossen auch das X-Chromosom ein, das aufgrund methodischer Besonderheiten in der GWAS-Literatur typischerweise bisher nur wenig Beachtung findet. Die Ergebnisse der GWAS-Analysen bildeten die Grundlage für weitere sog. „Post-GWAS“-Analysen, mit dem Ziel die gewonnenen Erkenntnisse auf funktionell-molekularer, klinischer und epidemiologischer Sicht einzuordnen. Im 3. und letzten Teil dieser Arbeit wurden schließlich die Ergebnisse der Field-Synopsis mit den neu generierten Ergebnissen der GWAS verglichen. Zusammenfassend stellt diese Arbeit eine der derzeit umfangreichsten Analysen zur Genetik der Dystonie weltweit dar.

4.1 Field-Synopsis

4.1.1 Zusammenfassung

Nach umfangreicher Literaturrecherche und sorgfältiger Prüfung wurden 42 unabhängige Publikationen zu genetischen Assoziationsstudien identifiziert. Hierin waren Daten von 45 Polymorphismen in mindestens zwei unabhängigen Datensätzen vorhanden, die nachfolgend metaanalysiert wurden. In der Metaanalyse zeigten 5 Polymorphismen einen Hinweis auf eine nominal signifikante Assoziation mit dem Risiko, eine Form der isolierten Dystonie zu entwickeln. Von diesen blieb nur die DNA-Variante rs4532 in *DRD1* nach Bonferroni-Korrektur signifikant. Zusätzlich wurden die identifizierten Datensätze in weitere klinisch-diagnostische Dystonie-

Gruppen und, wenn genügend Daten vorhanden waren, auch nach ethnischer Abstammung unterteilt und nachfolgend ebenfalls metaanalysiert. Die DNA-Variante rs4532 zeigte hierbei auch innerhalb der Gruppe der fokalen Dystonien eine signifikante Assoziation. Des Weiteren zeigte die DNA-Variante rs11655081 im *ARSG*-Gen nach Bonferroni-Korrektur eine Assoziation mit der Musikerdystonie. Vor unserer Field-Synopsis waren lediglich vier Metaanalysen von genetischen Dystonie-Daten veröffentlicht (35–38), eine systematisch-quantitative Übersicht über die in der Dystonie untersuchten genetischen Polymorphismen existierte nach meinem Kenntnisstand nicht. Dieser Teil der Dissertation wurde in der Dezember 2018 Ausgabe der Fachzeitschrift *Parkinsonism and Related Disorders* als Research Artikel unter der Rubrik „Original Research“ veröffentlicht (25). In der Januar 2019 Ausgabe derselben Zeitschrift wurde als Kommentar zu unserer Field-Synopsis ein Editorial veröffentlicht (106), der u. a. die Wichtigkeit derartiger systematischer Ansätze lobend hervorhob.

4.1.2 Einordnung

Obwohl die Ergebnisse der Field-Synopsis durch die Zusammenführung aller bis dato veröffentlichten genetischen Assoziationsdaten *qua Design* auf den größtmöglichen Datensätzen basierten ist einschränkend anzumerken, dass die Stichprobengrößen der Metaanalysen insgesamt trotzdem vergleichsweise klein waren. Z. B. umfassten die Daten der Metaanalysen zu *DRD1* nur Daten aus drei unabhängigen Stichproben (39, 107, 108) mit insgesamt 315 erkrankten und 1.172 gesunden Probanden. Bei *ARSG* umfassten die metaanalysierten Daten mit 243 erkrankten und 1.097 gesunden Probanden etwa gleich viele Stichproben, hier allerdings nur aus zwei unabhängigen Datensätzen, die zudem alle derselben Publikation entstammten (39). Der Median des Stichprobenumfangs über alle Metaanalysen hinweg betrug nur 2.882 Personen (25), verglichen mit 4.500 für Alzheimer (47) und 7.500 für Parkinson (48) aus vorher von unserer Arbeitsgruppe publizierten ähnlichen Field-Synopsen. Eine in der Publikation durchgeführte formale Abschätzung der statistischen Trennschärfe („Power“) ergab, dass diese für die allermeisten untersuchten Varianten kleiner als 80% und somit sehr gering war. Lediglich für ORs ≥ 1.5 erreichte unsere Studie eine Power von $>80\%$. Nur Assoziationsstudien mit Stichproben deutlich >3000 würden genügend Power haben, um SNPs mit einer geringeren Effektstärke zu erkennen (48). Dasselbe gilt für Analysen in Bezug auf einen möglichen Publikationsbias und die Identifizierungen anderer Verzerrungen.

Andere Einschränkung der Field-Synopsis beziehen sich auf die Auswahl der Suchparameter, durch die einige in Frage kommende Publikationen möglicherweise nicht identifiziert wurden. Auch besteht die Möglichkeit eines fehlerhaften Datenextraktions- und Darstellungsverfahrens. Solche Fehler sind bei Verfolgung systematischer Ansätze allerdings mehr oder weniger unvermeidlich.

Durch die große Erfahrung unserer Arbeitsgruppe bei der Durchführung systematischer Field-Synopsen (47, 48, 109–111) wird die Gefahr dieser Probleme allerdings minimiert und sollte daher keinen nennenswerten Einfluss auf die Hauptschlussfolgerungen der Arbeit gehabt haben.

Vor unserer Studie waren nur für zwei Gene, *TOR1A* (35, 36) und *BDNF* (37, 38), Metaanalysen über genetische Assoziationsdaten bei Dystonien verfügbar. Diese Publikationen berichten aber widersprüchliche Ergebnisse. Die Studie von Siokas et al., 2017 (36) ergab z. B. eine signifikante Auswirkung der *TOR1A*-Varianten rs1182 auf die Entwicklung der fokalen Dystonie. Des Weiteren ergab die *TOR1A*-Variante rs1801968 eine signifikante Assoziation mit dem Schreibkrampf. Die Studie von Groen et al., 2013 (35), hingegen kam zu einem gegenteiligen Ergebnis. In unserer Field-Synopsis ergaben keine der beiden *TOR1A*-Varianten eine signifikante Assoziation mit der fokalen Dystonie (siehe Tabelle 3.2 A, reproduziert aus Ohlei et al, 2018 (25)). Obwohl die Ergebnisse meiner Metaanalysen auf einen nominalen Effekt der zweiten *TOR1A*-Variante rs1801968 beim Schreibkrampf hindeuteten (siehe Tabelle 3.2 F, reproduziert aus Ohlei et al, 2018 (25)), war dieses Ergebnis nach Bonferroni-Korrektur nicht mehr signifikant. Bei dem *BDNF* Gen wurde in mehreren Quellen der Literatur nur ein einziger Polymorphismus rs6265 untersucht. Diese DNA-Variante zeigte keinen Hinweis auf eine Assoziation zur Dystonie in der Metaanalyse von Gómez-Garre et al., 2014 (37) , obwohl eine gewisse Unterstützung für die Assoziation mit der Dystonie in der Publikation von Sako et al., 2015 (36) berichtet wird. In meiner Arbeit wurde dahingegen kein Assoziationsnachweis der *BDNF*-Variante rs6265 mit dem Risiko für die Dystonie festgestellt (25).

4.2 GWAS der isolierten Dystonie und klinischen Subgruppen

4.2.1 Zusammenfassung

Eine Konklusion der Field-Synopsis war, dass zusätzliche, großangelegte Assoziationsstudien notwendig sind, um die Genetik der isolierten Dystonie und ihrer Subformen weiter zu entschlüsseln. Die Methode der Wahl ist heute die Durchführung von genomweiten Assoziationsanalysen z. B. als GWAS, die auch in dieser Arbeit verwendet wurden. Die Daten der Dystonie-Patienten und deren DNA-Proben wurden im Kontext des Teilprojekts P6 der DFG-Forschungsgruppe „ProtectMove“ gesammelt und umfassten insgesamt die Proben von 2.838 Patienten und 5.212 Kontrollprobanden aus Deutschland und den USA. Obwohl bei der kombinierten Analyse aller Patienten mit der isolierten Dystonie keine genomweit signifikanten ($p < 5 \times 10^{-8}$) Signale gefunden wurden, traten einige interessante Assoziationen bei den GWAS-Analysen der Dystonie-Subgruppen zu Tage. Bei den Patienten mit kranialer Dystonie wurden z. B. 12 genomweit signifikante SNPs identifiziert, die über vier Genloci verteilt sind. Das stärkste Assoziationssignal mit einem p-Wert von $5,52 \times 10^{-09}$ wies die Variante rs143591361 innerhalb eines nicht-kodierenden Bereichs auf Chromosom 6q12 zwischen den beiden lncRNAs (*RP11-542F9.1* und

RP11-406O16.1) auf. In dieser Region wurden weitere genomweit signifikante Signale für acht SNPs nachgewiesen, die in starkem LD mit dem *lead*-SNP standen. Ein weiteres genomweit signifikantes Signal wurde mit SNP rs3094685 innerhalb der *HLA*-Region auf Chromosom 6q21.33 beobachtet. Zwei weitere hochsignifikante Signale, mit rs12022195 und rs76357386 haben keine weiteren signifikanten Varianten in Kopplungsgleichgewicht in der unmittelbaren Nähe und müssen daher kritisch bewertet werden (s.u.). Die einzige andere Dystonie-Subgruppe mit genomweit signifikanten Assoziationsbefunden war die Dystonie der oberen Extremitäten, wo zwei SNPs (rs4081356 auf Chromosom 1p31.3 und rs182982101 auf Chromosom 3q12.3) diese Schwelle unterschritten. Beide Polymorphismen liegen in der Nähe von lncRNA-Sequenzen (d. h. *RP11-393N21.1* und *RP11-221J22.1*; Tabelle 3.5). Die meisten in diesen Analysen hervorgehobenen Top-SNPs liegen in oder in der Nähe von nicht-kodierenden RNA-Genen (vorwiegend lncRNAs), was darauf hindeutet, dass viele der assoziierten SNPs ihre Wirkung durch Beeinflussung der Expression von bisher weitgehend unbekanntem Zielgenen entfalten können.

Interessanterweise zeigten die signifikanten Signale der kranialen Dystonie und der Dystonie der oberen Extremitäten keine Assoziation mit anderen klinischen Dystonie-Gruppen, was auf Unterschiede in den pathophysiologischen Mechanismen zwischen den klinischen Dystonie-Gruppen hindeuten könnte. Das X-Chromosom wurde aufgrund methodischer Besonderheiten separat zu den Autosomen analysiert. Obwohl dabei verschiedene X-Chromosom-weit signifikante Assoziationen identifiziert wurden, erreichte keines dieser Signale den Grenzwert für eine genomweit signifikante Assoziation.

Bei den im Zuge der Post-GWAS-Analysen durchgeführten Gen-basierten GWAS-Berechnungen wurden ein weiteres genomweit signifikantes Signal ($p < 2,7 \times 10^{-06}$) mit dem Gen *TNKS* (Chromosom 8; $p = 2,22 \times 10^{-07}$) in der Subgruppe der segmentalen und multifokalen Dystonie festgestellt. Weitere Post-GWAS-Analysen zeigten z. B. dass Gene, die mit dem Dystonie-Risiko assoziiert sind, auch im Bereich des Gehirns, v.a. des Zerebellums (Kleinhirn) stärker exprimiert werden und dass die phänotypische Varianz durch die hier analysierten SNPs zu einem Anteil von ca. 8-12 % erklärt wird. Dieser Wert zeigte eine bemerkenswerte Konstanz über die meisten der analysierten Subgruppen und Studienpopulationen (deutsch vs. amerikanisch). In Bezug zu den anderen hier analysierten komplex-genetischen Erkrankungen gab es insgesamt nur wenige Überlappungen bezogen auf die genetische Architektur. Die stärkste Korrelation wurde bei der fokalen Dystonie mit der Zwangsneurose ($r = 0,49$, $p = 1,0 \times 10^{-03}$), die statistisch signifikanteste mit BMI ($r = -0,23$, $p = 4 \times 10^{-06}$) beobachtet. Interessanterweise habe ich keine signifikante genetische Korrelation mit GWAS-Befunden der Parkinson-Krankheit, einer anderen, häufigeren Bewegungsstörung, gefunden, obwohl diese einige klinische Gemeinsamkeiten mit der Dystonie aufweist. Ein

Manuskript zur Veröffentlichung der in dieser Arbeit beschriebenen GWAS-Befunde befindet sich derzeit in Vorbereitung.

4.2.2 Einordnung

Obwohl in der Literatur wurden bereits drei Dystonie-GWAS veröffentlicht wurden (eine zur Musikerdystonie (39) und zwei zur zervikalen Dystonie (40, 41); letztere wurde erst im Juli 2021, also nach Abschluss der Analysen dieser Dissertation publiziert) gab es zu Beginn der hier durchgeführten Analysen keine weiteren GWAS zu anderen Formen der isolierten Dystonie. D.h. für diese Dystonieformen stellt die vorliegende Arbeit nach meinem Kenntnisstand die erste GWAS weltweit dar. Weitere Stärken bestehen in der für eine seltene Erkrankung vergleichsweise großen Stichprobe ($n = 2.838$ Fälle und 5.212 Kontrollen kombiniert), die im Feld erstmalige Berechnung von Assoziationsergebnissen auf dem X-Chromosom (XWAS), sowie die Durchführung vieler weiterführender Analysen, die zumindest partiell die Interpretation der erhobenen Ergebnisse erleichtern.

Trotz dieser Stärken müssen auch einige Limitationen dieser Arbeit in Betracht gezogen werden, zuerst die der Stichprobengröße. Obwohl ich in dieser Arbeit die weltweit größten Dystonie-Datensätze, die je einer GWAS unterzogen wurden, analysieren konnte, sind die verfügbaren Datensätze insbesondere für die Dystonie-Subgruppen teilweise immer noch vergleichsweise klein. Im Kontext von GWAS z. B. der Parkinson-Krankheit, d.h. der häufigsten Bewegungsstörung des Menschen, sind die Fallzahlen (und damit auch die statistische Power) mit über >50.000 Fällen und >1 Millionen Kontrollprobanden (97) derzeit wesentlich höher. Dies ist zu einem der mit $\sim 1\%$ deutlich höheren Prävalenz des Parkinsons im Vergleich zur Dystonie ($\sim 0.01-0.03\%$) geschuldet, was die Kollektion von erkrankten Probanden wesentlich erleichtert. In Tabelle 4.1 habe ich die statistische Trennschärfe für verschiedene Kombinationen von Effektstärken (OR) und Allelfrequenzen (MAF) für die hier durchgeführten GWAS-Metaanalysen aufgeführt. Dabei bin ich von einer Prävalenz der Dystonie von $0,03\%$ bei der deutschen und amerikanischen Bevölkerung ausgegangen. Die Power-Schätzungen zeigen, dass die Trennschärfe zur Detektion genomweit signifikanter ($p < 5 \times 10^{-08}$) oder genomweit suggestiv signifikanter ($p < 1 \times 10^{-06}$) Varianten mit ORs $< 1,15$ über alle MAF-Bereiche hinweg nur gering ($< 30\%$) war. Das heißt im Umkehrschluss, dass nur Analysen mit wesentlich größeren Stichproben ($n > 40.000$) auch genügend „Power“ haben, um SNPs mit derart geringen Effektstärken (OR: $1,15$; MAF = 10%) zu identifizieren.

Tabelle 4.1 Power-Schätzungen für die in diesem Bereich durchgeführten isolierten Dystonie Metaanalysen. Die Stichprobe umfasst 2.838 Dystonie Patienten und 5.212 gesunde Probanden. Die Schätzwerte wurden mit dem Power Calculator (112) für eine Reihe von OR-, MAF- und Typ-I-Fehlerkombinationen berechnet, um eine 80% Power zu erzielen. Alle Schätzungen gehen von einer Krankheitsprävalenz von 0.03 % und von additiven genetischen Effekten aus. Die grau schattierten Zellen zeigen die OR-MAF-Kombinationen mit mindestens 80% Power für Assoziationen.

Allelic OR	Power für $\alpha = 5 \times 10^{-08}$ ($\alpha = 1 \times 10^{-06}$)				
	MAF = 0,01	MAF = 0,05	MAF = 0,1	MAF = 0,25	MAF = 0,5
1,15	0 (0)	0 (0)	0 (0,01)	0,03 (0,10)	0,07 (0,18)
1,25	0 (0)	0,01 (0,03)	0,09 (0,21)	0,58 (0,78)	0,74 (0,88)
1,50	0 (0,01)	0,55 (0,76)	0,98 (0,99)	1 (1)	1 (1)
2,00	0,21 (0,41)	1 (1)	1 (1)	1 (1)	1 (1)
2,50	0,82 (0,93)	1 (1)	1 (1)	1 (1)	1 (1)

Aufgrund des Vorliegens von nur drei vorherigen GWAS aus dem Dystonie-Feld ist ein konkreter Ergebnisvergleich mit Vorläuferstudien nur begrenzt möglich. In der bereits beschriebenen GWAS zur Musikerdystonie wurde eine genomweit signifikante Assoziation mit SNP rs11655081 im *ARSG* Gen berichtet. In den Ergebnissen meiner Arbeit war dieser und andere SNPs in der *ARSG* Region allerdings nicht signifikant mit der aufgabenspezifischen Dystonie assoziiert (siehe Tabelle A 2 im Anhang). Mögliche Gründe dafür sind die Unterschiede bei der Dystonie-Gruppen Klassifizierung (hier: allgemein aufgabenspezifisch; vormals: Musikerdystonie) und eine ungenügende statistische Trennschärfe aufgrund der eher kleinen Stichprobengröße in dieser Arbeit. Allerdings war die Stichprobengröße in der Vorläufer-GWAS mit 243 Fällen und 1.097 Kontrollen noch kleiner, so dass auch das Vorliegen eines falsch-positiven Befundes in der GWAS von Lohmann et al. 2014 (39) nicht ausgeschlossen werden kann. Dies würde auch die dort berichtete und mit 4,4 für komplex-genetische Erkrankungen ungewöhnlich hohe OR erklären, die auf eine verzerrte GWAS-Teststatistik hinweisen könnte. Letztlich kann hier nur die Analyse weiterer unabhängiger Stichproben Klarheit schaffen.

Die zweite Dystonie-GWAS (40) hat sich mit der Analyse der zervikalen Dystonie befasst und beschreibt mögliche Assoziationen mit SNPs in der Nähe des *NALCN*-Gens. Allerdings erfüllten die Ergebnisse dieser GWAS nach derzeit gängigen Standards nicht die Kriterien der genomweit signifikanten Assoziation, da das Top-Signal aus Mok et al., 2014 (40) mit einem p-Wert von $9,76 \times 10^{-07}$ mehr als eine Größenordnung von diesem Grenzwert entfernt lag (Tabelle A 2 im Anhang). Aufgrund der viel größeren Fallzahl der Dystonie Patienten hier ($n = 817$) im Vergleich zu Mok et al., 2014 ($n = 212$) (40) ist die Wahrscheinlichkeit eines falsch-positiven Befunds in der Vorläuferstudie hoch.

Die dritte Dystonie-GWAS (41) hat sich ebenfalls mit der Analyse der zervikalen Dystonie GWAS befasst. Dabei wurden genomweit signifikante Assoziation mit dem SNP rs2219975 im *COL8A1*-Gen

berichtet. In den Ergebnissen meiner Arbeit war dieser SNP in der *COL8A1*-Region allerdings nicht signifikant mit der zervikalen Dystonie assoziiert. Darüber hinaus zeigte keines der in Sun et al., 2021 (41) beschriebenen Top Ergebnisse in den hier vorgelegten Daten eine nennenswerte Assoziation (Tabelle A 2 im Anhang). Aufgrund der viel größeren Fallzahl hier ($n = 6.029$) vs. Sun et al., 2021 ($n = 2.410$) (41), ist auch hier ein falsch-positiver Befund in der Vorläuferstudie nicht ausgeschlossen, bzw. sogar wahrscheinlich. Hierzu passt, dass die Autoren von Sun et al., 2021 (41) ihren genomweiten Befund mit rs2219975 selbst in einer unabhängigen Stichprobe von 473 Fällen mit zervikaler Dystonie und 481 Kontrollen nicht replizieren konnten (41).

Auch bei einigen in meiner Arbeit als genomweit signifikant identifizierten Ergebnissen kann ein falsch-positiver Befund nicht ausgeschlossen werden. Zu nennen wäre hier v. a. SNP rs182982101 der mit einer OR von 7,48 und $p = 2,95 \times 10^{-8}$ eine sehr starke Assoziation mit der Dystonie der oberen Extremitäten zeigte. Allerdings handelt es sich bei rs182982101 um eine seltene DNA-Variante mit einer MAF von 0,7% in der „Normalbevölkerung“ nicht finnischer, nord-west-europäischer Abstammung (113). Abgesehen davon, dass die Teststatistiken bei derart seltenen Varianten naturgemäß anfällig sind für Ausreißereffekte, können DNA-Varianten mit geringen Allelenhäufigkeiten ($<5\%$) auch nur mit geringer Genauigkeit imputiert werden (114). Um diesem Effekt entgegenzuwirken, wurden hier nur die SNPs analysiert, die einem hohen Imputations-Qualitäts-Score ($r^2 = 0,7$) aufwiesen. Dennoch kann ein falsch positives Ergebnis nicht ausgeschlossen werden und eine unabhängige Replikation dieses und der anderen Assoziationsergebnisse ist notwendig, um das hier vorgestellte Ergebnis sicherer interpretieren zu können. Eine weitere mögliche Limitation dieser Studie ist bedingt durch die laborexperimentelle Prozessierung. Aus logistischen Gründen wurden die DNA-Proben der ProtectMove Gesamtstichprobe (zu der auch DNA-Samples der PD-Fälle und Kontrollen sowie Samples mit XDP gehörten) in unterschiedlichen Chargen in das Genotypisierungslabor geschickt und dort auch in unterschiedlichen Labor-Batches prozessiert. Hierbei konnten nicht immer Fälle und Kontrollen in demselben molekulargenetischen Experiment prozessiert werden, so dass die Gefahr von Batch-Effekten besteht. Die Wahrscheinlichkeit solcher Verzerrungen wurde allerdings auf analytischer Seite minimiert durch die Verwendung immer desselben Genotypisierungsarrays und die gemeinsame Prozessierung aller Proben nach einem standardisierten, semi-automatisierten Verfahren. Des Weiteren wurden die Genotyp-Cluster-Plots sämtlicher genomweit assoziierter SNPs *post-hoc* noch einmal visuell bewertet, um Unterschiede im Clustering über die Labor-Batches hinweg auszuschließen. Trotz dieser Vorsichtsmaßnahmen kann nicht ausgeschlossen werden, dass einige der hier berichteten Ergebnisse evtl. falsch-positive Befunde darstellen.

Weitere mögliche Limitationen betreffen die Post-GWAS-Analysen, z. B. die Schätzungen der SNP-basierten Heritabilität, die stark von der verwendeten Prävalenz abhängig ist. Leider gibt es in der einschlägigen Literatur nur sehr wenige verlässliche neuere Prävalenzdaten zur Dystonie, insbesondere bzgl. der hier analysierten klinischen Subtypen. Gerade die Prävalenzschätzungen in der amerikanischen Bevölkerung sind z. T. sehr alt (siehe Tabelle 3.9 in Abschnitt 3.2.5.4), neuere Daten waren aber nach meinem Kenntnisstand zur Schriftlegung dieser Arbeit nicht verfügbar. Von diesen Ungenauigkeiten können daher auch die Schätzungen der SNP-basierten Heritabilität betroffen sein. Gegen eine größere Verzerrung der Daten spricht allerdings, dass sämtliche Heritabilitätschätzungen in den Bereich zwischen 8% und 12% fielen und damit, trotz Unterschieden in der Krankheitsprävalenz, z. B. zwischen der deutschen und amerikanischen Population, sehr konsistent ausfielen (siehe Tabelle 3.9). Schließlich können auch die Berechnungen der genetischen Korrelationen der Dystonie mit anderen komplex-genetischen Phänotypen, die hier mithilfe der LDSC-Methode geschätzt wurden, bestimmten Verzerrungen unterliegen, weil die Genauigkeit dieser Methode v.a. von den Stichprobengrößen der zu vergleichenden GWAS abhängig ist. Zur Minimierung dieses Effekts wurden zum einen hier nur Phänotypen berücksichtigt, für die GWAS-Ergebnisse von größeren Datensätzen vorlagen (alle $n > 6.500$, die meisten deutlich darüber, s. Tabelle 3.11). Auf der Dystonie-Seite habe ich für die LDSC-Analysen darüber hinaus nur die drei größten klinischen hier verfügbaren Gruppen ($n > 6.029$ s. Tabelle 3.11) berücksichtigt. Trotzdem kann nicht ausgeschlossen werden, dass ein Teil der LDSC-Ergebnisse durch die begrenzte Stichprobengröße verfälscht ist, d.h. die Korrelationen unter- oder überschätzt wurden.

4.3 Korrespondenz der Ergebnisse aus beiden Teilen der Arbeit

Die Field-Synopsis ergab Hinweise auf signifikante Signale zwischen dem Dystonie-Risiko und Varianten in *DRD1* und *ARSG* (25). Obwohl es eine gewisse Übereinstimmung, insbesondere in Bezug auf die Effektrichtung, mit dem Signal von SNP rs4532 in *DRD1* in den GWAS-Analysen dieser Arbeit zur isolierten (OR: 1,10; $p = 0,0164$) und fokalen Dystonie (OR: 1,11; $p = 0,0149$) Dystonie gab, sind diese Assoziationsbefunde sehr weit von der Schwelle der genomweiten Signifikanz entfernt. Dies mag zum einen daran liegen, dass unsere Studie für kleine Effektstärken nicht ausreichend statistische Power besaß (s.o.), zum anderen aber auch daran, dass die Rolle von *DRD1* in der Risikoarchitektur der Dystonie nicht so groß ist wie ursprünglich beschrieben. Eine ähnliche Konklusion lässt sich für den zweiten signifikanten SNP der Field-Synopsis, d.h. rs11655081 in *ARSG*, ziehen: Dieser SNP zeigte in der Subgruppe der aufgabenspezifischen Dystonie (die zu 66% aus Patienten mit Musikerdystonie besteht) eine lediglich nominal signifikante Assoziation (OR: 1,59; $p = 6,0 \times 10^{-02}$) deren Effektstärke deutlich unter der der Vorläuferpublikation lag (OR: 4,42; $p = 2,0 \times 10^{-9}$). Mögliche Gründe für diese Diskrepanz sind die Unterschiede in der Klassifizierung der

Dystonie-Gruppen und/oder ein falsch-positives Ergebnis der Vorläuferstudie (39). Keine anderen der nominal signifikanten Assoziationen der Field-Synopsis erreichten nennenswerte Resultate in den hier durchgeführten GWAS-Analysen. Für diese sowie alle weiteren genetischen Assoziationsergebnisse dieser Arbeit gilt, dass weitere unabhängige Stichproben getestet werden müssen, um weiterführende Konklusionen treffen zu können.

5 Ausblick

In diesem letzten Kapitel möchte ich noch kurz mögliche weiterführende Analysen und Nachfolgeuntersuchungen skizzieren, die als mögliche „nächste Schritte“ nachfolgend zu den Analysen dieser Dissertation angesehen werden können, deren Durchführung aber den Rahmen dieser Arbeit überschritten hätten. Außer den hier angedachten Analyseprojekten gibt es allerdings noch eine Vielzahl weiterer Möglichkeiten die im Kontext dieser Arbeit generierten Daten zu verwenden, mit dem Ziel die genetische Basis der Dystonie noch besser zu verstehen und damit evtl. die individuelle Vorhersage der Krankheitsprädisposition zu ermöglichen.

5.1 Genomweite Analyse weiterer Phänotypen der Dystonie

Das Ziel dieser Arbeit war es, die genetischen Faktoren (SNPs) zu identifizieren, die einen direkten Einfluss auf die Krankheitsentwicklung der Dystonie haben, d.h. die Beschreibung von Risikovarianten. In weiteren Analysen könnten auch genetische Faktoren identifiziert werden, welche andere Aspekte der Krankheitsexpressivität beeinflussen, d.h. die Durchführung von GWAS-Analysen, in denen z. B. bestimmte quantitative Dystonie-relevante Phänotypen (anstelle des hier verwendeten Fall-Kontroll-Status) als abhängige Variablen Verwendung finden. Hierfür könnten z. B. das Alter zu Beginn der Erkrankung, der Schweregrad der Anfangssymptome und die Geschwindigkeit des Verlaufs (d.h. Krankheitsprogression) verwendet werden. Von diesen genannten Phänotypen war bei Durchführung der Analysen dieser Arbeit nur das Alter zu Beginn der Erkrankung bekannt. Für die eigentlichen GWAS könnten die bereits imputierten und auf ihre Qualität hin kontrollierten SNPs dieser Arbeit als unabhängige Variablen übernommen werden. Ein Vorteil dieses aufgrund der quantitativen Natur der abhängigen Variablen auch als „quantitative trait locus“ (QTL)-Analysen bezeichneten Vorgehens ist, dass die quantitativen Variablen aufgrund ihrer größeren Varianz meist eine erhöhte statistische Power nach sich ziehen und dass es aufgrund der Nichtbeachtung der gesunden Kontrollgruppen keine Anfälligkeit für Batch-Effekte gibt.

5.2 Genomweite Analyse in Mutationsträgern

Neben der standardmäßig vom Hersteller (Illumina, Inc.) auf dem GSA platzierten genetischen Varianten wurde im Kontext der ProtectMove-Studie noch weitere Varianten inkludiert (auf der sog. „shared custom content“ Partition des GSA), die größtenteils bereits bekannte Mutationen der untersuchten Bewegungsstörungen darstellen. In zukünftigen Analysen könnten die Dystonie-Patienten und Kontrollen so auf das Vorhandensein einer der auf dem Custom-Content eingeschlossenen Dystonie-Mutationen hin analysiert werden. Falls hierdurch eine genügend große Anzahl an Probanden (v.a. Kontrollprobanden) als Träger solcher Mutationen identifiziert werden,

könnten z. B. zukünftige GWAS-Analysen gezielt nach genetischen Faktoren der reduzierten Penetranz suchen, also nach Faktoren, die Mutationsträger davor „schützen“ die Krankheit zu entwickeln. Auch hierfür könnten die bereits qualitätskontrollierten und imputierten SNPs aus dieser Arbeit übernommen werden und mit den in Abschnitt 2.2.3.3. beschriebenen Methoden berechnet werden.

5.3 Genomweite Analysen anderer genetischer Varianten

Typische GWAS basieren meist auf SNP-Assoziationen und sind daher in Bezug auf ihre genetische Auflösung limitiert, weil es außer SNPs im humanen Genom auch noch andere genetische Varianten gibt, wie z. B. copy number variants (CNVs) oder short tandem repeats (STRs), die ebenfalls zur individuellen Prädisposition für eine Krankheit neben den SNPs beitragen können. Anders als bei den beiden oben erwähnten weiterführenden Analysen müssten hierfür die Genotypisierungsrohdaten vor den entsprechenden Analysen allerdings noch einmal komplett neu prozessiert und qualitätskontrolliert werden, z. B. mittels spezifischer Referenzpanels. Dafür könnten die Phänotypen als abhängige Variable direkt aus dieser Arbeit übernommen werden.

6 Zusammenfassung

Die Dystonie hat neben erworbenen Auslösern auch genetisch vererbare Ursachen, die monogener und polygener Natur sein können. Obwohl in der Literatur mehr als zwei Dutzend genetische Loci in Verbindung mit monogenen Formen der Dystonie beschrieben wurden, ist nur wenig bekannt über die Genetik der polygenen Form der Erkrankung. Um das Verständnis der genetischen Architektur dieser Dystonie-Form zu erweitern, wurden im Rahmen dieser Arbeit die folgenden Strategien verfolgt: 1. die Erstellung einer Field-Synopsis, inklusive systematischer Metaanalysen, von publizierten genetischen Assoziationsstudien, 2. die Durchführung von GWAS-Analysen bei mehr als 8.000 DNA-Proben aus Deutschland und den USA, und 3. der Vergleich der neuen GWAS-Ergebnisse mit den Befunden der Field-Synopsis.

Für die Field-Synopsis identifizierte ich 3.575 Artikel, von denen 42 die Einschlusskriterien erfüllten. SNPs, für die mindestens zwei unabhängige Datensätze vorlagen, wurden metaanalysiert. Insgesamt habe ich so 134 Metaanalysen zu 45 DNA-Varianten in 17 Genloci durchgeführt. Hierbei zeigten mehrere Varianten nominal signifikante Assoziationssignale, obwohl nur zwei Varianten in den Genen *DRD1* und *ARSG* signifikant nach Korrektur für multiples Testen blieben.

Im zweiten Teil der Arbeit habe ich mittels des GWAS-Ansatzes nach neuen genetischen Risikofaktoren der isolierten Dystonie und ihrer klinischen Unterformen gesucht. Dafür wurden über 8.000 DNA-Proben von deutschen und amerikanischen Fällen und Kontrollpersonen genotypisiert. In sechs Metaanalysen über beide Datensätze hinweg konnte ich insgesamt sechs neue genomweit signifikante Risikoloci mit der kranialen Dystonie und der Dystonie der oberen Extremitäten identifizieren. Bei den Gen-basierten GWAS-Analysen trat ein weiteres genomweit signifikantes Signal zwischen Markern in *TNKS* und der segmentalen und multifokalen Dystonie zu Tage. Post-GWAS-Analysen zeigten u. a., dass die SNP-basierte Heritabilität der meisten Dystonie-Subtypen bei ca. 10% lag und dass es nur moderat signifikante Überlappungen mit den GWAS-Befunden anderer komplex-genetischer Phänotypen gab. Es zeigten nur sehr wenige SNPs aus den Metaanalysen der Field-Synopsis in den neuen GWAS-Daten eine nominale Signifikanz.

Zusammengefasst stellt diese Arbeit aufgrund ihrer Stichprobengröße und der Verwendung des GWAS-Ansatzes die derzeit größte Abhandlung zur Architektur der komplex-genetischen Form der isolierten Dystonie dar. Obwohl alle hier identifizierten potentiellen Dystonie-Risikoloci konsistente Ergebnisse in den Subgruppen der deutschen und amerikanischen Probanden zeigten, bedürfen die hier beschriebenen Ergebnisse weiterer Replikation in unabhängigen Datensätzen, bevor sie als gesichert angesehen werden können.

Literaturverzeichnis

1. Ceballos-Baumann A, Conrad B. 9 Dystonien. In: Ceballos-Baumann A, Conrad B, editors. *Bewegungsstörungen*. 2nd ed. Stuttgart: Georg Thieme Verlag; 2005. Available from: <https://doi.org/10.1055/b-0033-621>
2. Albanese A, Bhatia K, Bressman SB, Delong MR, Fahn S, Fung VSCC, et al. Phenomenology and classification of dystonia: A consensus update. *Mov Disord*. 2013 Jun 15;28(7):863–73. Available from: <https://doi.org/10.1002/mds.25475>
3. Fahn S. Classification of movement disorders. *Mov Disord*. 2011 May 1;26(6):947–57. Available from: <https://doi.org/10.1002/mds.23759>
4. Klein C, Lohmann K, Marras C, Münchau A. Hereditary Dystonia Overview. *GeneReviews®*. University of Washington, Seattle; 1993. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20301334>
5. Lohmann K, Klein C. Genetics of dystonia: What's known? What's new? What's next? *Mov Disord*. 2013 Jun 15;28(7):899–905. Available from: <https://doi.org/10.1002/mds.25536>
6. Castelon Konkiewitz E, Trender-Gerhard I, Kamm C, Warner T, Ben-Shlomo Y, Gasser T, et al. Service-based survey of dystonia in Munich. *Neuroepidemiology*. 2002;21(4):202–6. Available from: <https://doi.org/10.1159/000059525>
7. Nutt JG, Muentner MD, Aronson A, Kurland LT, Melton LJ. Epidemiology of focal and generalized dystonia in Rochester, Minnesota. *Mov Disord*. 1988;3(3):188–94. Available from: <https://doi.org/10.1002/mds.870030302>
8. Rondot P, Marchand MP, Dellatolas G. Spasmodic Torticollis — Review of 220 Patients. *Can J Neurol Sci / J Can des Sci Neurol*. 1991;18(2):143–51. Available from: <https://doi.org/10.1017/S0317167100031619>
9. Chan J, Brin MF, Fahn S. Idiopathic cervical dystonia: Clinical characteristics. *Mov Disord*. 1991;6(2):119–26. Available from: <https://doi.org/10.1002/mds.870060206>
10. Jinnah HA, Berardelli A, Comella C, Defazio G, Delong MR, Factor S, et al. The focal dystonias: Current views and challenges for future research. Vol. 28, *Movement Disorders*. *Mov Disord*; 2013. p. 926–43. Available from: <https://doi.org/10.1002/mds.25567>
11. Ludlow CL, Adler CH, Berke GS, Bielamowicz SA, Blitzer A, Bressman SB, et al. Research priorities in spasmodic dysphonia. Vol. 139, *Otolaryngology - Head and Neck Surgery*.

- Mosby Inc.; 2008. p. 495. Available from: <https://doi.org/10.1016/j.otohns.2008.05.624>
12. Zwirner P, Dressler D, Kruse E. Spasmodic laryngeal dyspnea: A rare manifestation of laryngeal dystonia. *Eur Arch Oto-Rhino-Laryngology*. 1997;254(5):242–5. Available from: <https://doi.org/10.1007/BF00874096>
 13. Chitkara A, Meyer T, Keidar A, Blitzer A. Sieger’s dystonia: First report of a variant of spasmodic dysphonia. Vol. 115, *Annals of Otolaryngology, Rhinology and Laryngology*. Annals Publishing Company; 2006. p. 89–92. Available from: <https://doi.org/10.1177/000348940611500201>
 14. Karp BI. Botulinum toxin physiology in focal hand and cranial dystonia [Internet]. Vol. 4, *Toxins*. Multidisciplinary Digital Publishing Institute (MDPI); 2012. p. 1404–14. Available from: <https://doi.org/10.3390/toxins4111404>
 15. Ceballos-Baumann A. Idiopathische Dystonien. Springer, Berlin, Heidelberg; 2018 p. 1–17. Available from: https://doi.org/10.1007/978-3-662-44768-0_132-1
 16. Liuzzi D, Gigante AF, Leo A, Defazio G. The anatomical basis of upper limb dystonia: lesson from secondary cases. Vol. 37, *Neurological Sciences*. Springer-Verlag Italia s.r.l.; 2016. p. 1393–8. Available from: <https://doi.org/10.1007/s10072-016-2598-6>
 17. Stahl CM, Frucht SJ. Focal task specific dystonia: a review and update. *J Neurol*. 2017 Jul 1;264(7):1536–41. Available from: <https://doi.org/10.1007/s00415-016-8373-z>
 18. Sadnicka A, Kassavetis P, Pareés I, Meppelink AM, Butler K, Edwards M. Task-specific dystonia: Pathophysiology and management. Vol. 87, *Journal of Neurology, Neurosurgery and Psychiatry*. BMJ Publishing Group; 2016. p. 968–74. Available from: <http://dx.doi.org/10.1136/jnnp-2015-311298>
 19. Shamim EA, Chu J, Scheider LH, Savitt J, Jinnah HA, Hallett M. Extreme task specificity in writer’s cramp. *Mov Disord*. 2011 Sep;26(11):2107–9. Available from: <https://doi.org/10.1002/mds.23827>
 20. García-Ruiz PJ, Val J del, Losada M, Campos JM. Task-specific dystonia of the lower limb in a flamenco dancer. Vol. 17, *Parkinsonism and Related Disorders*. Elsevier; 2011. p. 221–2. Available from: <https://doi.org/10.1016/j.parkreldis.2010.12.001>
 21. Müller U, Kupke KG. The genetics of primary torsion dystonia. *Hum Genet*. 1990 Jan; 84(2):107–15. Available from: <https://doi.org/10.1007/BF00208922>
 22. Jinnah HA, Sun Y V. Dystonia genes and their biological pathways. Vol. 129, *Neurobiology*

- of Disease. Academic Press Inc.; 2019. p. 159–68. Available from:
<https://doi.org/10.1016/j.nbd.2019.05.014>
23. Marras C, Lang A, van de Warrenburg BP, Sue CM, Tabrizi SJ, Bertram L, et al. Nomenclature of genetic movement disorders: Recommendations of the international Parkinson and movement disorder society task force. *Movement Disorders* Apr, 2016 p. 436–57. Available from: <https://doi.org/10.1002/mds.26527>
 24. Lohmann K, Klein C. Update on the Genetics of Dystonia. *Current Neurology and Neuroscience Reports* Mar. Available from: <https://doi.org/10.1007/s11910-017-0735-0>
 25. Ohlei O, Dobricic V, Lohmann K, Klein C, Lill CM, Bertram L. Field synopsis and systematic meta-analyses of genetic association studies in isolated dystonia. *Parkinsonism Relat Disord*. 2018 Dec 1;57:50–7. Available from:
<https://doi.org/10.1016/j.parkreldis.2018.07.0184>
 26. Schaaf CP, Zschocke J. *Basiswissen Humangenetik*. Berlin, Heidelberg: Springer Berlin Heidelberg;. (Springer-Lehrbuch). Available from: <http://link.springer.com/10.1007/978-3-642-28907-1>
 27. Klein C. Genetics in dystonia. 2014 Jan 1;20(SUPPL.1):S137–42. Available from:
[https://doi.org/10.1016/S1353-8020\(13\)70033-6](https://doi.org/10.1016/S1353-8020(13)70033-6)
 28. Ozelius LJ, Hewett JW, Page CE, Bressman SB, Kramer PL, Shalish C, et al. The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. *Nat Genet*. 1997;17(1):40–8. Available from: <https://doi.org/10.1038/ng0997-40>
 29. Djarmati A, Schneider SA, Lohmann K, Winkler S, Pawlack H, Hagenah J, et al. Mutations in THAP1 (DYT6) and generalised dystonia with prominent spasmodic dysphonia: a genetic screening study. *Lancet Neurol*. 2009 May;8(5):447–52. Available from:
[https://doi.org/10.1016/S1474-4422\(09\)70083-3](https://doi.org/10.1016/S1474-4422(09)70083-3)
 30. Fuchs T, Saunders-Pullman R, Masuho I, Luciano MS, Raymond D, Factor S, et al. Mutations in GNAL cause primary torsion dystonia. *Nat Genet*. 2013 Jan;45(1):88–92. Available from:
<https://doi.org/10.1038/ng.2496>
 31. Charlesworth G, Plagnol V, Holmström KM, Bras J, Sheerin UM, Preza E, et al. Mutations in ANO3 cause dominant craniocervical dystonia: Ion channel implicated in pathogenesis. *Am J Hum Genet*. 2012 Dec 7;91(6):1041–50. Available from:
<https://doi.org/10.1016/j.ajhg.2012.10.024>

32. Olschewski L, Jesús S, Kim HJ, Tunc S, Löns S, Junker J, et al. Role of ANO3 mutations in dystonia: A large-scale mutational screening study. *Park Relat Disord*. 2019 May 1; 62:196–200. Available from: <https://doi.org/10.1016/j.parkreldis.2018.12.030>
33. Lill CM, Mashychev A, Hartmann C, Lohmann K, Marras C, Lang AE, et al. Launching the movement disorders society genetic mutation database (MDSGene). Vol. 31, *Movement Disorders*. John Wiley and Sons Inc.; 2016. p. 607–9. Available from: <https://doi.org/10.1002/mds.26651>
34. O'Connor LJ, Schoech AP, Hormozdiari F, Gazal S, Patterson N, Price AL. Extreme Polygenicity of Complex Traits Is Explained by Negative Selection. *Am J Hum Genet*. 2019 Sep 5;105(3):456–76. Available from: <https://doi.org/10.1016/j.ajhg.2019.07.003>
35. Groen JL, Ritz K, Tanck MW, van de Warrenburg BP, van Hilten JJ, Aramideh M, et al. Is TOR1A a risk factor in adult-onset primary torsion dystonia? *Mov Disord*. 2013 Jun;28(6):827–31. Available from: <https://doi.org/10.1002/mds.25381>
36. Siokas V, Dardiotis E, Tsironi EE, Tsigoulis G, Rikos D, Sokratous M, et al. The role of tor1a polymorphisms in dystonia: A systematic review and meta-analysis. Gonzalez-Alegre P, editor. *PLoS One*. 2017 Jan 12;12(1):e0169934. Available from: <https://doi.org/10.1371/journal.pone.0169934>
37. Gómez-Garre P, Huertas-Fernández I, Cáceres-Redondo MT, Alonso-Canovas A, Bernal-Bernal I, Blanco-Ollero A, et al. BDNF Val66Met polymorphism in primary adult-onset dystonia: A case-control study and meta-analysis. *Mov Disord*. 2014 Jul; 29(8):1083–6. Available from: <https://doi.org/10.1002/mds.25938>
38. Sako W, Murakami N, Izumi Y, Kaji R. Val66Met polymorphism of brain-derived neurotrophic factor is associated with idiopathic dystonia. *J Clin Neurosci*. 2015 Mar;22(3):575–7. Available from: <https://doi.org/10.1016/j.jocn.2014.08.014>
39. Lohmann K, Schmidt A, Schillert A, Winkler S, Albanese A, Baas F, et al. Genome-wide association study in musician's dystonia: A risk variant at the arylsulfatase G locus? *Mov Disord*. 2014 Jun;29(7):921–7. Available from: <http://doi.wiley.com/10.1002/mds.25791>
40. Mok KY, Schneider SA, Tratzuni D, Stamelou M, Edwards M, Kasperaviciute D, et al. Genomewide association study in cervical dystonia demonstrates possible association with sodium leak channel. *Mov Disord*. 2014 Feb;29(2):245–51. Available from: <http://doi.wiley.com/10.1002/mds.25732>

41. Sun YV, Li C, Hui Q, Huang Y, Barbano R, Rodriguez R, et al. A Multi-center Genome-wide Association Study of Cervical Dystonia. *Mov Disord*. 2021 Jul 28; Available from: <https://doi.org/10.1002/mds.28732>
42. Risch NJ. Searching for genetic determinants in the new millennium. Vol. 405, *Nature*. Nature Publishing Group; 2000. p. 847–56. Available from: <https://doi.org/10.1038/35015718>
43. Bush WS, Moore JH. Chapter 11: Genome-Wide Association Studies. *PLoS Comput Biol*. 2012 Dec;8(12). Available from: <https://doi.org/10.1371/journal.pcbi.1002822>
44. Deng Y, Tsao BP. Genetics of Human SLE. In: Dubois' *Lupus Erythematosus and Related Syndromes: Eighth Edition*. Elsevier Inc.; 2013. p. 35–45.
45. Kumar KR, Davis RL, Tchan MC, Wali GM, Mahant N, Ng K, et al. Whole genome sequencing for the genetic diagnosis of heterogenous dystonia phenotypes. *Park Relat Disord*. 2019 Dec 1;69:111–8. Available from: <https://doi.org/10.1016/j.parkreldis.2019.11.004>
46. Gómez-Garre P, Huertas-Fernández I, Cáceres-Redondo MT, Alonso-Canovas A, Bernal-Bernal I, Blanco-Ollero A, et al. Lack of validation of variants associated with cervical dystonia risk: A GWAS replication study. *Mov Disord*. 2014;29(14):1825–8. Available from: <https://doi.org/10.1002/mds.26044>
47. L. Bertram, M.B. McQueen, K. Mullin, D. Blacker, R.E. Tanzi, Systematic metaanalyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat. Genet*. 2007;39 17–23. Available from: <https://doi.org/10.1038/ng1934>
48. Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, Schjeide B-MM, et al. Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. Myers AJ, editor. *PLoS Genet*. 2012 Mar 15 ;8(3):e1002548. Available from: <https://doi.org/10.1371/journal.pgen.1002548>
49. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009 Jul;6(7). Available from: <https://doi.org/10.1371/journal.pmed.1000097>
50. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007 Sep;81(3):559–75. Available from: <https://doi.org/10.1086/519795>
51. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK:

- rising to the challenge of larger and richer datasets. *Gigascience*. 2015 Dec 25;4(1):7. Available from: <https://doi.org/10.1186/s13742-015-0047-8>
52. Meyer E, Carss KJ, Rankin J, Nichols JME, Grozeva D, Joseph AP, et al. Mutations in the histone methyltransferase gene *KMT2B* cause complex early-onset dystonia. *Nat Genet*. 2017 Feb 19 ;49(2):223–37. Available from: <https://doi.org/10.1038/ng.3740>
 53. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. *Nature*. 2015 Sep 30;526(7571):68–74. Available from: <https://doi.org/10.1038/nature15393>
 54. Clarke GM, Anderson CA, Pettersson FH, Cardon LR, Morris AP, Zondervan KT. Basic statistical analysis in genetic case-control studies. *Nat Protoc*. 2011 Feb ;6(2):121–33. Available from: <https://doi.org/10.1038/nprot.2010.182>
 55. Lohmann K, Uflacker N, Erogullari A, Lohnau T, Winkler S, Dendorfer A, et al. Identification and functional analysis of novel *THAP1* mutations. *Eur J Hum Genet*. 2012 Feb 17 ;20(2):171–5. Available from: <https://doi.org/10.1038/ejhg.2011.159>
 56. Newman JRBB, Sutherland GT, Boyle RS, Limberg N, Blum S, O’Sullivan JD, et al. Common polymorphisms in dystonia-linked genes and susceptibility to the sporadic primary dystonias. *Park Relat Disord*. 2012 May;18(4):351–7. Available from: <https://doi.org/10.1016/j.parkreldis.2011.11.024>
 57. Newman JRBB, Todorovic M, Silburn PA, Sutherland GT, Mellick GD. Lack of reproducibility in re-evaluating associations between *GCH1* polymorphisms and Parkinson’s disease and isolated dystonia in an Australian case-control group. *Park Relat Disord*. 2014 Jun;20(6):668–70. Available from: <https://doi.org/10.1016/j.parkreldis.2014.02.014>
 58. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986 Sep 1 ;7(3):177–88. Available from: [https://doi.org/10.1016/0197-2456\(86\)90046-2](https://doi.org/10.1016/0197-2456(86)90046-2)
 59. Goeman JJ, Solari A. Multiple hypothesis testing in genomics. *Stat Med [Internet]*. 2014 May 20 [cited 2020 May 12];33(11):1946–78. Available from: <http://doi.wiley.com/10.1002/sim.6082>
 60. Slatkin M. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat Rev Genet*. 2008 Jun;9(6):477–85. Available from: <https://doi.org/10.1038/nrg2361>
 61. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to

- detect publication bias in meta-analysis. *JAMA*. 2006 Feb 8;295(6):676–80. Available from: <http://doi.org/doi:10.1001/jama.295.6.676>
62. Mollenhauer B, Trautmann E, Sixel-Döring F, Wicke T, Ebentheuer J, Schaumburg M, et al. Nonmotor and diagnostic findings in subjects with de novo Parkinson disease of the DeNoPa cohort. *Neurology*. 2013 Oct 1;81(14):1226–34. Available from: <https://doi.org/10.1212/WNL.0b013e3182a6cbd5>
 63. Krawczak M, Nikolaus S, von Eberstein H, Croucher PJP, El Mokhtari NE, Schreiber S. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet*. 2006 Mar;9(1):55–61. Available from: <https://doi.org/10.1159/000090694>
 64. Kasten M, Hagenah J, Graf J, Lorwin A, Vollstedt E-J, Peters E, et al. Cohort Profile: a population-based cohort to study non-motor symptoms in parkinsonism (EPIPARK). *Int J Epidemiol*. 2013 Feb;42(1):128-128k. Available from: <https://doi.org/10.1093/ije/dys202>
 65. Wang A, Costello S, Cockburn M, Zhang X, Bronstein J, Ritz B. Parkinson’s disease risk from ambient exposure to pesticides. *Eur J Epidemiol*. 2011 Jul;26(7):547–55. Available from: <https://doi.org/10.1007/s10654-011-9574-5>
 66. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016 Oct 1;48(10):1284–7. Available from: <https://doi.org/10.1038/ng.3656>
 67. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. 2016 Oct 1;48(10):1279–83. Available from: <https://doi.org/10.1038/ng.3643>
 68. Hong S, Prokopenko D, Dobricic V, Kilpert F, Bos I, Vos SJB, et al. Genome-wide association study of Alzheimer’s disease CSF biomarkers in the EMIF-AD Multimodal Biomarker Discovery dataset. *Transl Psychiatry*. 2020 Dec 22;10(1):403. Available from: <https://doi.org/10.1038/s41398-020-01074-z>
 69. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006 Aug;38(8):904–9. Available from: <https://doi.org/10.1038/ng1847>.
 70. Abdellaoui A, Hottenga JJ, Knijff P De, Nivard MG, Xiao X, Scheet P, et al. Population structure, migration, and diversifying selection in the Netherlands. *Eur J Hum Genet*. 2013

- Nov;21(11):1277–85. Available from: <https://doi.org/10.1038/ejhg.2013.48>.
71. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010 Sept;26(17):2190–2191. Available from: <https://doi.org/10.1093/bioinformatics/btq340>
 72. Smith SM, Douaud G, Chen W, Hanayik T, Alfaro-Almagro F, Sharp K, et al. An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nat Neurosci* 2021 245. 2021 Apr 19;24(5):737–45. Available from: <https://doi.org/10.1038/s41593-021-00826-4>
 73. Stouffer, S. A., Suchman, E. A., Devinney, L. C., Star, S. A., & Williams RM. The American soldier: Adjustment during army life. (Studies in social psychology in World War II), Vol. 1. Princeton Univ. Press. 1949. Available from: <https://psycnet.apa.org/record/1950-00790-000>
 74. König IR, Loley C, Erdmann J, Ziegler A. How to Include Chromosome X in Your Genome-Wide Association Study. *Genet Epidemiol*. 2014 Feb 1;38(2):97–103. Available from: <http://doi.wiley.com/10.1002/gepi.21782>
 75. Narasimhan V, Danecek P, Scally A, Xue Y, Tyler-Smith C, Durbin R. BCFtools/RoH: A hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics*. 2016 Jun 1;32(11):1749–51. Available from: <https://doi.org/10.1093/bioinformatics/btw044>
 76. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009 Aug;25(16):2078–9. Available from: <https://doi.org/10.1093/bioinformatics/btp352>
 77. Clayton D. Testing for association on the X chromosome. *Biostatistics*. 2008 Oct;9(4):593–600. Available from: <https://doi.org/10.1093/biostatistics/kxn007>
 78. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017 Dec 1;8(1). Available from: <https://doi.org/10.1038/s41467-017-01261-5>
 79. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010 Jul 3;38(16). Available from: <https://doi.org/10.1093/nar/gkq603>
 80. Kircher M, Witten DM, Jain P, O’roak BJ, Cooper GM, Shendure J. A general framework for

- estimating the relative pathogenicity of human genetic variants. *Nat Genet.* 2014;46(3):310–5. Available from: <https://doi.org/10.1038/ng.2892>
81. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research* 2012, 22(9):1790-1797. *Genome Res.* 2012 Sep;22(9):1790–7. Available from: <https://doi.org/10.1101/gr.137323.112>
 82. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput Biol.* 2015 Apr 1;11(4). Available from: <https://doi.org/10.1371/journal.pcbi.1004219>
 83. Watanabe K, Umićević Mirkov M, de Leeuw CA, van den Heuvel MP, Posthuma D. Genetic mapping of cell type specificity for complex traits. *Nat Commun.* 2019 Dec 1;10(1). Available from: <https://doi.org/10.1038/s41467-019-11181-1>
 84. Aguet F, Barbeira AN, Bonazzola R, Brown A, Castel SE, Jo B, et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science.* 2020 Sep 1;369(6509):1318–30. Available from: <https://doi.org/10.1126/science.aaz1776>
 85. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era — concepts and misconceptions. *Nat Rev Genet* 2008 Apr;9(4):255-66. Available from: <https://doi.org/10.1038/nrg2322>. Epub 2008
 86. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet.* 2011 Jan 7;88(1):76–82. Available from: <https://doi.org/10.1016/j.ajhg.2010.11.011>
 87. Butler, Anthony G., Philip O. Duffey, Maurice R. Hawthorne, and Michael P. Barnes. An Epidemiologic Survey of Dystonia within the Entire Population of Northeast England over the Past Nine Years. *Advances in Neurology.* 2004;94:95–99 Available from: <https://pubmed.ncbi.nlm.nih.gov/14509660/>
 88. Ortiz R, Scheperjans F, Mertsalmi T, Pekkonen E. The prevalence of adult-onset isolated dystonia in Finland 2007-2016. *PLoS One.* 2018 Nov 1;13(11). Available from: <https://doi.org/10.1371/journal.pone.0207729>.
 89. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet.* 2015 Nov 1;47(11):1236–41. Available from: <https://doi.org/10.1038/ng.3406>

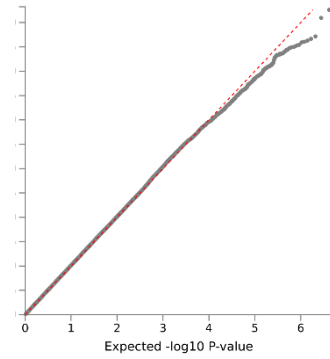
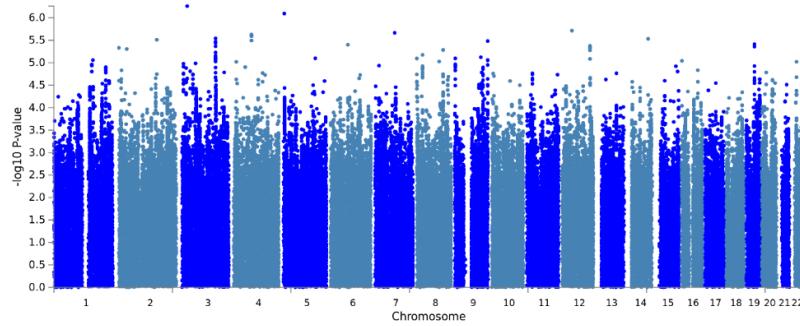
90. Tunc S, Graf J, Tadic V, Brüggemann N, Schmidt A, Al-Khaled M, et al. A population-based study on combined markers for early Parkinson's disease. *Mov Disord*. 2015 Apr 1;30(4):531–7. Available from: <https://doi.org/10.1002/mds.26100>
91. Kannarkat GT, Cook DA, Lee J-K, Chang J, Chung J, Sandy E, et al. Common Genetic Variant Association with Altered HLA Expression, Synergy with Pyrethroid Exposure, and Risk for Parkinson's Disease: An Observational and Case-Control Study. *NPJ Park Dis*. 2015;1. Available from: <https://doi.org/10.1038/npjparkd.2015.2>
92. McGurk L, Rifai OM, Bonini NM. TDP-43, a protein central to amyotrophic lateral sclerosis, is destabilized by tankyrase-1 and -2. *J Cell Sci*. 2020 Jun 23;133(12). Available from: <https://doi.org/10.1242/jcs.245811>
93. Tanji K, Mori F, Shirai F, Fukami T, Seimiya H, Utsumi J, et al. Novel tankyrase inhibitors suppress TDP-43 aggregate formation. *Biochem Biophys Res Commun*. 2021 Jan 22;537:85–92. Available from: <https://doi.org/10.1016/j.bbrc.2020.12.037>
94. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet*. 2011 Mar 11;88(3):294–305. Available from: <https://doi.org/10.1016/j.ajhg.2011.02.002>
95. Brainstorm Consortium TB, Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, et al. Analysis of shared heritability in common disorders of the brain. *Science*. 2018 Jun 22;360(6395):eaap8757. Available from: <https://doi.org/10.1126/science.aap8757>
96. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet*. 2019 Mar 1;51(3):404–13. Available from: <https://doi.org/10.1038/s41588-018-0311-9>
97. Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol*. 2019 Dec 1 ;18(12):1091–102. Available from: <https://pubmed.ncbi.nlm.nih.gov/31701892/>
98. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet*. 2019 Mar 1; 51(3):431–44. Available from: <https://doi.org/10.1038/s41588-019-0344-8>
99. Ruderfer DM, Ripke S, McQuillin A, Boocock J, Stahl EA, Pavlides JMW, et al. Genomic

- Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell*. 2018 Jun 14;173(7):1705-1715.e16. Available from: <https://doi.org/10.1016/j.cell.2018.05.046>
100. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet*. 2018 May 1;50(5):668–81. Available from: <https://doi.org/10.1038/s41588-018-0090-3>
 101. Arnold PD, Askland KD, Barlassina C, Bellodi L, Bienvenu OJ, Black D, et al. Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol Psychiatry*. 2018 May 1;23(5):1181–8. Available from: <https://doi.org/10.1038/mp.2017.154>
 102. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet*. 2018 Aug 1;50(8):1112–21. Available from: <https://doi.org/10.1038/s41588-018-0147-3>
 103. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700 000 individuals of European ancestry. *Hum Mol Genet*. 2018;27(20):3641–9. Available from: <https://doi.org/10.1093/hmg/ddy271>
 104. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018 Oct 11;562(7726):203–9. Available from: <https://doi.org/10.1038/s41586-018-0579-z>
 105. Karlsson Linnér R, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat Genet*. 2019 Feb 1;51(2):245–57. Available from: <https://doi.org/10.1038/s41588-018-0309-3>
 106. Gan-Or Z, Mencacci NE, Nalls MA. Dystonia; a roadmap is needed for future genetic studies. Vol. 58, *Parkinsonism and Related Disorders*. Elsevier Ltd; 2019. p. 9–11. Available from: <https://doi.org/10.1016/j.parkreldis.2018.12.021>
 107. Misbahuddin A, Placzek MR, Chaudhuri KR, Wood NW, Bhatia KP, Warner TT. A polymorphism in the dopamine receptor DRD5 is associated with blepharospasm. *Neurology*. 2002 Jan 8;58(1):124–6. Available from: <https://doi.org/10.1212/WNL.58.1.124>

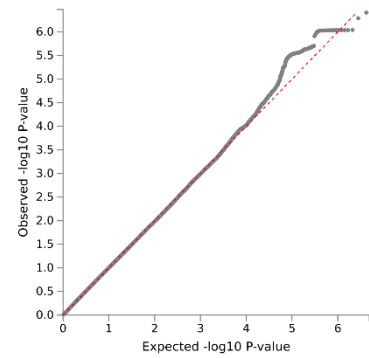
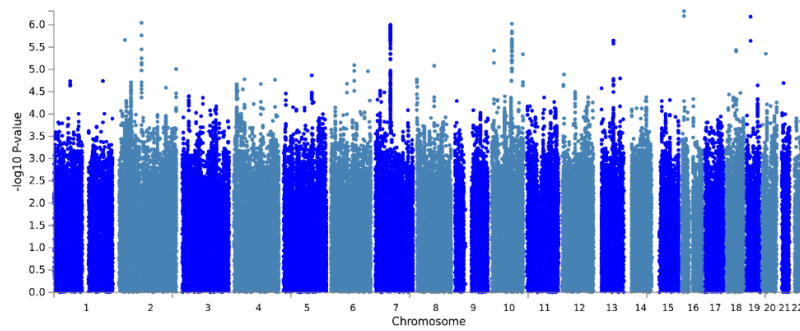
108. Placzek MR, Misbahuddin A, Ray Chaudhuri K, Wood NW, Bhatia KP, Warner TT. Cervical dystonia is associated with a polymorphism in the dopamine (D5) receptor gene. *J Neurol Neurosurg Psychiatry*. 2001 Aug;71(2):262–4. Available from: <http://dx.doi.org/10.1136/jnnp.71.2.262>
109. Allen NC, Bagade S, McQueen MB, Ioannidis JPA, Kavvoura FK, Khoury MJ, et al. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet*. 2008 Jul 1;40(7):827–34. Available from: <https://doi.org/10.1038/ng.171>
110. Castaldi PJ, Cho MH, Cohn M, Langerman F, Moran S, Tarragona N, et al. The COPD genetic association compendium: a comprehensive online database of COPD genetic associations. *Hum Mol Genet*. 2010 Feb 1;19(3):526–34. Available from: <https://doi.org/10.1093/hmg/ddp519>
111. Chatzinasiou F, Lill CM, Kypreou K, Stefanaki I, Nicolaou V, Spyrou G, et al. Comprehensive Field Synopsis and Systematic Meta-analyses of Genetic Association Studies in Cutaneous Melanoma. *JNCI J Natl Cancer Inst*. 2011 Aug 17;103(16):1227–35. Available from: <https://doi.org/10.1093/jnci/djr219>
112. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet*. 2006 Feb 15;38(2):209–13. Available from: <https://doi.org/10.1038/ng1706>
113. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020 May 28;581(7809):434–43. Available from: <https://doi.org/10.1038/s41586-020-2308-7>
114. Fuchsberger C, Abecasis GR, Hinds DA. Minimac2: Faster genotype imputation. *Bioinformatics*. 2015 Mar 1;31(5):782–4. Available from: <https://doi.org/10.1093/bioinformatics/btu704>
115. Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. Feero WG, Guttmacher AE, editors. *N Engl J Med*. 2010 Jul 8;363(2):166–76. Available from: <http://www.nejm.org/doi/10.1056/NEJMra0905980>

Anhang A

A



B



C

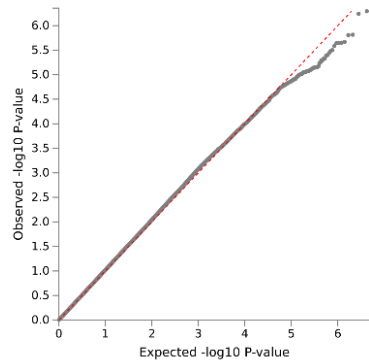
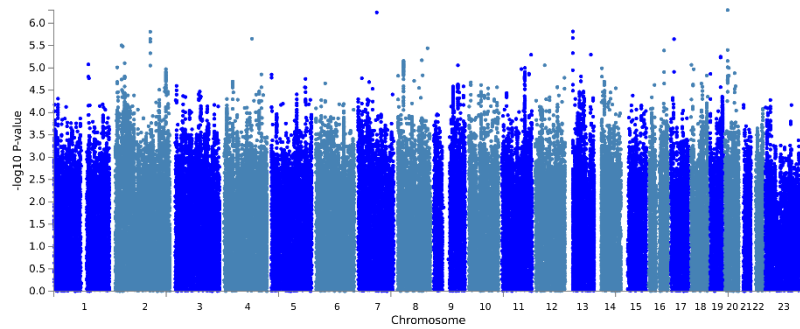
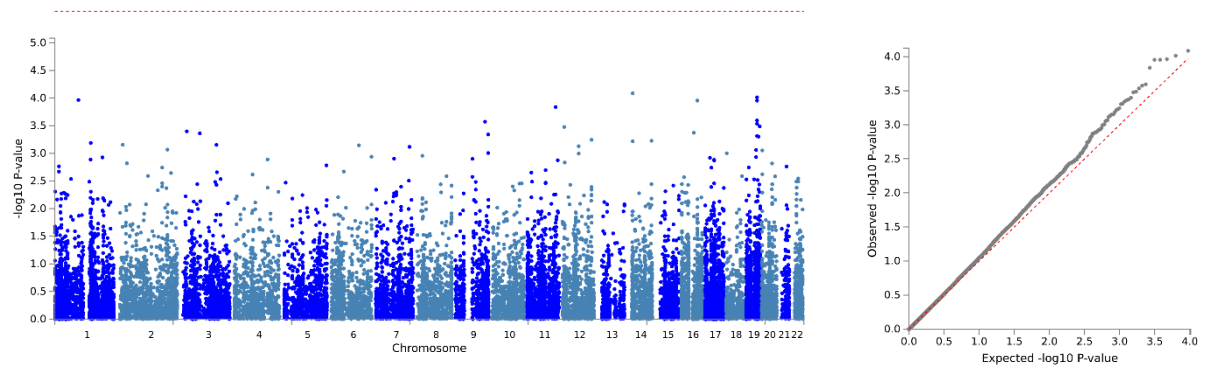
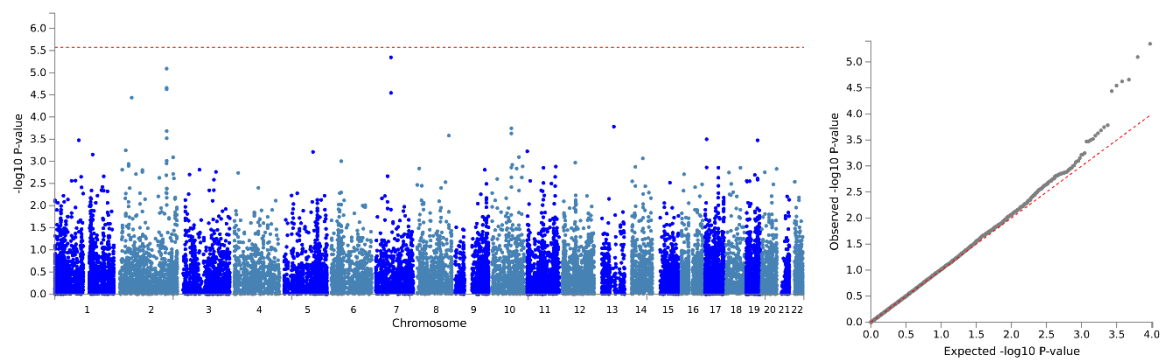


Abbildung A 1 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der **fokalen Dystonie** A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie kennzeichnet das genomweite Signifikanzniveau (5×10^{-8}). NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

A



B



C

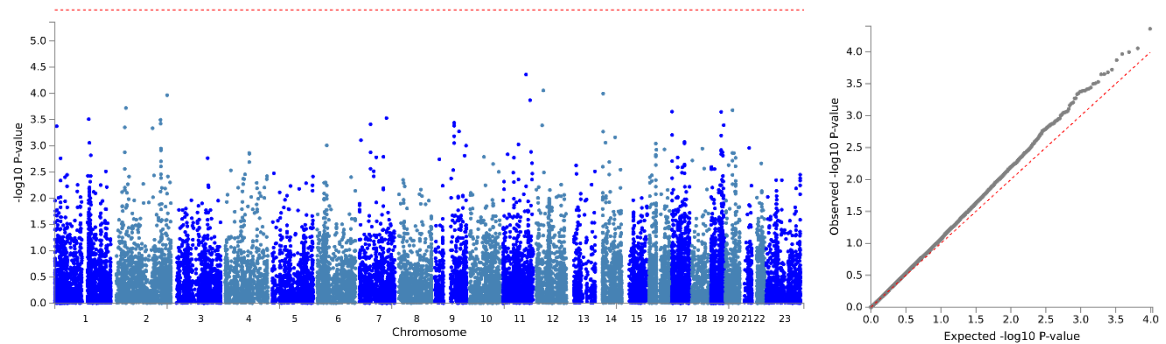
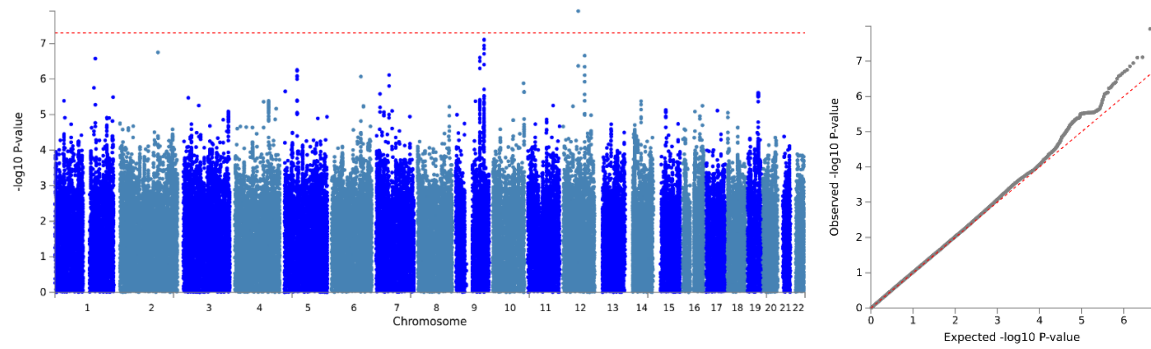
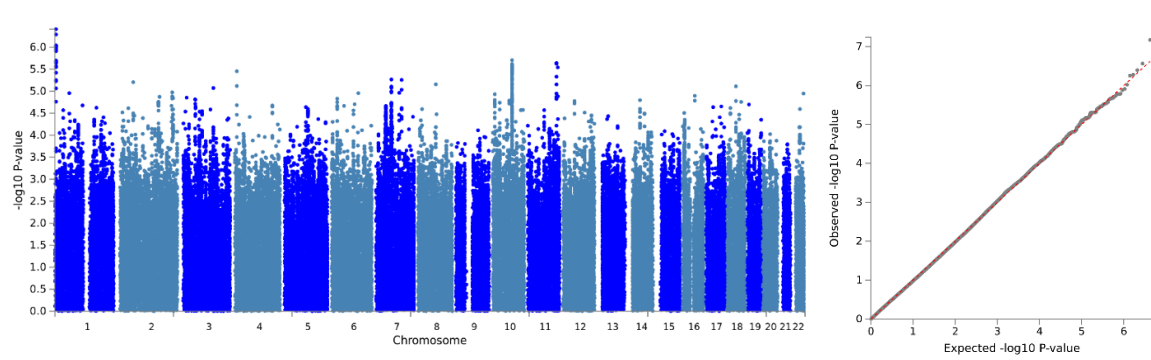


Abbildung A 2 Manhattan- und QQ-Plots der genomweiten Assoziationsergebnisse auf Gen-Ebene mit **fokaler Dystonie**. A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie zeigt den Schwellenwert für genomweite Signifikanz, d. h. $\alpha = 5 \times 10^{-6}$. NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

A



B



C

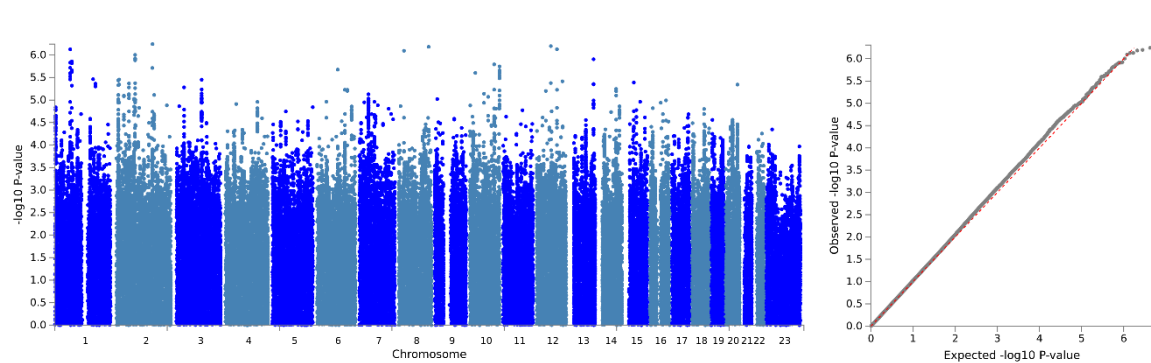
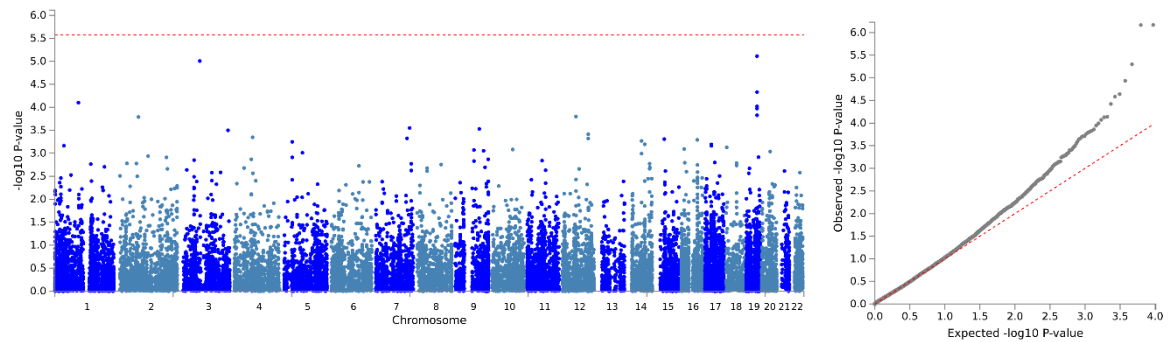
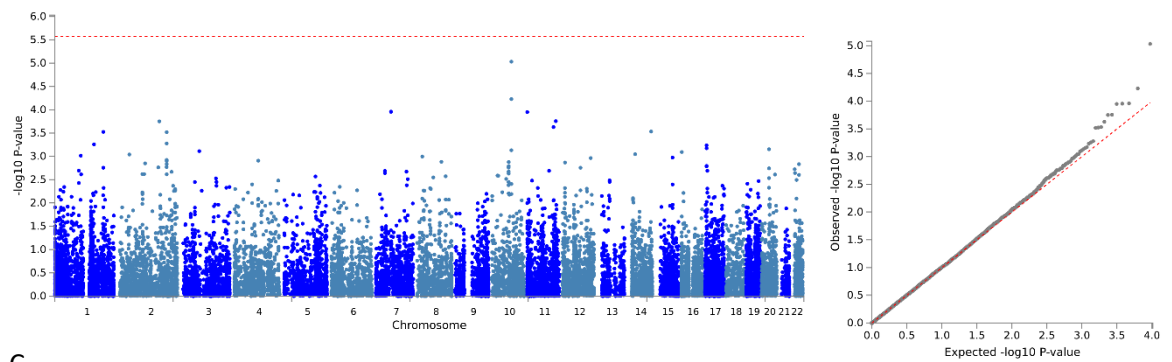


Abbildung A 3 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der **zervikalen Dystonie** A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie kennzeichnet das genomweite Signifikanzniveau (5×10^{-8}). NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

A



B



C

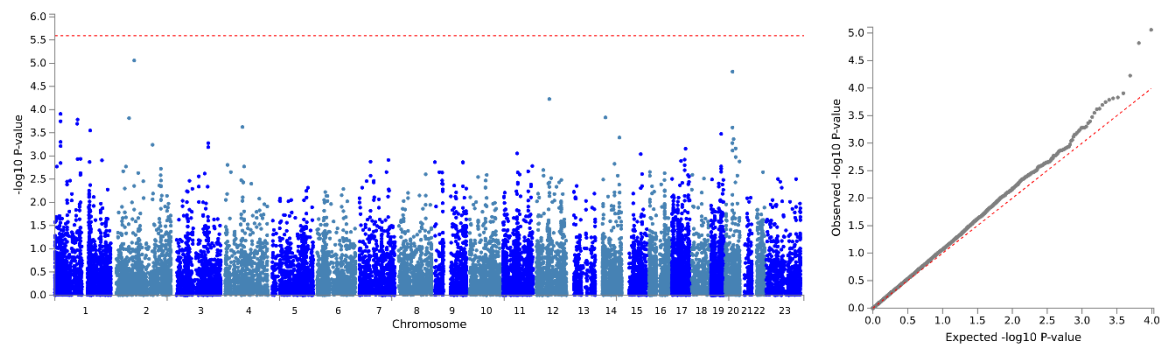


Abbildung A 4 Manhattan- und QQ-Plots der genomweiten Assoziationsergebnisse auf Gen-Ebene mit **zervikaler Dystonie**. A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie zeigt den Schwellenwert für genomweite Signifikanz, d. h. $\alpha = 5 \times 10^{-6}$. NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

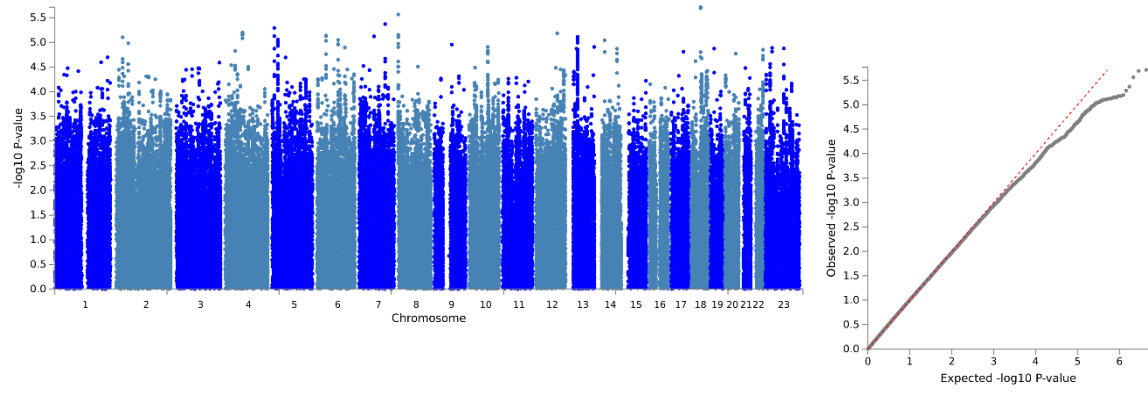


Abbildung A 5 Manhattan- und QQ-Plot der genomweiten Assoziationsergebnisse auf SNP-Ebene mit **laryngealer Dystonie** im amerikanischen Datensatz. Das Diagramm wurde mit FUMA erstellt.

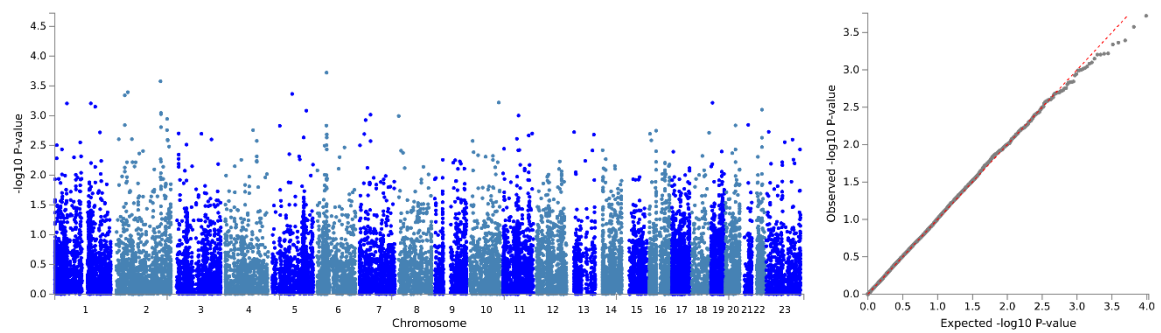
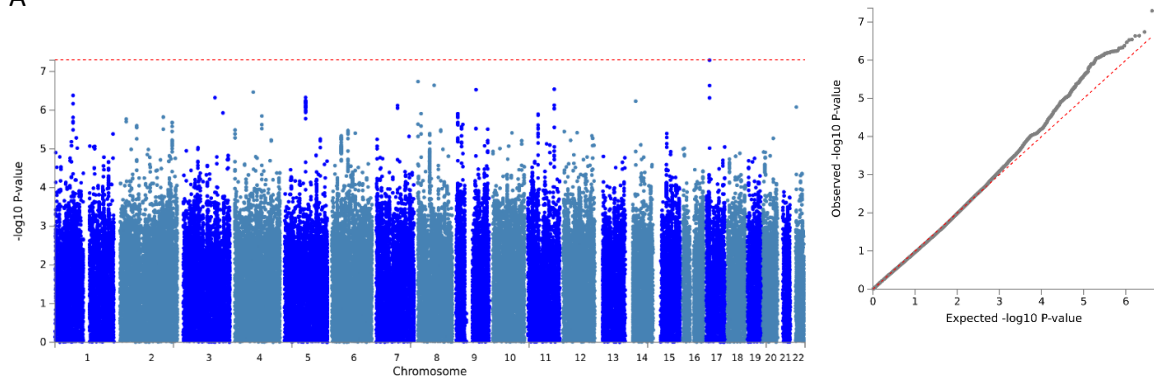
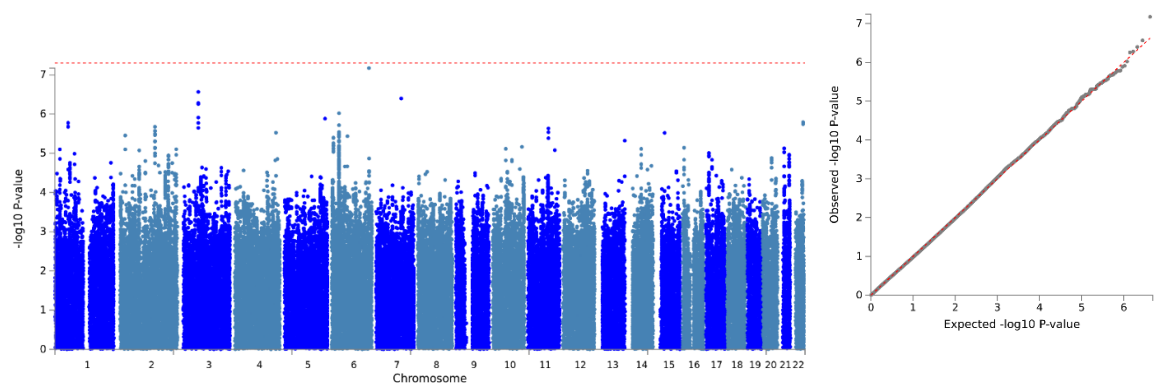


Abbildung A 6 Manhattan- und QQ-Plot der genomweiten Assoziationsergebnisse auf Gen-Ebene mit **laryngealer Dystonie** im amerikanischen Datensatz. Das Diagramm wurde mit FUMA erstellt.

A



B



C

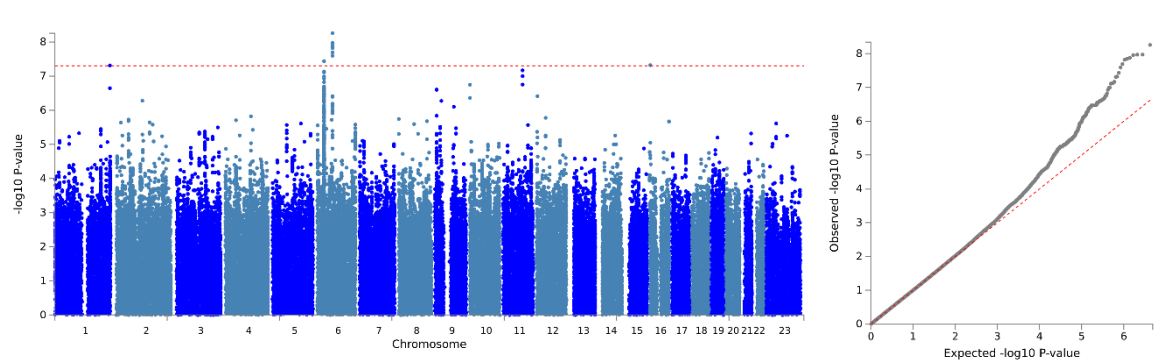
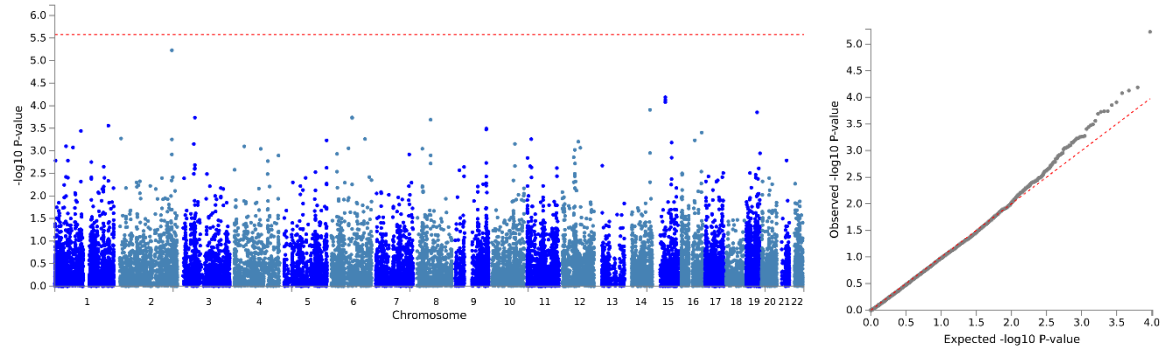
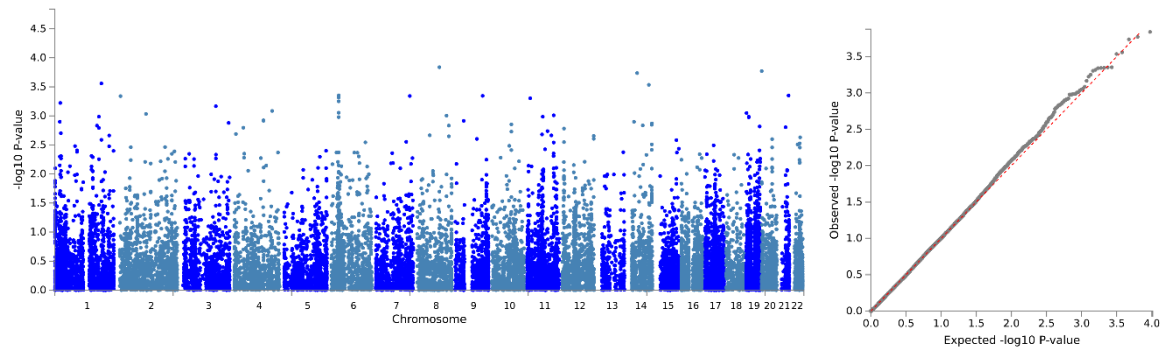


Abbildung A 7 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der **kranialen Dystonie** A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie kennzeichnet das genomweite Signifikanzniveau (5×10^{-08}). NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

A



B



C

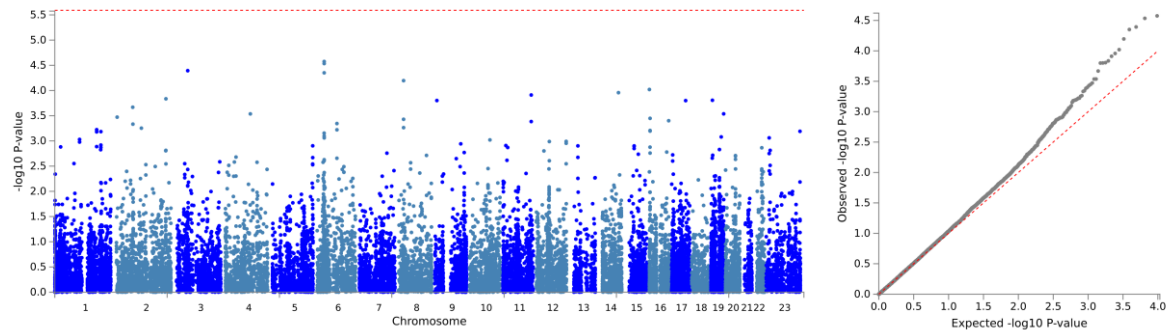
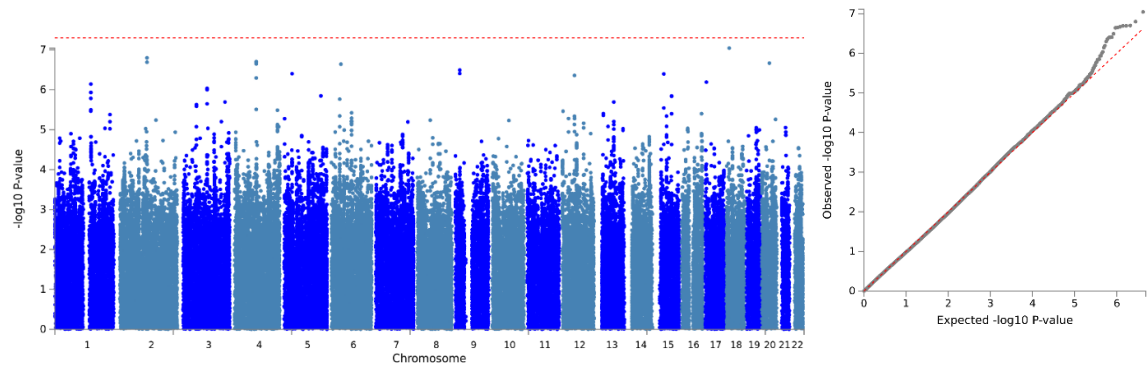
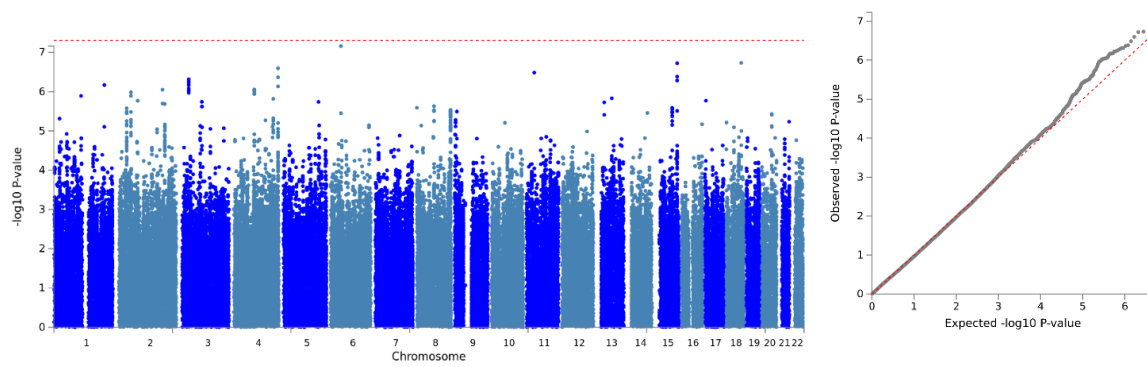


Abbildung A 8 Manhattan- und QQ-Plots der genomweiten Assoziationsergebnisse auf Gen-Ebene mit **kranialer Dystonie**. A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie zeigt den Schwellenwert für genomweite Signifikanz, d. h. $\alpha = 5 \times 10^{-6}$. NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

A



B



C

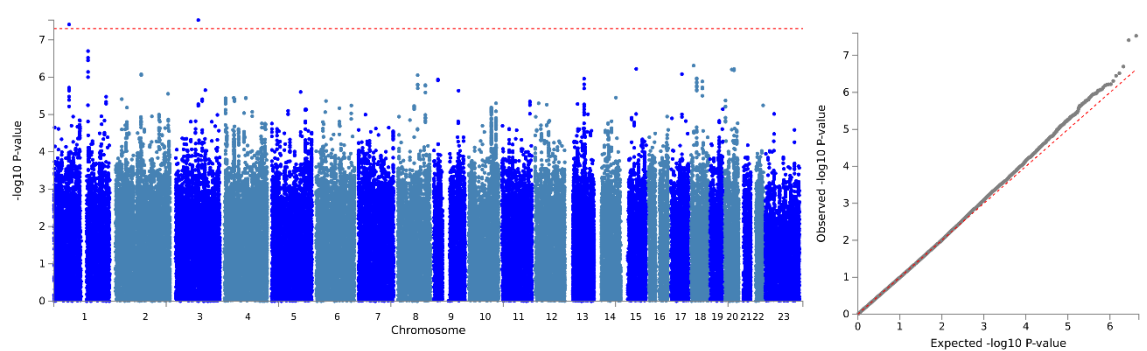
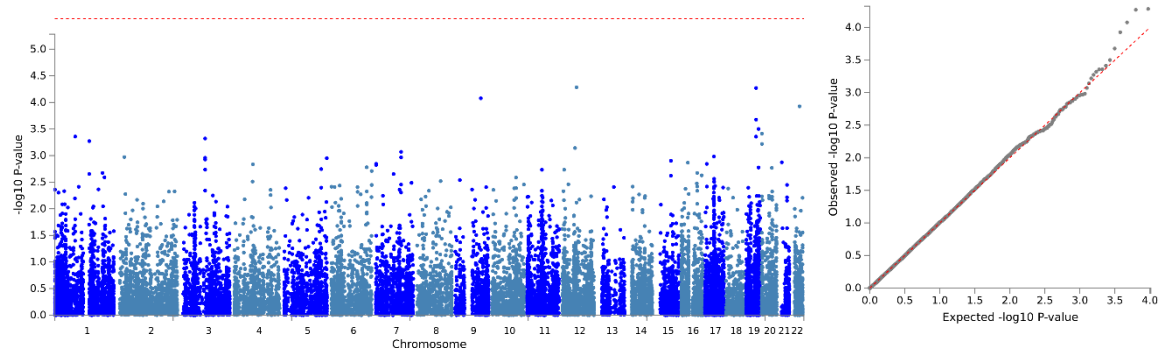
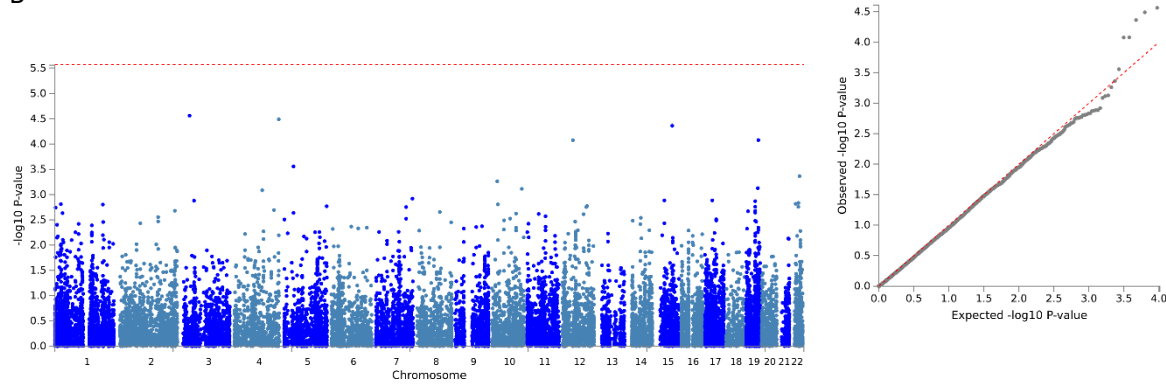


Abbildung A 9 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der **Dystonie der oberen Extremitäten** A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie kennzeichnet das genomweite Signifikanzniveau (5×10^{-08}). NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

A



B



C

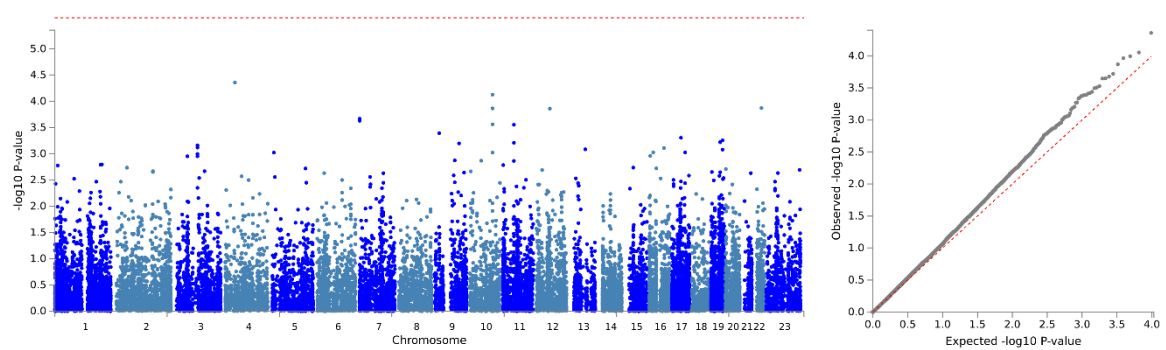


Abbildung A 10 Manhattan- und QQ-Plots der genomweiten Assoziationsergebnisse auf Gen-Ebene mit **Dystonie der oberen Extremitäten**. A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie zeigt den Schwellenwert für genomweite Signifikanz, d. h. $\alpha = 5 \times 10^{-6}$. NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

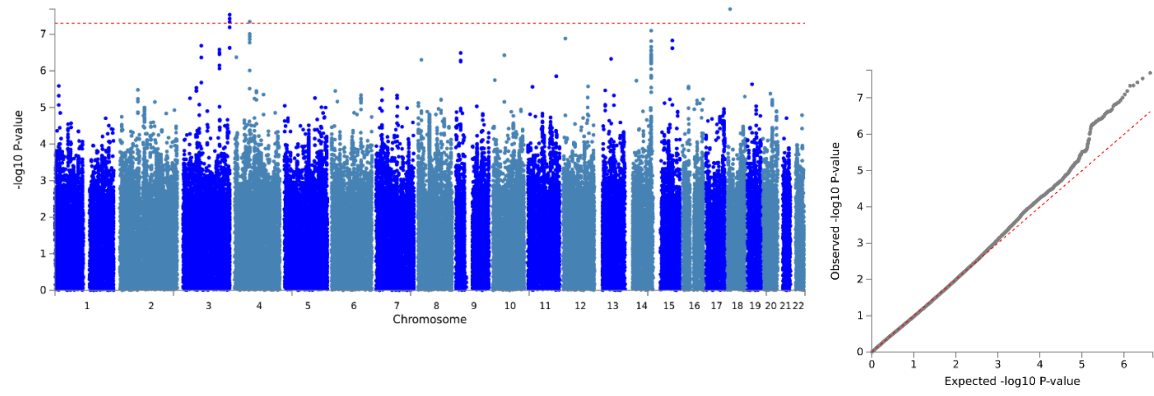


Abbildung A 11 Manhattan- und QQ-Plot der genomweiten Assoziationsergebnisse auf SNP-Ebene mit **aufgabenspezifischer Dystonie** im amerikanischen Datensatz. Das Diagramm wurde mit FUMA erstellt.

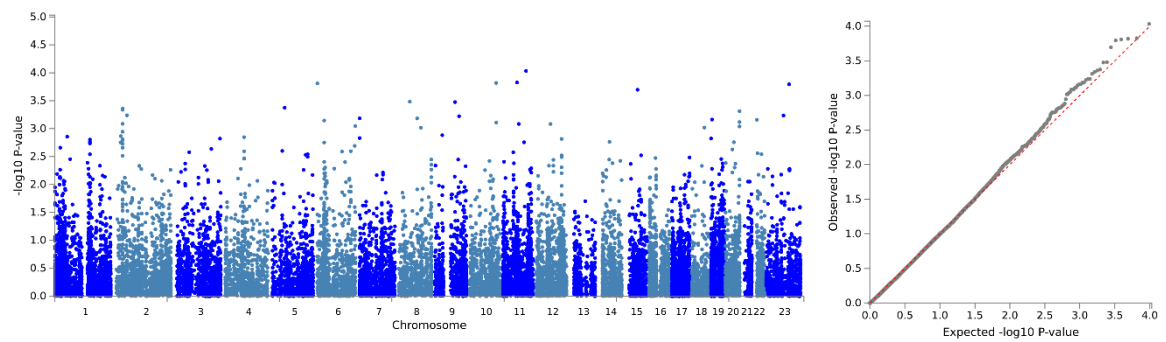
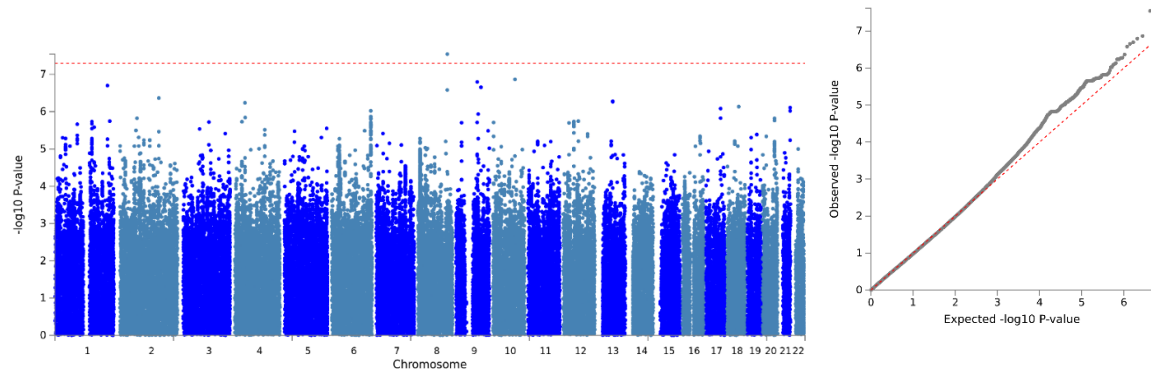
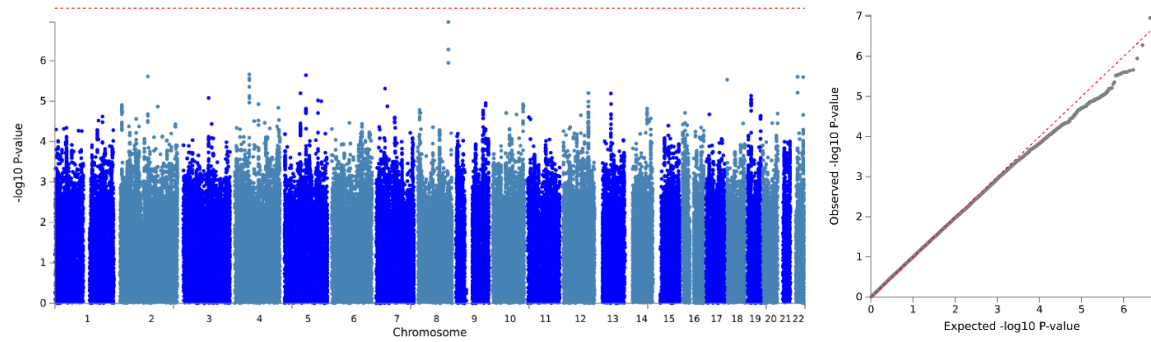


Abbildung A 12 Manhattan- und QQ-Plot der genomweiten Assoziationsergebnisse auf Gen-Ebene mit **aufgabenspezifischer Dystonie** im amerikanischen Datensatz. Das Diagramm wurde mit FUMA erstellt.

A



B



C

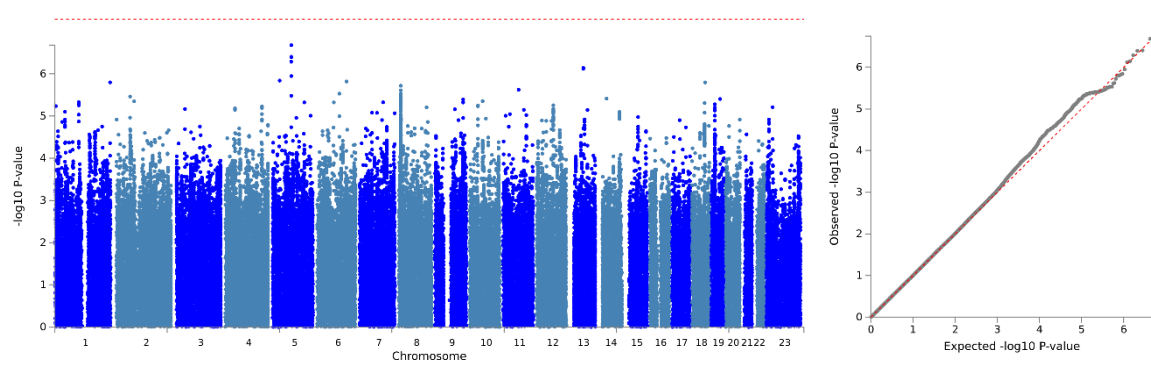
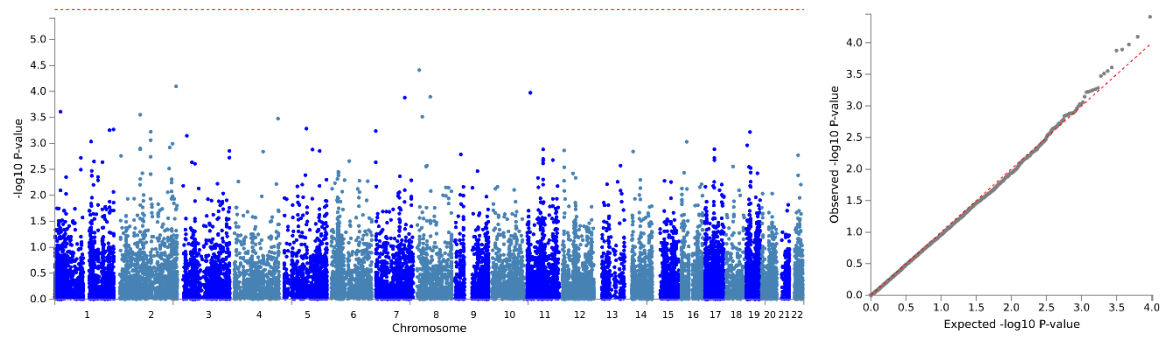
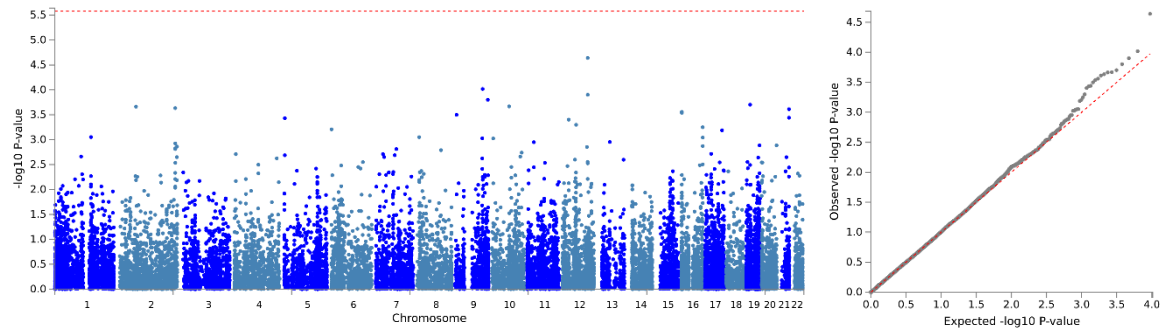


Abbildung A 13 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der **segmentalen und multifokalen Dystonie**. A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie kennzeichnet das genomweite Signifikanzniveau (5×10^{-08}). NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

A



B



C

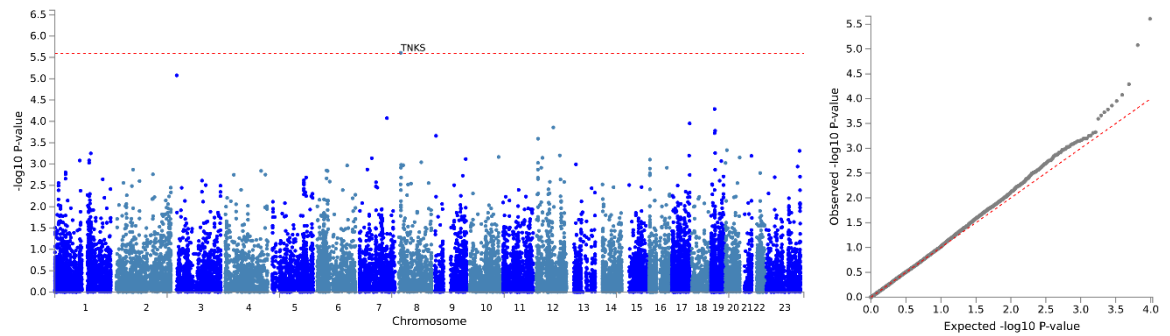


Abbildung A 14 Manhattan- und QQ-Plots der genomweiten Assoziationsergebnisse auf Gen-Ebene mit **der segmentalen und multifokalen Dystonie**. A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie zeigt den Schwellenwert für genomweite Signifikanz, d. h. $\alpha = 5 \times 10^{-6}$. NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

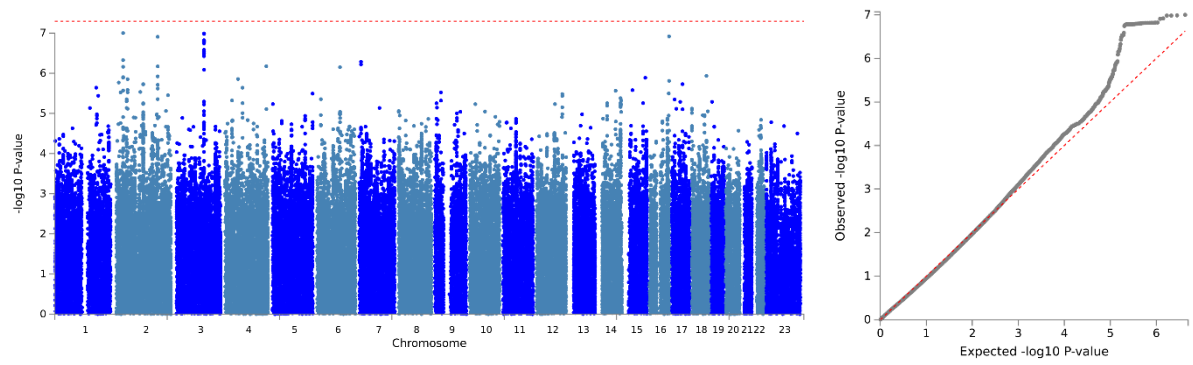


Abbildung A 15 Manhattan- und QQ-Plot der genomweiten Assoziationsergebnisse auf SNP-Ebene mit **generalisierter Dystonie** im amerikanischen Datensatz. Das Diagramm wurde mit FUMA erstellt.

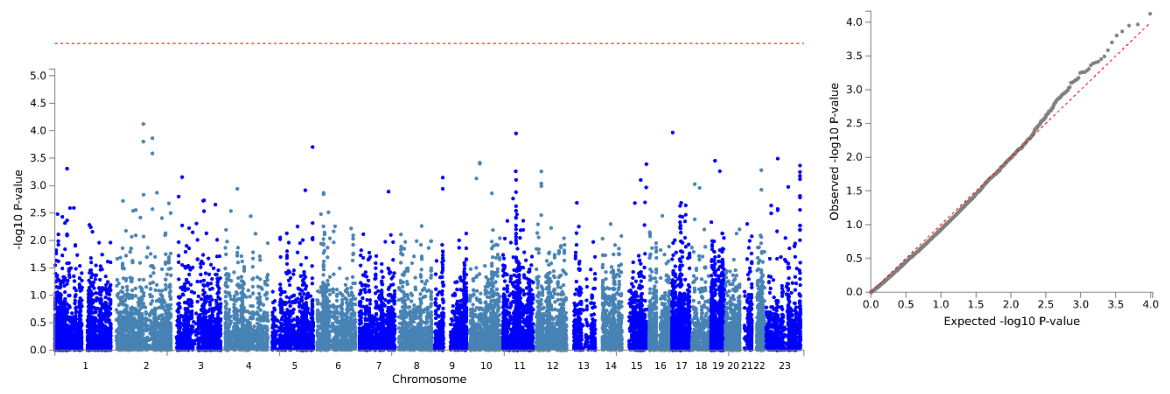
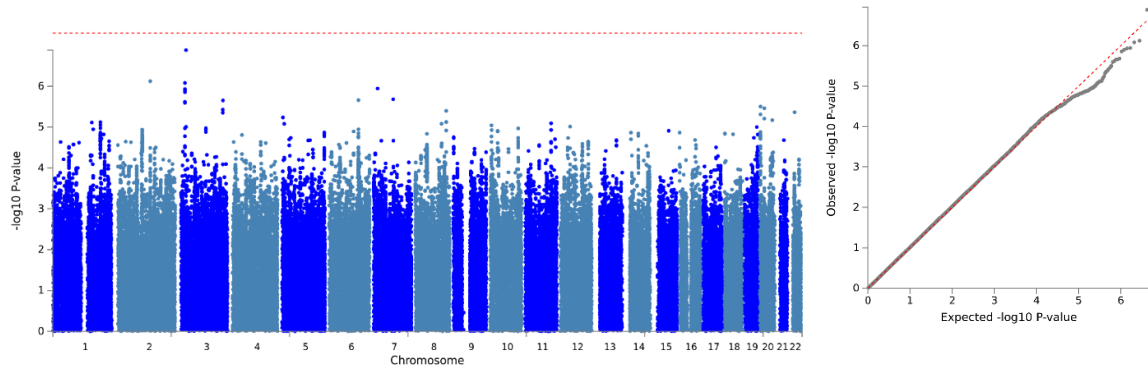
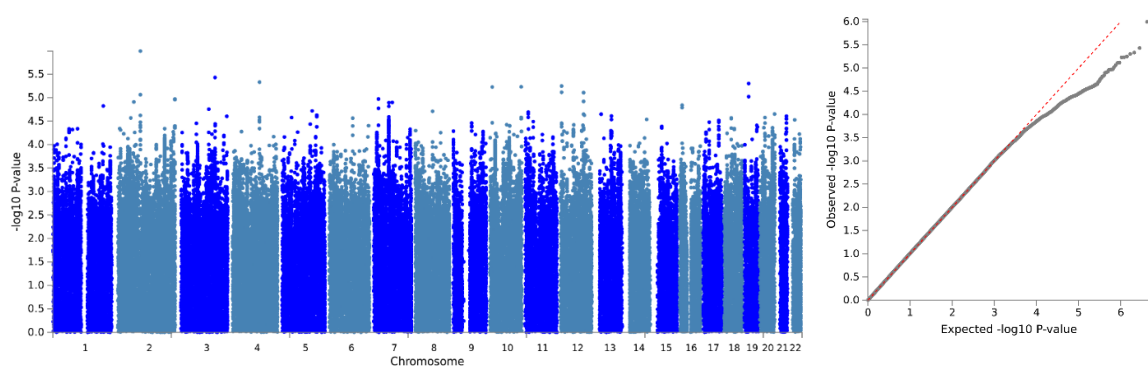


Abbildung A 16 Manhattan- und QQ-Plot der genomweiten Assoziationsergebnisse auf Gen-Ebene mit **generalisierter Dystonie** im amerikanischen Datensatz. Das Diagramm wurde mit FUMA erstellt.

A



B



C

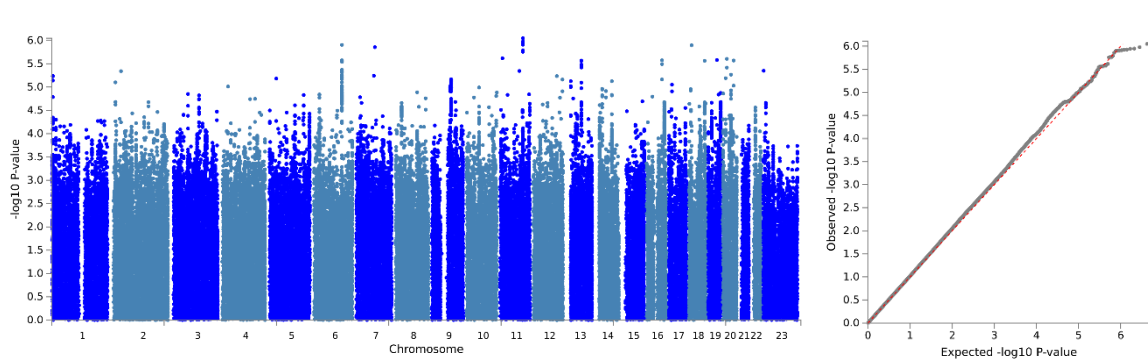
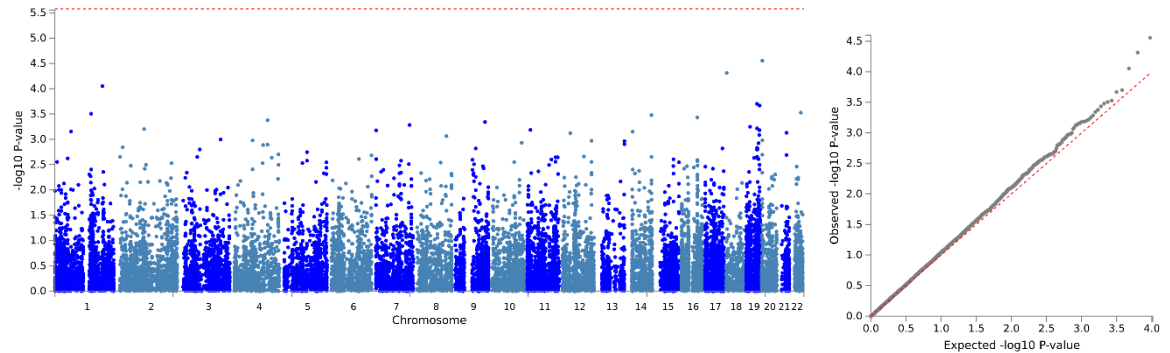
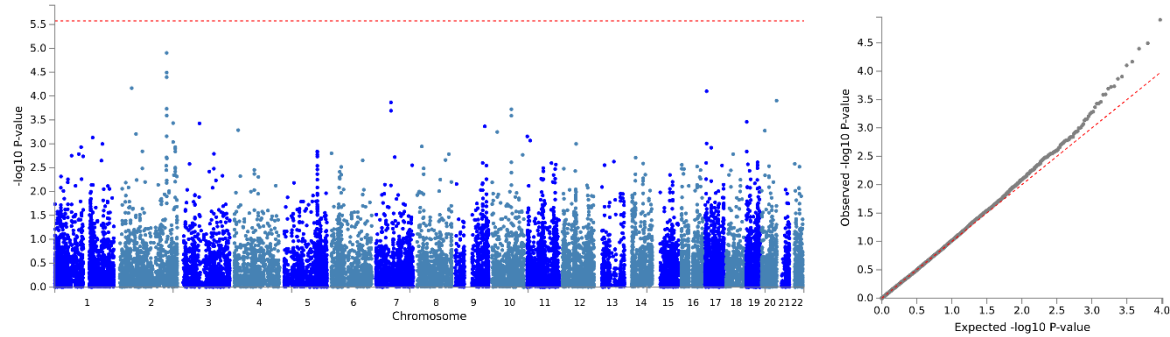


Abbildung A 17 Manhattan und QQ-Plots der genomweiten Assoziationsergebnisse auf SNP-Ebene mit der **isolierten Dystonie**. A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie kennzeichnet das genomweite Signifikanzniveau (5×10^{-8}). NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

A



B



C

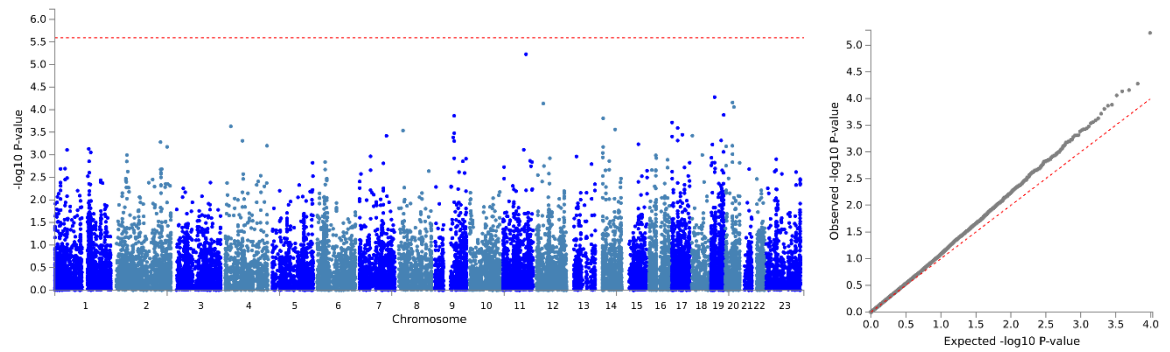
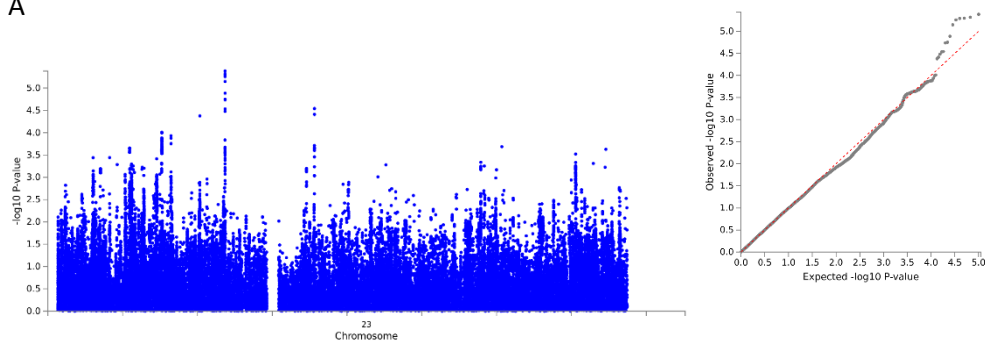


Abbildung A 18 Manhattan- und QQ-Plots der genomweiten Assoziationsergebnisse auf Gen-Ebene mit **isolierter Dystonie**. A) Deutscher Datensatz. B) Amerikanischer Datensatz. C) Metaanalyse über A+B. Die Abbildungen wurden mit FUMA erstellt. Die rote Linie zeigt den Schwellenwert für genomweite Signifikanz, d. h. $\alpha = 5 \times 10^{-6}$. NB: Die Ergebnisse des X-Chromosoms werden nur für die Ergebnisse der Metaanalyse angezeigt, da für dieses Chromosom ein anderes Analyseprotokoll verwendet wurde (Methoden).

A



B

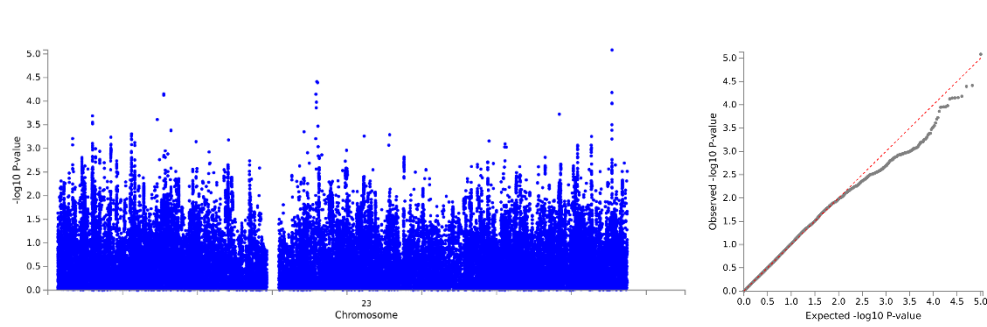


Abbildung A 19 Manhattan- und QQ-Plots für die Ergebnisse der XWAS-Metaanalysen in **fokale Dystonie**. Die Abbildungen wurden mit FUMA erstellt. A) der weiblichen Studienteilnehmer. B) der männlichen Studienteilnehmer

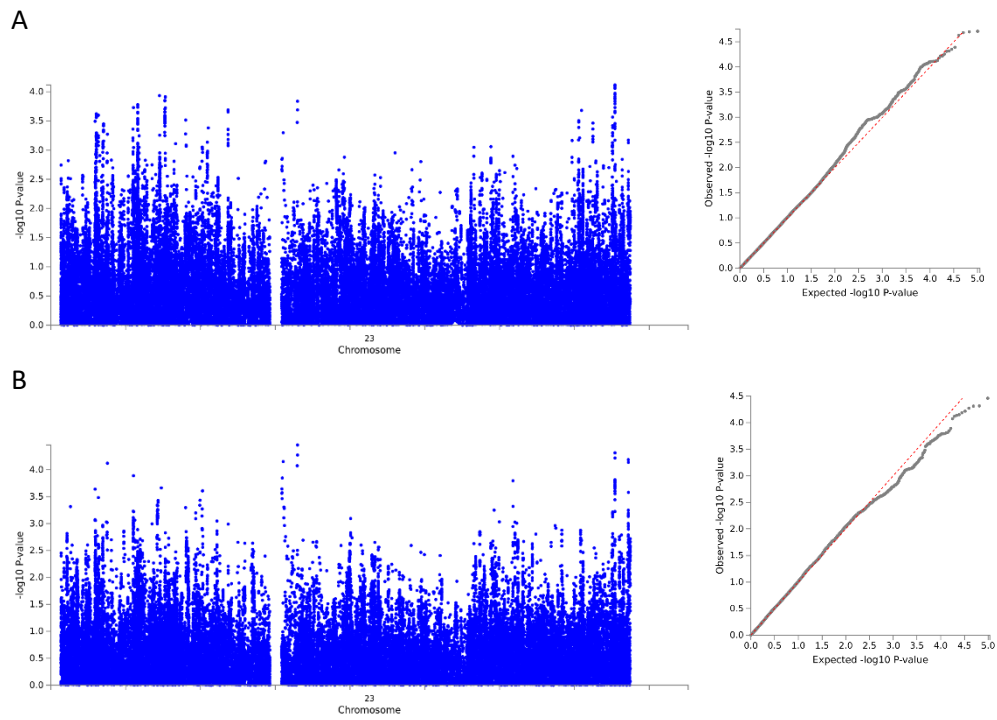
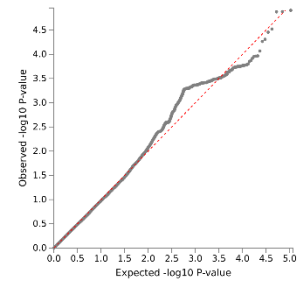
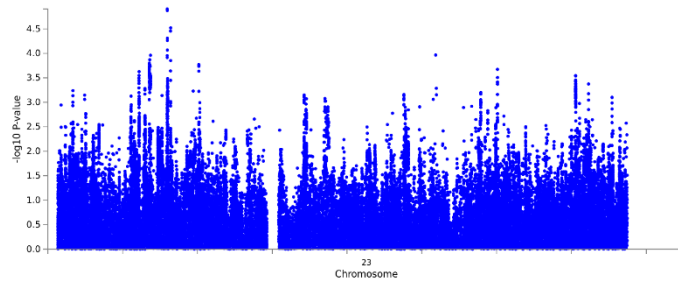


Abbildung A 20 Manhattan- und QQ-Plots für die Ergebnisse der XWAS-Metaanalysen in **zervikale Dystonie**. Die Abbildungen wurden mit FUMA erstellt. A) der weiblichen Studienteilnehmer. B) der männlichen Studienteilnehmer

A



B

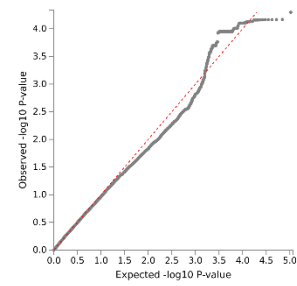
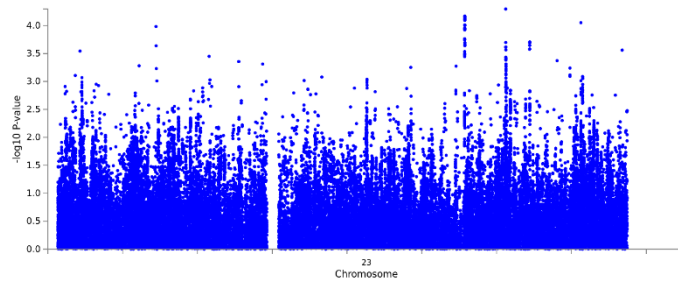
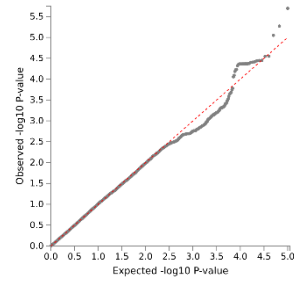
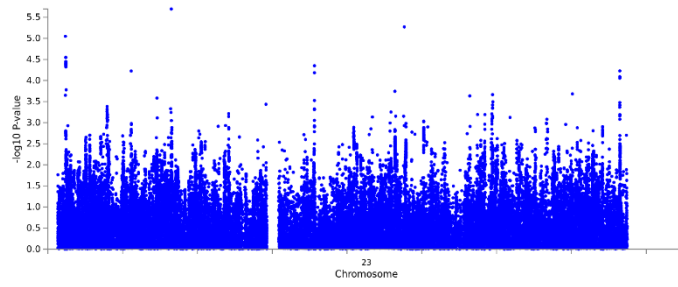


Abbildung A 21 Manhattan- und QQ-Plots für die Ergebnisse der XWAS-Metaanalysen in **laryngeale Dystonie**. Die Abbildungen wurden mit FUMA erstellt. A) der weiblichen Studienteilnehmer. B) der männlichen Studienteilnehmer

A



B

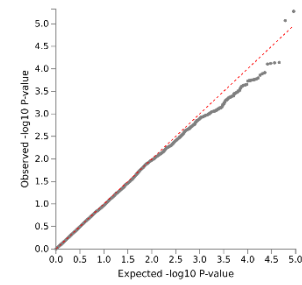
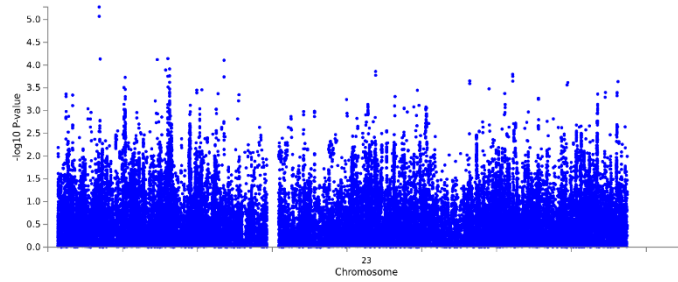
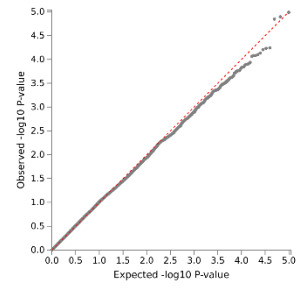
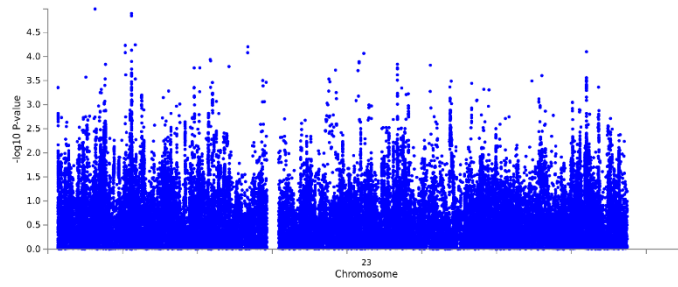


Abbildung A 22 Manhattan- und QQ-Plots für die Ergebnisse der XWAS-Metaanalysen in **kraniale Dystonie**. Die Abbildungen wurden mit FUMA erstellt. A) der weiblichen Studienteilnehmer. B) der männlichen Studienteilnehmer

A



B

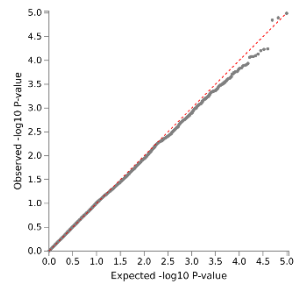
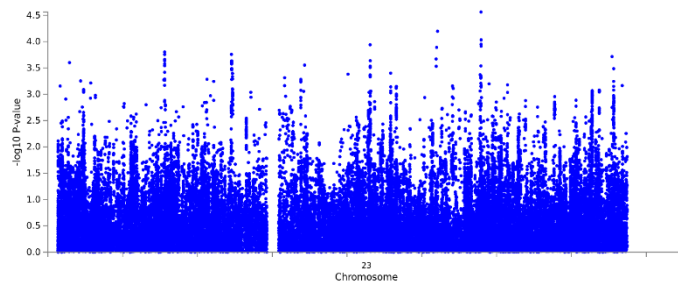
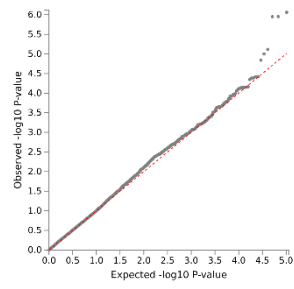
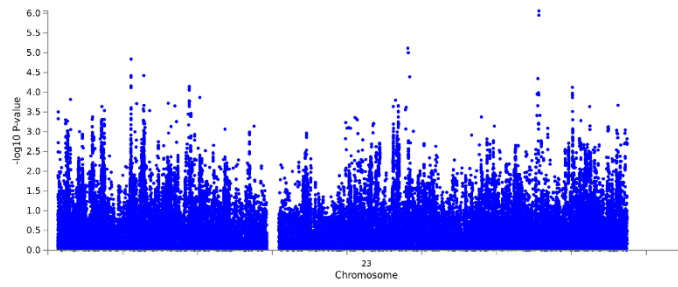


Abbildung A 23 Manhattan- und QQ-Plots für die Ergebnisse der XWAS-Metaanalysen in **Dystonie der oberen Extremitäten**. Die Abbildungen wurden mit FUMA erstellt. A) der weiblichen Studienteilnehmer. B) der männlichen Studienteilnehmer

A



B

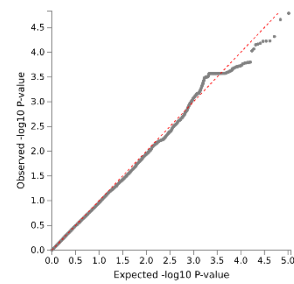
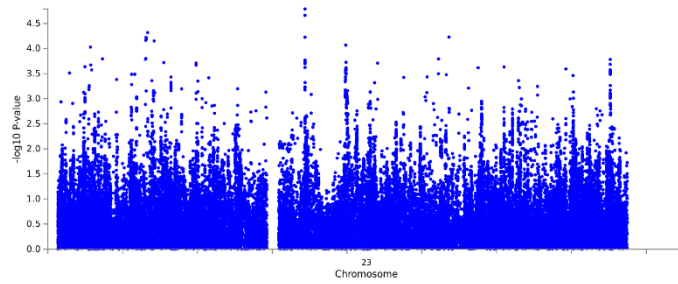
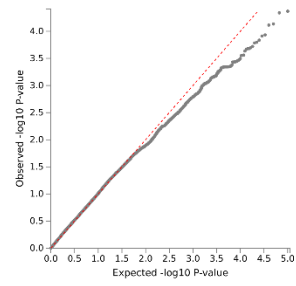
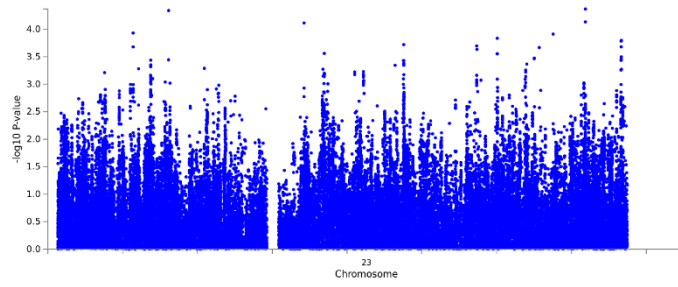


Abbildung A 24 Manhattan- und QQ-Plots für die Ergebnisse der XWAS-Metaanalysen in **aufgabenspezifische Dystonie**. Die Abbildungen wurden mit FUMA erstellt. A) der weiblichen Studienteilnehmer. B) der männlichen Studienteilnehmer

A



B

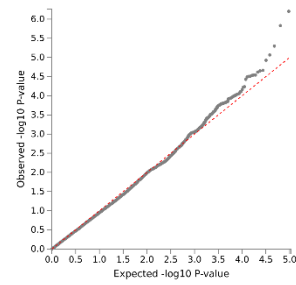
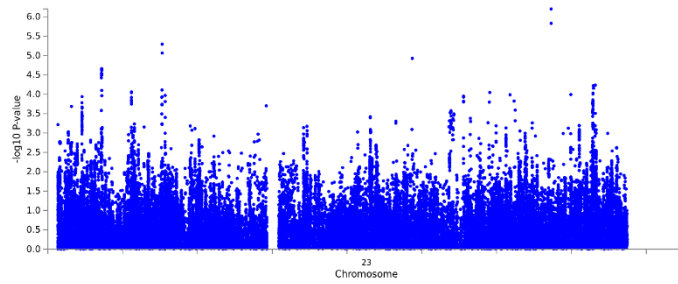


Abbildung A 25 Manhattan- und QQ-Plots für die Ergebnisse der XWAS-Metaanalysen in **segmentale und multifokale Dystonie**. Die Abbildungen wurden mit FUMA erstellt. A) der weiblichen Studienteilnehmer. B) der männlichen Studienteilnehmer

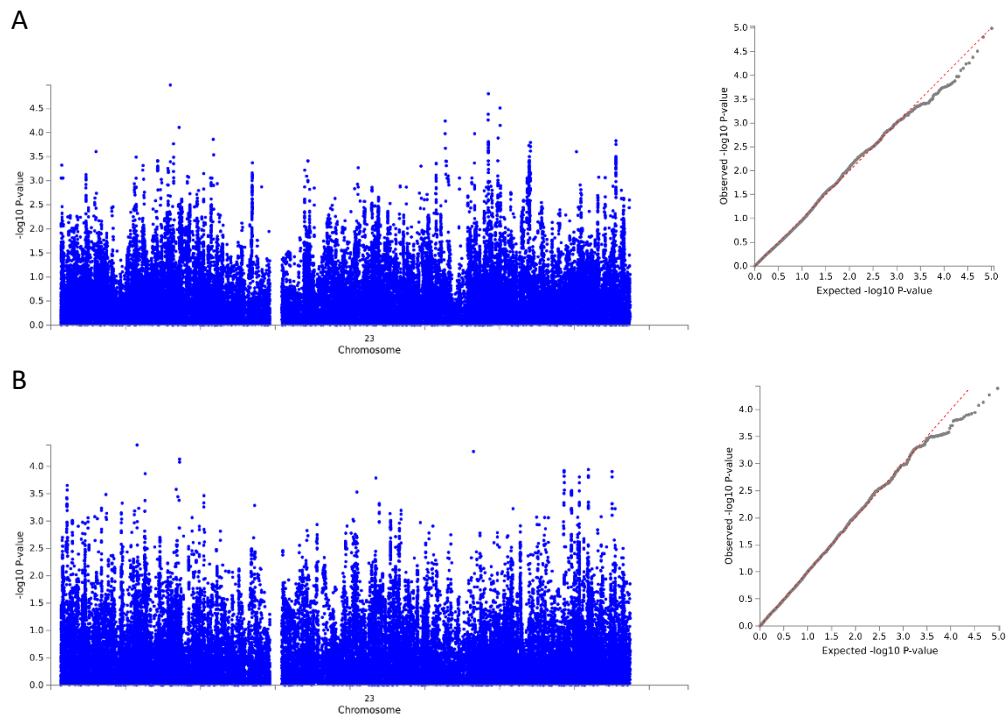
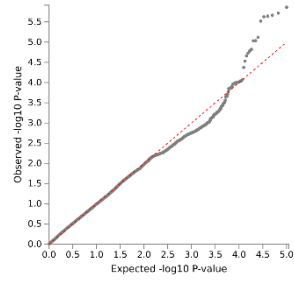
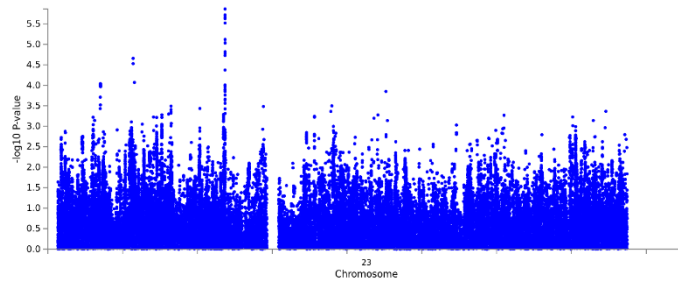


Abbildung A 26 Manhattan- und QQ-Plots für die Ergebnisse der XWAS-Metaanalysen in **generalisierte Dystonie**. Die Abbildungen wurden mit FUMA erstellt. A) der weiblichen Studienteilnehmer. B) der männlichen Studienteilnehmer

A



B

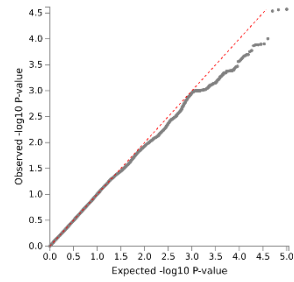
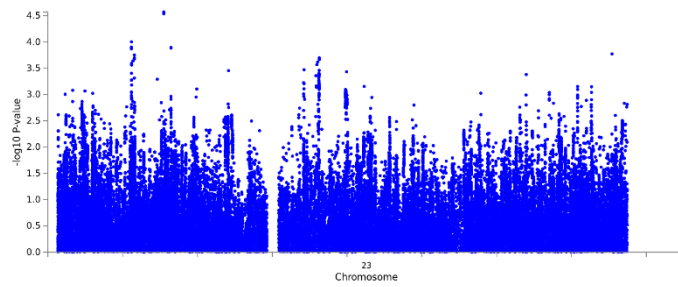


Abbildung A 27 Manhattan- und QQ-Plots für die Ergebnisse der XWAS-Metaanalysen in **isolierte Dystonie**. Die Abbildungen wurden mit FUMA erstellt. A) XWAS der weiblichen Studienteilnehmer. B) XWAS der männlichen Studienteilnehmer

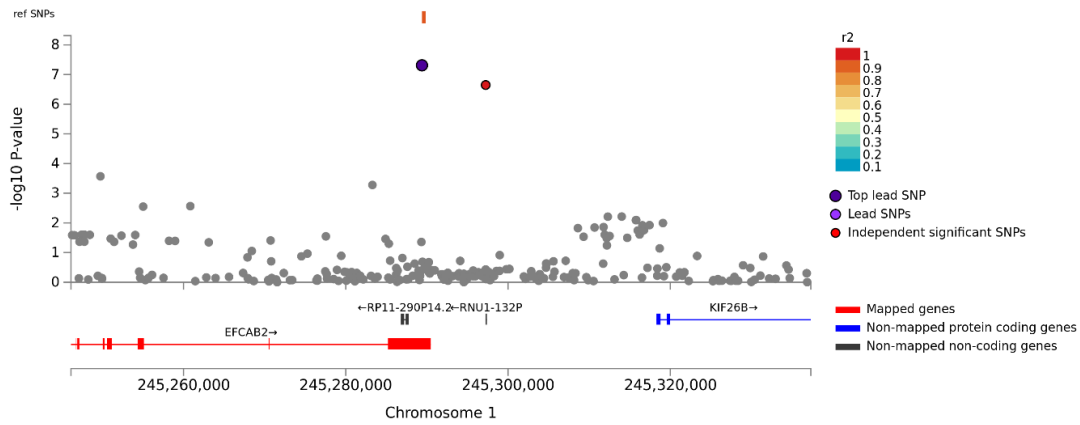


Abbildung A 28 Locus-Zoom Plot der neu entdeckten Region auf chr. 1q44 für die **kraniale Dystonie**. Der *lead*-SNP in dieser Region ist rs12022195. r^2 quantifiziert das Kopplungsungleichgewicht zwischen den *lead*-SNP und anderen SNPs in der Region. Die Genzuordnungen und die Schätzungen des Kopplungsungleichgewichts wurden mit Hilfe des FUMA-Tools annotiert (Methoden).

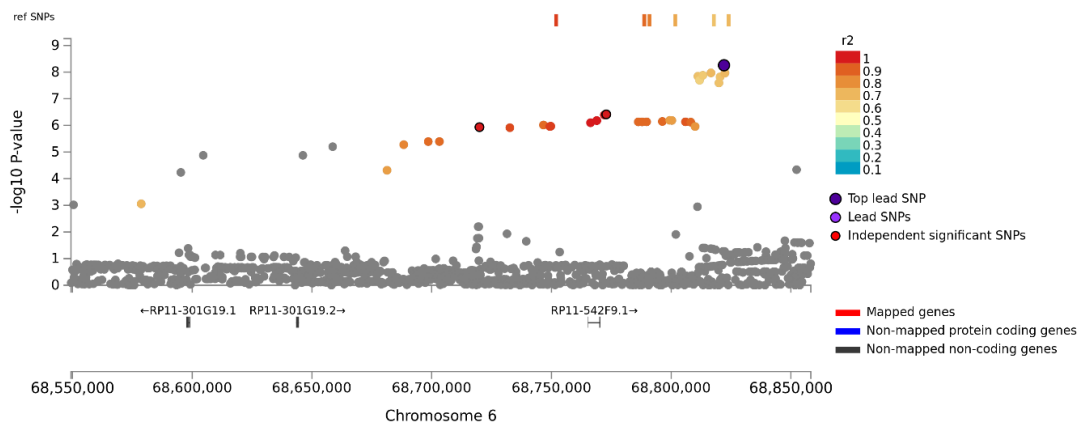


Abbildung A 29 Locus-Zoom Plot der neu entdeckten Region auf chr. 6q12 für die **kraniale Dystonie**. Der *lead*-SNP in dieser Region ist rs143591361. r^2 quantifiziert das Kopplungsungleichgewicht zwischen den *lead*-SNP und anderen SNPs in der Region. Die Genzuordnungen und die Schätzungen des Kopplungsungleichgewichts wurden mit Hilfe des FUMA-Tools annotiert (Methoden).

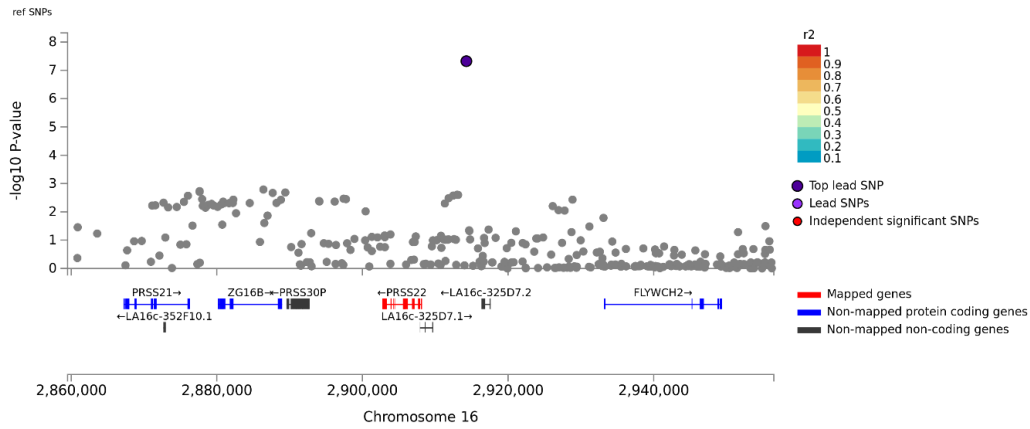


Abbildung A 30 Locus-Zoom Plot der neu entdeckten Region auf chr. 16p13.3 für die **kraniale Dystonie**. Der *lead*-SNP in dieser Region ist rs76357386. r^2 quantifiziert das Kopplungsungleichgewicht zwischen den *lead*-SNP und anderen SNPs in der Region. Die Genzuordnungen und die Schätzungen des Kopplungsungleichgewichts wurden mit Hilfe des FUMA-Tools annotiert (Methoden).

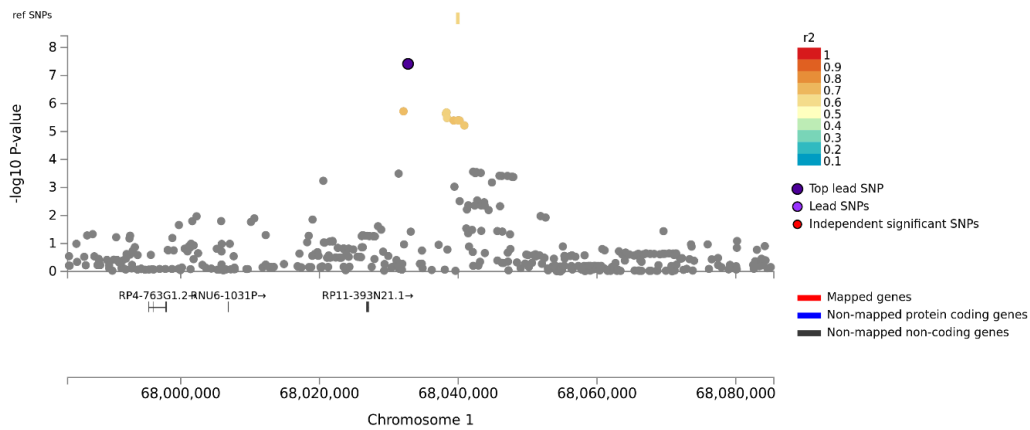


Abbildung A 31 Locus-Zoom Plot der neu entdeckten Region auf chr. 1p31.3 für die **Dystonie der oberen Extremitäten**. Der *lead*-SNP in dieser Region ist rs4081356. r^2 quantifiziert das Kopplungsungleichgewicht zwischen den *lead*-SNP und anderen SNPs in der Region. Die Genzuordnungen und die Schätzungen des Kopplungsungleichgewichts wurden mit Hilfe des FUMA-Tools annotiert (Methoden).

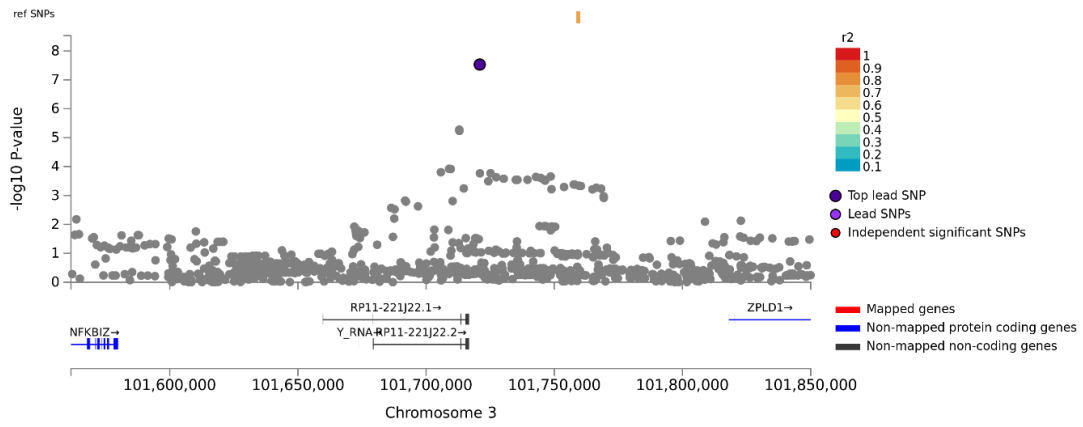


Abbildung A 32 Locus-Zoom Plot der neu entdeckten Region auf chr. 3q12.3 für die **Dystonie der oberen Extremitäten**. Der *lead*-SNP in dieser Region ist rs182982101. r^2 quantifiziert das Kopplungsungleichgewicht zwischen den *lead*-SNP und anderen SNPs in der Region. Die Genzuordnungen und die Schätzungen des Kopplungsungleichgewichts wurden mit Hilfe des FUMA-Tools annotiert (Methoden).



UNIVERSITÄT ZU LÜBECK

Universität zu Lübeck · Ratzeburger Allee 160 · 23538 Lübeck

Frau Prof. Dr. C. Klein
Institut für Neurogenetik

im Haus



Ethik-Kommission

Vorsitzender:
Herr Prof. Dr. med. A. Katalinic
Stellv. Vorsitzender:
Herr Prof. Dr. med. F. Gieseler

Geschäftsstelle:
Frau Dr. phil. A. Hüppe
Tel.: +49 451 500 5854
Fax: +49 451 500 5872
ethikkommission@uni-luebeck.de

Aktenzeichen: 16-039

Datum: 03.03.2016

zur Vorlage bei der DFG

Betreff: Votum zur FOR 2488 „Reduced penetrance in hereditary movement disorders: Elucidating mechanisms of endogenous disease protection“

Sehr geehrte Frau Professor Klein,

hiermit bestätigen wir den Eingang Ihres Funding Proposal for the Research Unit zur Vorlage bei der DFG sowie der für die ethische Beratung erforderlichen Antragsunterlagen für die einzelne Teilprojekte.

Die vorgelegten Unterlagen wurden unter berufsethischen, medizinisch-wissenschaftlichen und berufsrechtlichen Gesichtspunkten geprüft. Es bestehen **keine ethischen Bedenken** gegen die Durchführung der in den Teilprojekten geschilderten Forschungsvorhaben.

Dieses Votum ist abschließend, insofern es sich auf die Studien, wie sie im Antrag beschrieben sind, bezieht und umfasst die folgenden Teilprojekte:

- P1 Markers and mechanisms of reduced penetrance in LRRK2 mutation carriers
Grünwald (Klein, Farrer)
- P2 Reduced penetrance in Parkin and PINK1 deficiency: Prevalence, genetic modifiers, and protective mechanisms (Seibler, Klein)
- P3 Identification and characterization of genes resulting in reduced penetrance in pink1 mutant flies (Vos, Rakovic)
- P4 Identification of genetic modifiers of the penetrance of THAP1 mutations (Kaiser, Lohmann)
- P5 Genetic modifiers of disease penetrance and expression in X-linked dystonia-parkinsonism (Westenberger, Bertram)
- P6 Identification of penetrance- and risk-modifying variants in dystonia using different genome-wide technologies (Bertram, Erdmann)
- P7 Identifying determinants of reduced penetrance by analysis of mutation carriers and by age at onset analyses in Parkinson's disease (Lill, König)

1

- P8 Mendelian randomization and path models to infer causality for movement disorders with reduced penetrance (Ziegler, Del Greco Miglianico)
- INF Set-up and maintenance of a Central Project Knowledge Base and Genome Data Analysis Core (Krawczak, Bertram)
- Z1 Coordination (Klein, Pramstaller, Kaiser)
- Z2 The ProtectMove Consortium Cohorts: Longitudinal insights into penetrance of movement disorders and genetic impact at the population level (Kasten, Brüggemann, Rolfs, Hicks).

Bei Änderung der Studiendesigns sollte der Antrag erneut vorgelegt werden. Über alle schwerwiegenden oder unerwarteten und unerwünschten Ereignisse, die im Rahmen der Durchführung der Einzelstudien auftreten, ist die Kommission umgehend zu benachrichtigen. Die Deklaration von Helsinki in der aktuellen Fassung fordert in § 35 dazu auf, jedes medizinische Forschungsvorhaben mit Menschen zu registrieren. Daher empfiehlt die Kommission grundsätzlich die Studienregistrierung in einem öffentlichen Register (z.B. unter www.drks.de). Die ärztliche und juristische Verantwortung der Studienleiter und der an der Studie teilnehmenden Ärzte bleibt entsprechend der Beratungsfunktion der Ethikkommission durch unsere Stellungnahme unberührt.

Mit freundlichem Gruß



Prof. Dr. med. A. Katalinic
Vorsitzender

Tabelle A 1 Die genetische Heterogenität der Dystonie-Datensätze, das mit dem genomischen Inflationsfaktor „Lambda“ (λ) geschätzt wurde.

Diagnose	# aus Deutschland	# aus USA	# Metaanalyse
<i>Fokale</i>	1.031	1.0176	1.0237
Zervikale	1.0487	1.0348	1.0414
Laryngeale	n.v.	1.0249	n.v.
Kraniale	0.9936	1.0101	1.0119
Dystonie der oberen Extremitäten	1.0192	0.9935	0.9923
Unbekannter Ort und andere ¹	n.v.	n.v.	n.v.
Aufgabenspezifische Dystonie ²	1.022	n.v.	n.v.
<i>Multifokale + segmentale</i>	1.014	1.0251	1.0157
<i>Generalisierte</i>	n.v.	0.9387	n.v.
<i>Unbekannter Ort und Hemidystonie</i>	n.v.	n.v.	n.v.
Alle isolierten Dystonien	1.0302	1.0204	1.0381

n.v. = nicht vorhanden, aufgrund der kleinen Stichprobengröße $n < 50$.

Tabelle A 2 Vergleich der Ergebnisse der aktuellen Dystonie-GWAS mit den Ergebnissen der in dieser Arbeit erstellten GWAS-Metaanalysen.

SNP	Chr.	Position	A1	EAF	OR	p Literatur	p GWAS	Gene
1. Results from Ohlei et al., 2018 ¹								
rs4532	5.q35.2	174870150	C	0.37	1.37	1.5×10^{-03}	1.6×10^{-02}	<i>DRD1</i>
2. Results from Lohmann et al., 2014 ¹								
rs11655081	17.q24.2	66382209	C	0.02	4.33	3.95×10^{-9}	6.0×10^{-2}	<i>ARSG</i>
3. Results from Mok et al, 2014 ¹								
rs61973742	13.q33.1	102083273	A	0.9384	4.848	9.76×10^{-7}	5.48×10^{-1}	<i>NALCN</i>
4. Results from Sun et al., 2021 ¹								
rs2219975	3.q12.1	99070019	C	0.06	2.123	3.04×10^{-8}	4.73×10^{-1}	<i>COL8A1</i>
n.a.	9.q33.3	126141933- 126692431	n.a.	n.a.	n.a.	1.23×10^{-6}	5.03×10^{-1}	<i>DENND1A</i>
rs147331823	9.q22.33	101222141	G	0.01	n.a.	3.07×10^{-8}	5.45×10^{-1}	<i>GABBR2</i>

A1 = in der Literatur beschriebenes Effekt-Allel.

EAF = Häufigkeit des Effektallels (A1) gemäß den Angaben in der Literatur.

OR = in der Literatur angegebene Odds Ratio

¹ Bei diesen Einträgen bezieht sich p GWAS auf die GWAS-Ergebnisse unserer Studie unter Verwendung des entsprechenden Phänotyps, d. h. isolierte Dystonie ($n_{\text{meta}} = 8050$; rs4532) für Ohlei et al., 2018 (25), aufgabenspezifische Dystonie ($n_{\text{meta}} = 4477$) für Lohmann et al., 2014 (39) and zervikale Dystonie ($n_{\text{meta}} = 6029$) für Mok et al, 2014 (40) und Sun et al, 2021(41).

DANKSAGUNG

Am Ende dieser Arbeit möchte ich mich bei allen Personen bedanken, die mich bei dieser Dissertation unterstützt haben. Ein besonderer Dank ist an Dr. Valerija Dobricic für die Vorbereitung der DNA-Proben im LIGA-Labor und Ihre aktive Unterstützung bei der Durchführung der Field-Synopsis gerichtet. Weiterhin möchte ich dem Team des IKMB am UKSH-Campus Kiel (v.a. Frau Tanja Wesse, Frau Sanaz Sedghpour Sabet, Herrn Dr. Micheal Wittig und Prof. Dr. Andre Franke) für die Durchführung der Genotypisierungsexperimente und Dr. Fabian Kilpert der Arbeitsgruppe LIGA für seine Hilfe bei der Qualitätskontrolle und Imputation der Genotypisierungsdaten danken. Dem gesamten Team der Arbeitsgruppe LIGA danke ich für die langjährige Unterstützung und Feedback während meiner zahlreichen Progress-Reports in den regelmäßigen Treffen der Arbeitsgruppe.

Ein ganz besonderer Dank gebührt Herrn Prof. Dr. Lars Bertram. Seine hervorragende Unterstützung und die Expertise im Bereich der GWAS-Analysen haben mir immer stets geholfen die Analysen der vorliegenden Dissertation durchzuführen und die Resultate zu interpretieren. Ich bedanke mich recht herzlich für seine Betreuung und die Zurverfügungstellung des Themas.

Ein herzliches Dankeschön auch an alle Dystonie-Patienten und Kontrollpersonen für ihre Teilnahme an dieser Studie. Des Weiteren möchte ich den MitarbeiterInnen der DFG-Forschungsgruppe FOR2488 für die Bereitstellung der DNA-Proben und die Zusammenstellung und Übermittlung der relevanten klinischen Informationen der hier eingeschlossenen Patienten und Kontrollpersonen danken (v.a. Prof. Dr. Meike Kasten, Prof. Dr. Norbert Brüggemann, Dr. Eva-Juliane Vollstedt und Prof. Dr. Christine Klein). Vielen Dank gebührt auch der PopGen Arbeitsgruppe (v.a. Dr. Lukas Tittmann und Prof. Dr. Wolfgang Lieb) und Frau Prof. Beate Ritz und Team an der UCLA, USA, zur Verfügungstellung von Kontrollproben aus den PopGen- und PEG-Studien.

Zu guter Letzt möchte ich mich bei meiner gesamten Familie und meiner Freundin Frau Tanja Kruse bedanken, auf deren Rückhalt und Liebe ich immer vertrauen konnte und kann.