Technische Universität München
TUM School of Medicine and Health

TUM

Molecular analysis of movement disorders – genomic and epigenomic approaches

Philip Wolfgang Harrer

Abstract

Movement disorders are a diverse group of neurological diseases characterized by irregular tone, posture, and insufficient control of voluntary and involuntary movements. The genetics of movement disorders include both monogenic modes of inheritance due to rare variants carrying a large effect size, often directly impacting the function of encoded proteins, and polygenic forms involving many loci with variants that individually have only mild effects on the disease phenotype. Functional investigations of monogenic forms usually are successful by focusing on the disease gene and the functions of its encoded protein while research on polygenic conditions needs to address the entire concert of the many small effects which can be done by studying consequences on other omics-levels such as epigenomics, transcriptomics, and proteomics.

In this work we analyzed the pathogenesis of movement disorders from these two sides:

i) We have identified variants in two genes, *NUP54* and *EIF4A2*, recognizing two novel monogenic forms of dystonia, and provided functional evidence for two potential disease mechanisms involving changes at the nuclear envelope and impaired regulation of protein translation by targeted protein analyses. These analyses aimed to unveil changes in expression patterns, shifts in subcellular locations, disruptions in protein-protein interactions, and the subsequent functional impact.

ii) In case of the polygenic movement disorder restless legs syndrome (RLS) we performed epigenome-wide association studies (EWAS) and epigenetic age estimations which allowed us to support the notion of altered neurodevelopment but not neurodegeneration playing a role in RLS, and to derive a reliable methylation risk score for RLS.

In conclusion, the study of monogenic as well as polygenic causes of movement disorders continues to unveil new genetic associations and insights into disease mechanisms. Functional analyses and omics approaches contribute to a deeper understanding of these complex neurological conditions.

Zusammenfassung

Bewegungsstörungen bilden eine heterogene Gruppe neurologischer Erkrankungen, welche durch einen unregelmäßigen Tonus, eine beeinträchtigte Körperhaltung sowie eine unzureichende Kontrolle der willkürlichen und unwillkürlichen Bewegungen gekennzeichnet sind. Die Genetik von Bewegungsstörungen umfasst sowohl monogene Vererbungsformen, die auf seltene Varianten mit großer Effektstärke zurückzuführen sind und sich häufig direkt auf die Funktion kodierter Proteine auswirken, als auch polygene Formen, bei denen Varianten in multiplen Loci beteiligt sind, welche im Einzelnen vergleichsweise geringe Effekte auf den Phänotyp der Erkrankung haben. Während funktionelle Untersuchungen zu monogenen Formen eine hohe Aufklärungsrate aufweisen, wenn sie sich auf das Krankheitsgen und die Funktionen des von ihm kodierten Proteins konzentrieren, muss die Ursachenforschung bei polygenen Erkrankungen die Gesamtheit der zahlreichen kleineren Effekte berücksichtigen. Dies kann durch die Untersuchung der Auswirkungen auf andere Omik-Ebenen wie Epigenom, Transkriptom und Proteom erfolgen.

In dieser Arbeit analysierten wir die Pathogenese von Bewegungsstörungen basierend auf monogenen und polygenen Einflüssen:

i) Wir haben Varianten in zwei Genen, *NUP54* und *EIF4A2*, identifiziert, die zwei neue, monogene Formen der Dystonie kennzeichnen, und durch gezielte Proteinanalysen funktionelle Anhaltspunkte für zwei potenzielle Krankheitsmechanismen erbracht, welche mit Veränderungen an der Kernhülle und einer beeinträchtigen Regulation der Proteintranslation verknüpft sind. Die Zielsetzung der durchgeführten Analysen umfasste die Identifizierung veränderter Expressionsmuster, Verschiebungen subzellulärer Lokalisationen, Störungen in Protein-Protein-Interaktionen und die daraus resultierenden funktionellen Auswirkungen.

ii) Im Rahmen der polygenen Bewegungsstörung Restless Legs Syndrom (RLS) führten wir epigenomweite Assoziationsstudien (EWAS) und epigenetische Altersabschätzungen durch, was die Bestätigung der Annahme ermöglichte, dass eine veränderte neuronale Entwicklung, nicht aber eine Neurodegeneration von

pathomechanistischer Bedeutung ist. Darüber hinaus konnten wir einen zuverlässigen epigenetischen Risikoscore für RLS generieren.

Zusammenfassend ermöglicht die Untersuchung monogener und polygener Ursachen von Bewegungsstörungen weiterhin neue genetische Assoziationen und Erkenntnisse über bisher unbekannte Krankheitsmechanismen. Funktionelle Analysen und Omik-Ansätze tragen zu einem tieferen Verständnis dieser komplexen neurologischen Erkrankungen bei.

Table of Contents

1. Introduction and methodology

The understanding of the role of genetic predisposition in disease etiology has evolved significantly over time, revealing different genetic patterns underlying diverse diseases, each necessitating adapted investigative approaches. Movement disorders, a heterogeneous group with widespread genetic conditions, exemplify this diversity.

1.1 Spectrum of genetic variants

Genetic variants represent a fundamental force that shapes phenotypic variation. Mutation rates vary from roughly 1 to 1.7 per 100 million nucleotides [1-3]. Based on the size of the human genome of approximately 3.2×10^9 base pairs, roughly 30 to 50 de novo point mutations and a lesser number of larger variants are introduced with each genome replication [1-3]. The distribution of variants, however, is not even across the human genome and varies between genomic regions [4]. Variants can differ in size, composition (non-structural vs. structural), location (within the coding vs. non-coding regions or intra- vs. intergenic regions), frequency and pathogenicity.

Variants contributing to the pathogenesis of a disease are referred to as pathogenic variants. The phenotype is the consequence of nonlinear interactions among genetic variants and nongenetic factors [5]. In addition to the variant with the largest effect size, often referred to as "causal" variant, other variants might lessen or accentuate the phenotype [5, 6].

The allele frequency (AF) is used to estimate how frequently a variant occurs in a population. Variants are often defined as common at AF > 5%, low-frequency at AF 1-5%, rare at AF < 1%, and very rare at AF < 0.1% [7] (**Figure 1**). Effect sizes of genetic variants are likely gradual with variants of high effect sizes and penetrance on the one end and variants with modest effects, often contributing to common traits, on the other end of the spectrum [8].
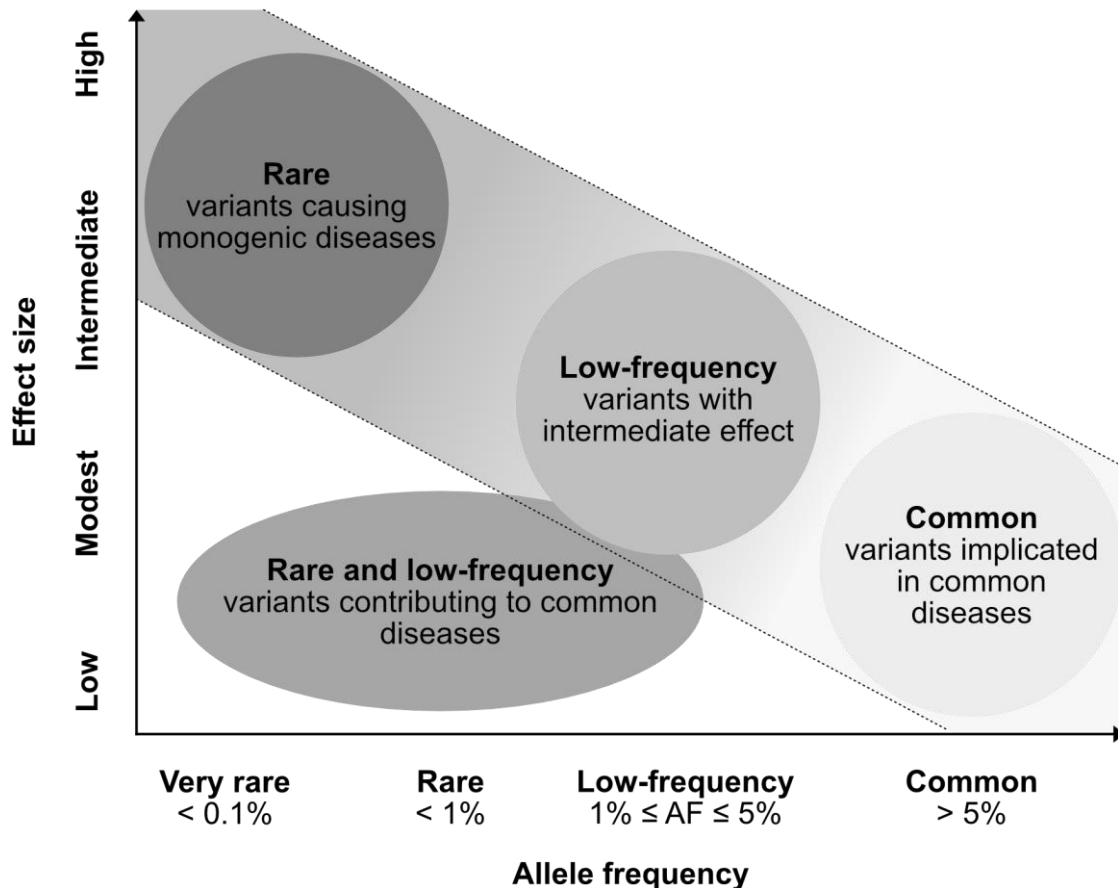
Figure 1: Relationship between allele frequency and effect size of genetic variants
(modified from *Manolio et al.* [9])

Common diseases, like the neurological movement disorder restless legs syndrome (RLS), arise from multiple predisposing factors, both genetic and non-genetic. According to the original "common disease - common variant" hypothesis [10-12], common diseases are caused by interaction of multiple loci of low or moderate effect on the disease phenotype [13]. Although common variants are mostly located outside of gene coding regions and often have no direct consequences on protein sequences, thousands of them have been associated with multiple traits and diseases by genome-wide association studies (GWAS) [14]. Though studies on expression quantitative trait loci (eQTLs) have broadly shown that non-coding variants can modulate gene expression [15], clear mechanistic models linking common variants and biological functions are still limited.

In contrast, deleterious variants with large effect sizes are comparatively rare within a population and often subjected to purifying selection [16]. With recent

advances in high-throughput next-generation sequencing (NGS) technologies, especially whole-exome sequencing (WES) and whole-genome sequencing (WGS), increasingly being applied in studies of both rare and common traits, a large number of rare variants is captured, that previously was not interrogated by genotyping platform designs [17]. This growth in availability of large datasets with rare variants enables researchers to further assess their impact on rare and common diseases, but also raises the question of how to functionally validate the impact of newly discovered variants [17].

1.2   Genetic architecture of human diseases

Genetic disorders can show monogenic (Mendelian disorders) or polygenic (multifactorial disorders) architectures. In monogenic (Mendelian) traits genetic alteration of a single gene or allele determines the phenotype. By contrast, polygenic traits result from the contributions of multiple, mostly common genomic variants in conjunction with influences of the physical and social environment. By now, variants in thousands of genes have been shown to cause rare, monogenic disorders [18], including severe developmental disorders manifesting from birth or early childhood [19]. However, the clinical phenotype in affected individuals can vary [20, 21]. Pleiotropy (influence of one gene on multiple seemingly unrelated phenotypic traits), incomplete penetrance (whether a disease-causing genotype displays the corresponding clinical phenotype or not in an affected individual), or variable expressivity (range of phenotypes that can be observed in affected individuals) can contribute to the variability of phenotypes caused by variants in the same gene or complicate the diagnosis [22]. Inherited common variants might further contribute to the range of the clinical presentation of rare human disorders caused by single deleterious, protein-coding variants [23]. Vice versa rare variants with weak effects can also play unique roles in the genetics of common diseases in humans [24, 25]. Rare variants with a strong effect can form monogenic subgroups of common diseases, which are typically polygenic, such as *BRCA1/2* or other candidate genes in monogenic forms of breast cancer [26].

1.2.1 Movement disorders

The term movement disorders first came up in 1968 when Stanley Fahn founded the first specialist clinic at the University of Pennsylvania [27]. Prior to this definition the term extrapyramidal disorders, originally introduced by Samuel Alexander Kinnier Wilson was often used to categorize neurological diseases secondary to disorders of the basal ganglia and its pathways [28]. However, several movement disorders such as dystonia weren't only associated with a basal ganglia pathology [27]. Despite further classification attempts based on the anatomic localization of the pathology [29] or clinical characteristics such as hypokinetic (decrease of movements) and hyperkinetic (excess of movements) movement disorders [27] the shortcomings of the classification became apparent with increasing knowledge of the phenomenology, pathology, and genetics [30].

Implementation of NGS in the last 20 years has provided a huge step forward in the identification of several genetic forms of hypokinetic and hyperkinetic movement disorders and the comprehension of pathological mechanisms [31]. NGS techniques, which are increasingly used in the clinical setting might hereby not only support the diagnostic procedure, but furthermore enable the identification of genetic variants in genes not suspected to be causative in the specific case under study, broadening the genotypic spectrum of the disease [32]. However, this also leads to a vast number of genetic variants of unknown significance, posing a major challenge to genetic counseling and making functional tests inevitable in case of the detection of a novel genetic variant to confirm its pathogenetic role.

1.3 Molecular approaches used in this work

Starting with the process of mapping and sequencing the human genome [33], new upcoming technologies have made it possible to obtain a large number of molecular measurements within a tissue or cell. These technologies can be applied to a biological system of interest to obtain a snapshot of the underlying biology. High-throughput measurement of such biological molecules enabling comprehensive analyses of biological components within a biological system are called omics [34].

Many research areas can be classified as omics. Examples include genomics, epigenomics, transcriptomics, proteomics, or metabolomics, corresponding to global analyses of genes, methylated DNA or modified histone proteins, RNA, proteins, or metabolites, respectively. Omics analyses can provide a comprehensive understanding of biological systems under study or help to associate omics-based molecular measurements with a clinical outcome of interest [34].

1.3.1 Genomics

Genomics focuses on the study of an organism's entire genetic material, including all of its genes and their interactions. It involves the sequencing, analysis, and interpretation of genetic data to understand genetic variations, gene functions, and the relationships between genes.

1.3.1.1 Whole exome sequencing

Whole exome sequencing (WES) is a widely used NGS method that only involves sequencing of the protein-coding regions of the genome. Inherited or acquired protein-coding variants represent the majority of disease-causing variants, accounting for upwards of 60% of all known causative genomic variants [35, 36].

WES is broadly used as a rapid and effective diagnostic tool. It can be employed at a prenatal period to detect fetal abnormalities [37] and postnatally following a phenotypic observation [38]. A significant benefit comes from the possibility to determine whether genetic abnormalities have been inherited from the parents, occurred spontaneously during gametogenesis (de novo), or in gestation (genetic mosaicism) [39]. Given an uncertain phenotype the success rate to provide a diagnosis ranges between 50% and 80% in newborn [37, 40] and between 25% and 50% for adult-onset phenotypes [41, 42].

1.3.1.2 Filtering workflow

To enable a systematic search for likely causative variants a stringent filtering scheme and rigorous clinical interpretation logic is required [43] (**Figure 2**).

As a first step, bioinformatic filtering was applied to all exome-generated variants using a validated in-house pipeline. Gene lists of interest were created via full-text queries in the "Online Mendelian Inheritance in Man" (OMIM) database [44] and inspection of the literature, considering the relevant phenotypic aspects of each index case under investigation. OMIM is a database of human genes and genetic disorders, providing phenotype-gene relationships including information on clinical features, inheritance, and information on molecular genetics in accordance to recent literature (e.g. *https://www.omim.org/entry/620427*; dystonia; NUP54; early onset with striatal lesions; autosomal recessive; impairment of nuclear pore function, based on Harrer et al. 2023 [45]) [44].

In the sets of predefined genes, a filtering process for protein-changing variants (including single-nucleotide variants, small insertions and deletions, and copy number variations) with an AF of 0.001 for variants in genes associated with autosomal-dominant and X-linked inheritance and 0.005 for variants in genes associated with autosomal-recessive inheritance was applied (in in-house controls and the Genome Aggregation Database (gnomAD) [46]). Cross-referenced databases like ClinVar [47], the Human Gene Mutation Database (HGMD) [48], and information already published on PubMed allowed the identification of known pathogenic alleles.

For further filtering of previously unreported variants various arguments supporting potential pathogenicity such as consistency of the variant type with the established mechanism of disease, variant location at an amino acid position previously implicated in disease, and variant location in a mutational hotspot or critical functional domain of the encoded protein were considered.

Known pathogenic variants and previously unreported variants with evidence supporting potential pathogenicity were subjected to co-segregation analyses whenever possible. The classification of all previously unreported variants was based on recommended diagnostic criteria of the American College of Medical Genetics and Genomics (ACMG) [49], providing international standards and guidelines for the interpretation of sequence variants.

Search for causative variants
in confirmed disease genes and genes
with previously described disease-association

Gene list of interest
based on the OMIM database, full-text queries
and systematic literature reviews

Protein changing variants
including loss-of-function, missense, indel-in variants,
and changes in gene dosage

| **Autosomal-dominant and X-linked inheritance** | **Autosomal-recessive inheritance** |

| **AF ≤ 0.001** in in-house controls and gnomAD | **AF ≤ 0.005** in in-house controls and gnomAD |

Known pathogenetic variants

Known pathogenetic variants

| Consistency with disease mechanism | Predicted deleterious effect | Consistency with disease mechanism | Predicted deleterious effect |
| Previously involved in disease | Located in mutation hotspot or crucial domain | Previously involved in disease | Located in mutation hotspot or crucial domain |

Segregation analyses
in all availabe family members (de novo/ co-segregation)

Classification of the variant
based on the ACMG criteria (pathogenic/ likely pathogenic)

Clinical interpretation
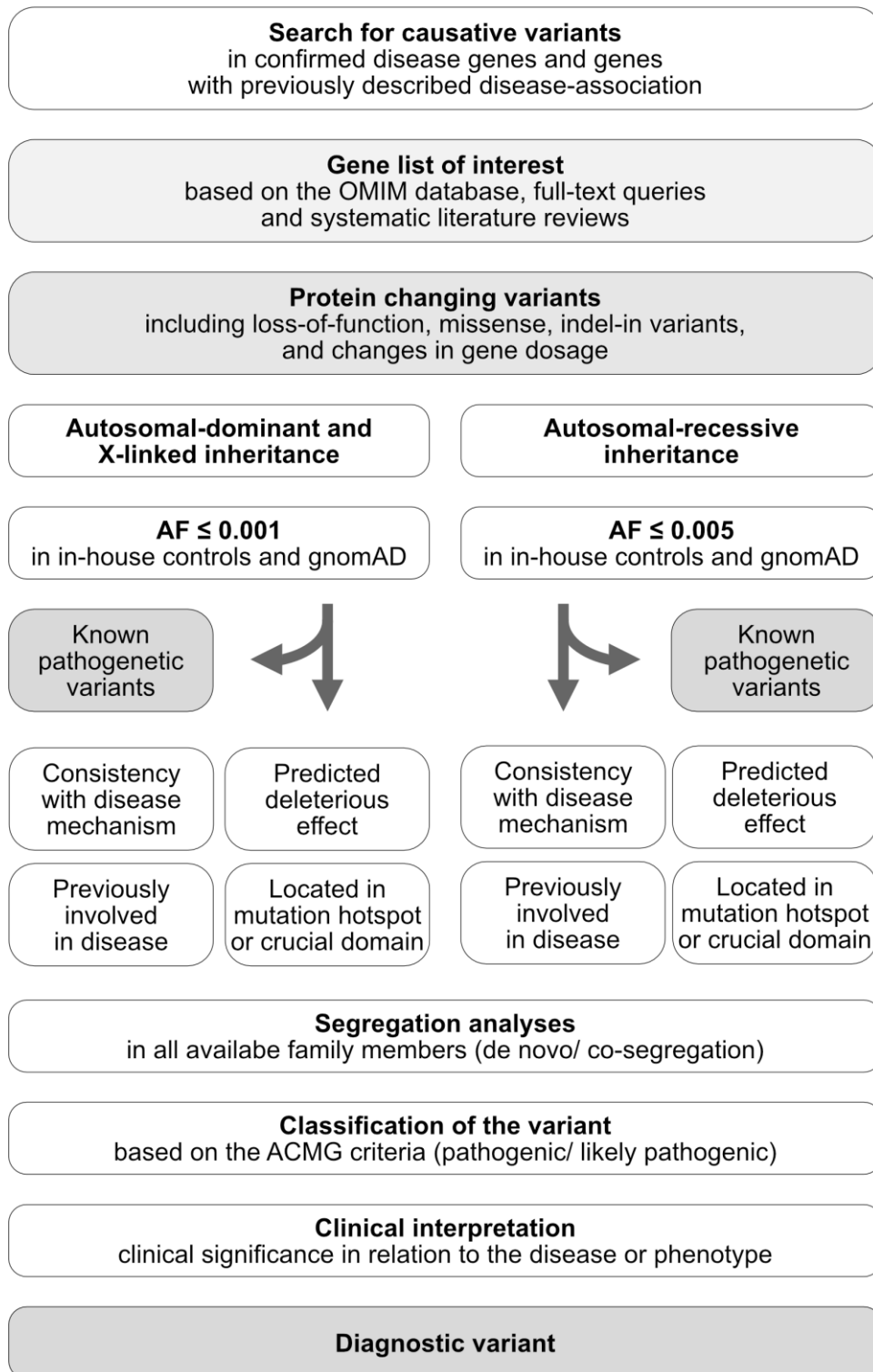clinical significance in relation to the disease or phenotype

Diagnostic variant

Figure 2: Identification of diagnostic variants in known disease genes

(modified from *Zech et al.* [43])

Priorization of candidate genes
in cases without diagnostic variant findings
after searching for variants in known disease genes

Variants in genes with fitting expression patterns
based on GTEx and The Human Protein Atlas

Autosomal-dominant and X-linked inheritance

Autosomal-recessive inheritance

Extremely rare variants, ≤ 1 carrier in in-house controls and gnomAD

AF ≤ 0.005 in in-house controls and gnomAD

LoF variants in genes with pLI ≥ 0.99

De-novo damaging (CADD ≥ 20) missense/indel-in variants in genes with z ≥ 2.5

LoF variants

De-novo damaging (CADD ≥ 20) missense/indel-in variants

Candidate genes

Genes (pLI ≥ 0.99) with de-novo LoF variant in ≥ 1 case

Genes (z ≥ 3.09) with de-novo missense/indel-in variants in ≥ 1 case

Genes with biallelic LoF variants in ≥ 1 cases

Genes with biallelic LoF + missense/indel-in variants in ≥ 2 cases

Genes (pLI ≥ 0.99) with LoF variants in ≥ 2 cases

Genes (z ≥ 2.5) with de-novo missense/indel-in variants in ≥ 2 cases

Genes with biallelic missense/indel-in variants in ≥ 2 cases

Promising candidate genes

Independent evidence
for a disease-relevant role of the candidate genes
(GeneMatcher and other databases)

Genes with de-novo or co-segregating variants
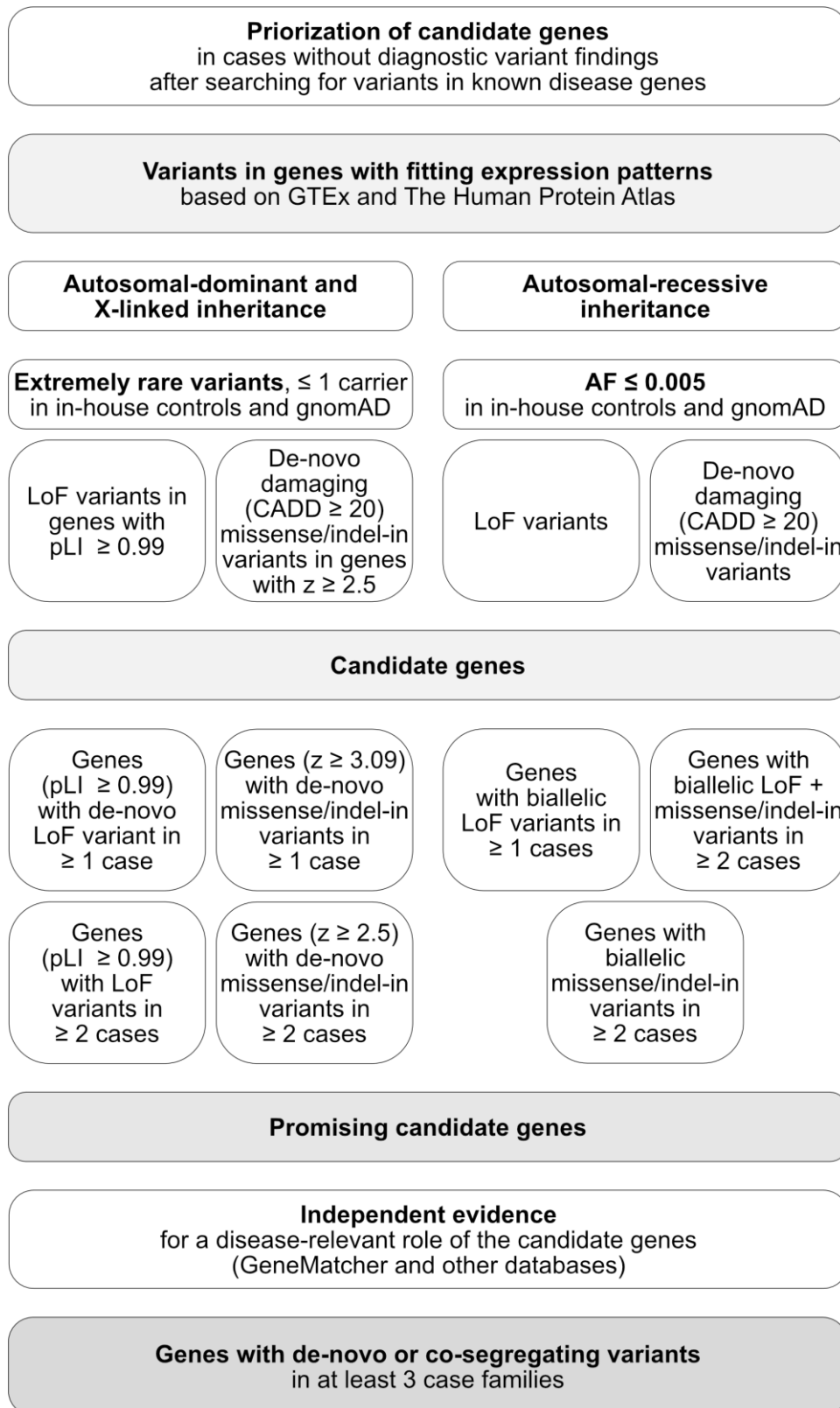in at least 3 case families

Figure 3: Identification of new candidate disease-causing genes

(modified from *Zech et al.* [43])

All variants remaining were manually reviewed to validate their clinical signifi-
cance in the context of the disease presentations under study. The conclusion on
the relationship of the variants in question and the observed phenotype was made
by a multidisciplinary team of geneticists, genetic scientists, neuropediatricians,
and neurologists.

In index cases without diagnostic variants, candidate causative variants and
genes were identified on a case-by-case basis based on their expression profile
in GTEx [50] and The Human Protein Atlas [51], allele frequency based on sug-
gested inheritance, suspected effects on the transcript of the gene using different
scores (Combined Annotation-Dependent Depletion (CADD) score used to
measure variant deleteriousness [52], probability of being loss-of-function intol-
erant [pLI] score, missense z-score), and independent evidence for a disease-
relevant role of the candidate genes using the matching tool GeneMatcher [53]
(**Figure 3**).

1.3.1.3   Functional validation of candidate variants by protein analyses

Adressing the challenge in molecular biology to understand the mapping of the
same genome to diverse phenotypes in different tissue types, development
stages and environmental conditions analyses on the transcriptomic [54] and pro-
teomic [55] expression levels can be promising tools. The set of all RNA tran-
scripts, including coding and non-coding, in an individual or a population of cells
is referred to as transcriptome, the total protein content of an organism as prote-
ome. However, not only the consideration of the levels on a large scale, but also
the specific elaboration of the effect of a potentially causal genetic variant on the
transcripts and protein structure of the encoded protein can be helpful in deci-
phering pathophysiological relationships. In this work we did not examine the en-
tire proteome but addressed specific aspects of selected proteins using various
targeted methods.

Protein structure prediction

The mRNA sequence defines the amino acid sequence (primary structure) of a
protein that drives the folding and intramolecular bonding of the linear amino acid
chain, determining its unique three-dimensional shape. Hydrogen bonding in

neighboring regions direct specific folding patterns, like alpha helices and beta sheets (secondary structure). The ensemble of formations and folds in a single linear chain of amino acids forms the tertiary, the interaction of multiple polypeptide chains or subunits the quaternary structure of a protein.

Through a large amount of different experimental designs the three-dimensional structures of around 100,000 unique proteins have been experimentally determined by now [56]. However, considering the billions of known protein sequences [57] and the large amount of time it still requires to experimentally determine a single protein structure, computational approaches are recently on the rise to enable large-scale structural protein analyses. The largest computational prediction tool to date is AlphaFold [58], using a neural network-based machine learning approach to predict the three-dimensional structure that a protein will adopt based on its amino acid sequence.

Protein expression and localization

Mass-spectrometry-based proteomic analyses, combining proteome experimentation and data analysis, offer broad information on protein composition, structure, expression, modification status, and the interactions and connections between proteins at an overall level [59]. With the substantial improvement of experimental technologies over the past decades, proteomics methods have been evolved from conventional methods, such as immunohistochemistry (IHC), immunofluorescence (IF) and Western blot (WB), to high-throughput methods such as tissue microarrays, protein pathway arrays and mass spectrometry [60]. Despite these advances, however, conventional methods such as WB, and IHC/IF staining are still applied frequently, offering the advantages of affordable, simple, but powerful targeted approaches in the determination of protein expression levels and their cellular localization [61].

To function properly every cell requires proteins to perform specific tasks in designated compartments. Each organelle in a cell is a specialized subunit within the cell fulfilling a specific function. The subcellular localization of a protein also determines its function as different organelles offer distinct environments containing a variety of physiological conditions and interaction partners. Mislocalization of proteins is often associated with cellular dysfunction and disease [62, 63].

Deploying the specific binding between an antibody and antigen the localization of specific proteins in cells and tissue can be detected and visualized using different microscopic techniques.

Protein - protein interactions

Critical biological processes that directly associate with our health like DNA replication, transcription, translation, and transmembrane signal transduction all rely on functional proteins, often being regulated through protein complexes, that are typically controlled via protein–protein interactions [64]. Aberrant interactions are associated with various diseases, including cancer, infectious, and neurodegenerative diseases [65].

There are several methods to investigate protein-protein interactions in cells, ranging from commonly-used co-immunoprecipitation [66] to more advanced fluorescence-based methods such as Fluorescence Resonance Energy Transfer [67]. Those methods, however, also have limitations and are dependent on genetic modification of the cells or cloning of genes, which are not feasible with patient samples. A suitable and sensitive assay to detect direct protein-protein interactions at endogenous levels of expression without genetic manipulation is the Proximity Ligation Assay (PLA) [68].

PLA permits the detection of the proximity of proteins in situ (distances < 40nm) at endogenous protein levels, where a close distance indicates high chances for functional interaction (**Figure 4**). The two proteins of interest are identified using specific antibodies against the target proteins, followed by secondary antibodies coupled with oligonucleotides (PLA probes) against the primary antibodies. After DNA amplification by polymerase chain reaction of the template created by ligation of hybridized connector oligonucleotides, the spots of proximity can be visualized by fluorescence microscopy using complementary detection oligos linked to fluorochromes.
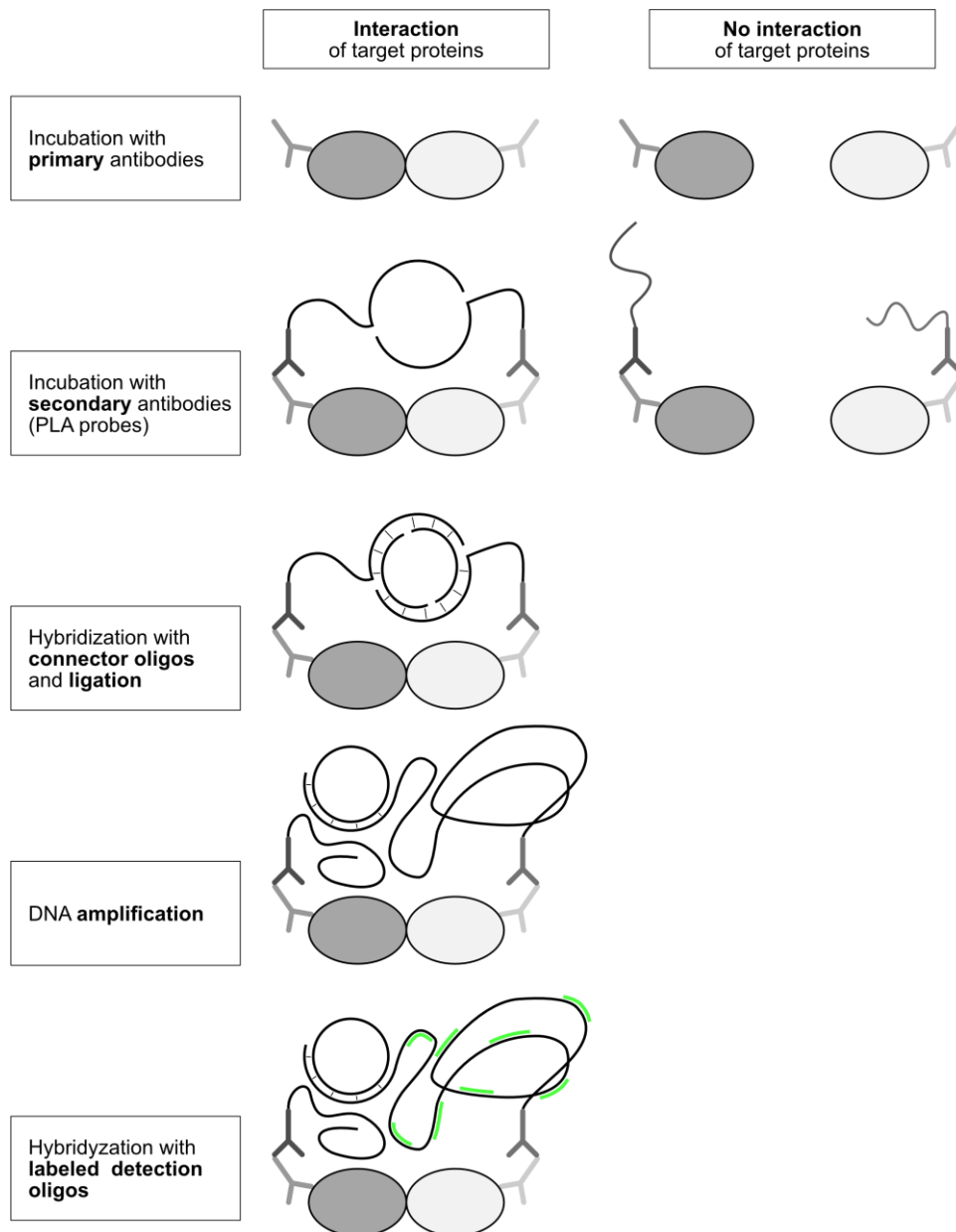
Figure 4: Schematic of the Proximity Ligation Assay

1.3.2 Epigenomics

Epigenetics is defined as changes in DNA that do not involve alterations to the underlying sequence [69]. These alterations include DNA methylations, histones, chromatin structure, and noncoding RNAs, which are modified throughout life in response to environmental and behavioral influences [70]. However, differentially expressed states can also be transmitted across cell divisions, once they are established, even in absence of the original signal [71, 72]. Epigenetic methylation patterns play an instructive part in transcriptional regulation by controlling the

accessibility of transcription factor binding sites at distal and proximal regulatory regions [73]. By maintaining these distinct gene expression profiles tissues can differentiate in function despite possessing identical or near-identical genomic DNA sequences [74]. The methylation state at the more than 30 million CpG sites across the human genome is involved in cellular differentiation throughout development and often maintained into adulthood [75].

Technical advances in detection of the relative proportion of methyl-groups (CH3) present at hundreds of thousands of CpG sites have enabled studies that perform genome-level, locus-agnostic examinations of DNA methylation levels, so called epigenome-wide association studies (EWAS). In contrast to static genetic estimates, EWAS are likely to capture tissue- and time-specific information distinct from that of static genetic estimates [76], which could provide further understanding how environmental factors may contribute the development of diseases [77].

1.3.2.1 Epigenome-wide association studies

The most characterized epigenetic marker is DNA methylation, a reversible modification to DNA involving the covalent addition of CH3 to the fifth carbon position of cytosine by DNA methyltransferases. In mammals, DNA methylation occurs mainly at cytosine-guanine dinucleotides (CpG sites). CpG sites are specific sequences of DNA bases where cytosines are followed by guanines. The 'p' indicates the phosphate bond separating the two residues in 5′ to 3′ direction [76, 78].

DNA methylation is primarily detected through the conversion of unmethylated cytosines with sodium bisulphite to uracil. Methylated and unmethylated cytosines can be distinguished using array-based or sequencing-based technologies. In each individual DNA molecule, cytosines are either methylated or not. Bulk tissue analyses are therefore averages across multiple cells [78]. High-throughput methylation arrays only capture a small proportion (~ 2%) of all possible methylation sites, however, current commercial arrays allow to get information of up to 96% of coding genes, known enhancer and promoter elements [79].

EWAS are used to examine the association between a large number of epigenetic variables at CpG sites and the phenotype of interest [76].

In contrast to the DNA sequence, which is relatively stable across cell types throughout lifetime, methylation patterns vary between cell types and are frequently modified by individual and environmental factors. Cell type-specific methylation patterns mark the cell-type lineages, allowing to determine the cell type of origin [80, 81]. At some genomic locations, DNA polymorphisms are associated with methylation, so called methylation quantitative trait loci [82]. Some of these associations are likely to imply a direct causal relationship between DNA variation and methylation patterns [83]. Furthermore, some CpG sites show a strong association with age [84], or other individual or environmental factors, like smoking [85].

1.3.2.2 Methylation risk scores

Risk scores describe weighted sums of risks for a phenotype transmitted by genetic, or epigenetic polymorphisms within an individual, where the weights used are coefficients from the respective association studies.

Trait prediction using methylation data requires at least two independent data sets for discovery and validation. The discovery set is used to identify CpG sites associated with the trait of interest, resulting in a list of probes and weights. To identify the CpG sites that most strongly associate with the outcome of interest a set of possible input variables is often selected based on a minimum p-value threshold in the EWAS results [76]. An additional, independent training set can help to further optimize the weighting and selection of probes included in the risk score [78].

To ensure a generalizable performance, the prediction model needs to be applied to an independent validation data set [78]. An association in the independent data, however, doesn't necessarily illustrate functional or causal mechanisms [76]. To summarize the performance of a model the sensitivity (true positive rate) and $1 -$ specificity (false-positive rate) are calculated and can be visualized as a curve in a plot of sensitivity versus $1-$ specificity, called a receiver operating

characteristic (ROC) curve (**Figure 5**). The area under the curve (AUC) provides a summary of the overall performance. A predictor that generates purely random predictions will result in a ROC curve along the diagonal and have an AUC roughly equal to 0.5, a perfect predictor will produce a curve with an AUC of 1. Informative but imperfect predictors will produce curves somewhere between these two extremes with AUC values ranking between 0.5 and 1 [76].
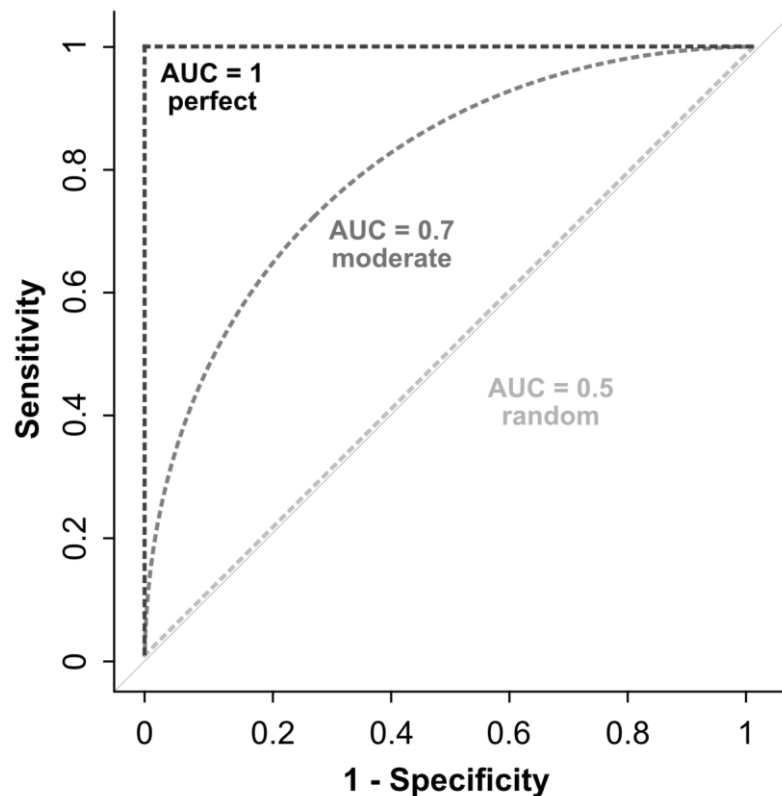

Figure 5: Visualization of risk score performance

(modified from *Yousefi et al.* [76])

1.3.2.3 Epigenetic age

Epigenetic clocks are mathematically derived age estimators based on DNA methylation or other epigenetic measurements [86]. The phenotypical age of a DNA source may be significantly higher (accelerated aging) or significantly lower (decelerated aging) than the chronological age [87]. Considering the variation of methylation patterns between different tissues, some epigenetic clocks have been optimized for specific tissue types and applications [88], like Shireby's

cortical clock for the assessment of epigenetic age in the human cortex and the identification of phenotypes associated with biological ageing in the brain [89].

1.3.2.4 Gene set enrichment analyses

Gene set enrichment analyses are widely used to provide insight into high-throughput gene expression data with many gene set analysis methods available [90]. Cellular processes are often associated with changes in the expression patterns of groups of genes that share biological functions or attributes, making a change in a group of these genes more biologically reliable and interpretable than a change in a single gene. Information about these sets of genes is available through public online databases such as KEGG [91] or FUMA [92]. Incorporating this information in the data analysis may support the identification of underlying biological processes or functions [90].

1.3.2.5 Differentially methylated regions

Differentially methylated regions (DMRs) are genomic regions with different methylation statuses among multiple samples (tissues, cells, individuals, etc.) and are regarded as possible functional regions involved in transcriptional regulation [93]. Identification of DMRs among multiple tissues can provide a comprehensive survey of differences in the methylation patterns. ~ 20% of the autosomal CpGs show a dynamic regulation within a normal developmental, often co-localizing with gene regulatory elements, particularly enhancers and transcription-factor-binding sites [94]. Although individual CpGs might have weak associations, DMRs can have stronger associations with the traits of interest [94].

1.4   Disorders studied in this work

1.4.1 Dystonia

Dystonia describes a class of movement disorders defined by manifestation of involuntary repetitive muscle contractions and unnatural postures, associated with both additional motor and non-motor features [95]. It can be the manifesting neurological sign of many disorders, either in isolation (isolated dystonia) or with

additional signs (combined dystonia). Most cases of isolated dystonia are caused by genetic alterations or are of unknown aetiology. Clinically, isolated dystonia can differ in manifestation, severity, and age at onset. Depending on the affected regions dystonia can be subclassified as focal (one region of the body is affected), segmental (adjacent regions of the body are affected), multifocal (non-contiguous regions of the body are affected) or generalized (several regions of the body are affected, including the trunk) [95]. The diagnosis of dystonia is largely based on clinical signs, and the classification and aetiological definition of this disorder remain a challenge, underlining the importance of establishing reliable biomarkers based on mechanistical insights [96].

1.4.1.1 Epidemiology

The overall prevalence of dystonia in the general population remains unclear [96]. Considering patients in movement disorder clinics individuals affected with dystonia represent around 20% [97, 98].

Isolated forms of dystonia often show a characteristic pattern regarding age at onset, sex, and anatomical distribution. Isolated generalized dystonia for example typically manifests in childhood or adolescence and has no sex predilection. Conversely, focal limb dystonia often manifests in the fourth decade of life and has a balanced female to male ratio, except for musician's dystonia [99]. The most common form of adult-onset dystonia is cervical dystonia, which affects women twice as often as men and has an age at onset in the fourth or fifth decade of life [99, 100].

1.4.1.2 Clinical presentation and diagnosis

The International Movement Disorders Society defines dystonia as a disorder with sustained or intermittent muscle contractions causing abnormal, often repetitive movements, postures, or both [95]. Tremor (rhythmic, though often inconstant patterns of movement that may precede the onset of dystonic posturing) is included as part of the motor phenomenology of dystonia [95]. Dystonic movement can spread to other body regions during movement of the primarily affected body part [101]. Stress and anxiety can worsen the symptoms. Susceptibility to

those can arise as part of the non-motor syndrome of dystonia, including abnormalities in sensory and perceptual functions, neuropsychiatric, cognitive and sleep domains [102].

Clinically, dystonia can be classified by age of onset, body distribution and temporal pattern (**Table 1**), including its relation to voluntary actions or triggers (**Table 2**). Forms of isolated idiopathic or genetic dystonia follow particular patterns with regard to age of onset, body distribution, and sex preponderance. Early onset dystonia (symptom presentation in the first three decades of life) for example typically manifests in the legs with subsequent generalization. Generalization occurs in 70% of patients, but the cranial region is often spared [103].

Table 1: Dystonia described by body distribution
(modified from *Balint et al.* [96])

Dystonia	Affected regions
Generalized dystonia	Trunk and at least two other body sites
Cervical dystonia	Neck, leading to abnormal postures of the head
Torticollis	Head turning to one side
Laterocollis	Tilt to the side
Retrocollis	Neck extension
Anterocollis	Neck flexion
Cranial dystonia	Face or voice
Laryngeal dystonia	Vocal cords, leading to a strangled, coarse voice with: • variations in pitch (adductor type) • whispering, breathy voice (abductor type)
Blepharospasm	Eye closing spasms
Oromandibular dystonia	Mouth and/or jaw, leading to involuntary perioral movements, mouth opening or mouth closing
Hemidystonia	Only one side of the body

Table 2: Dystonia characterized by the relation to voluntary actions or trigger
(modified from *Balint et al.* [96])

Dystonia	Action or trigger
Writer's cramp	Focal, affecting the hand and/or the forearm, manifesting as abnormal posturing when patients attempt to write, which increases as writing continues
Musician's dystonia	Manifests in the body parts involved when individuals play a musical instrument
Paroxysmal dystonia	Occurs only intermittently with certain triggers
Exercise-induced dystonia	Manifests after prolonged exercise as self-limiting episodes of dystonia in the exercised limb; the most typical manifestation is exercise-induced foot dystonia, which results in in-turning of feet after prolonged walking
Paroxysmal kinesigenic dyskinesia	Manifests as brief self-limiting episodes of dystonic posturing, triggered by sudden movements

1.4.1.3 Genetics

Since the introduction of powerful high-throughput sequencing platforms the number of gene-disease and variant-phenotype associations in individuals with dystonic disorders increased exponentially. In a recent large-scale whole-exome sequencing study of 708 index patients with diverse types of dystonic disorders 160 different monogenic, disease-causing variants in 78 distinct genes were identified (accounting for an overall diagnostic rate of ~19%) [43]. Novel implications of causative genetic defects of dystonic conditions include genes not previously linked to any human Mendelian disorders, genes which have originally been associated with other disease traits, and genes that are linked to more complex neurological or developmental phenotypes in which dystonia is one feature of the associated clinical presentation [104].

In contrast to the fast pace of new discoveries on the genetic level, the translation of these findings into biologically meaningful understanding of dystonia pathophysiology lags behind [105]. These mechanistical insights, however, are crucial for a better classification and stratification of patients, and for the perspective development of targeted clinical trials and directed treatments.

1.4.1.4   Pathophysiology

It is still debated which anatomical structures are involved in dystonia. Based on dystonia patients with discrete focal lesion in the basal ganglia [106] dystonia was initially assumed to be caused by an imbalance of the direct (facilitating motor activity) and indirect (reducing motor activity) pathways in the basal ganglia. Follow-up studies refined this assumption, suggesting impairments in a motor network including basal ganglia, cerebellum, thalamus, and cortex [107] in line with imaging studies outlining alterations in the volume of basal ganglia, cerebellum, and cortex in dystonia patients [108] and alterations in the connectivity of cerebello-thalamo-cortical circuits [109].

Due to the characteristic excess of muscle activity and co-contraction (simultaneous activation of muscles on opposite sides of a joint) several pathophysiological studies concentrated on a reduced excitability of inhibitory circuits in the spinal cord, brainstem and cortex which might predispose individuals to dystonia in the presence of other triggering factors [96, 110]. Alterations in synaptic plasticity in different brain regions could further promote the development of dystonia. Animal models have demonstrated persistent strengthening of synapses that leads to a long-lasting increase in signal transmission between neurons (increased long-term potentiation) and reduced or lost activity-dependent reduction in the efficacy of neuronal synapses (reduced longterm depression) in cortico-striatal projections [111, 112]. This could lead to a loss of inhibition of central circuits and promote excessive muscle contractions because of unwanted associations between activity in remote muscles and muscles directly involved in the task.

Considering molecular biological pathways in dystonia recent studies worked out convergences between seemingly unrelated dystonia-associated genes, contributing to the implementation of shared pathways including dopamine signaling, brain calcification and heavy metal accumulation, dysfunction of energy and endoplasmic reticulum homeostasis, calcium regulation, and gene-expression control [113, 114].

Dopamine

Defects of dopamine-linked neurotransmission have been associated with various disease traits in animals and human subjects including a wide range of movement disorders encompassing hypokinetic syndromes (Parkinson disease, atypical parkinsonism) as well as primarily hyperkinetic conditions such as chorea and dystonia [115]. One of the first monogenic etiologies for dystonia was a consequence of dopamine signaling abnormalities, identified by the Japanese neurologist Masaya Segawa as dopa-responsive dystonia in the 1970s [116]. The causal variant was located in the gene encoding GTP cyclohydrolase I (GCH1), responsible for the rate-limiting step of tetrahydrobiopterin biosynthesis, an essential cofactor for production of levodopa [116].

Overall molecular defects and their associated regional dopaminergic changes have been linked to presynaptic synthesis and metabolism of dopamine, dopamine-related postsynaptic signal transduction and second messenger-induced effect response, development, structural maintenance, and survival of dopaminergic neurons [117].

Brain calcification and heavy metal

Metals such as manganese, copper, and iron are essential for enzymatic reactions in various physiological processes, such as regulation of gene expression, neurotransmitter generation, and electron transport [117]. Deficiencies and accumulation of these metals lead to several harmful intracellular consequences, including mitochondrial dysfunction, calcium dyshomeostasis, buildup of damaged molecules, compromised DNA repair, reduction in neurogenesis, and impaired energy metabolism [118]. Resulting impairments of neurodevelopment, neurotransmission, and neurodegeneration often manifest with movement disorders such as dystonia [119]. Known genetic causes of dystonia associated with toxic metal deposition in the central nervous system include manganese transporter disorders, Wilson's disease (accumulation of copper), and the group of syndromes of neurodegeneration with brain iron accumulation [114].

Energy homeostasis

Mitochondrial dysfunction is an important cause of several movement disorders including various types of dystonia [120]. Impairments in oxidative

phosphorylation especially affect the energy-dependent muscle tissue and the central nervous system [121]. Due to this susceptibility mitochondrial disorders almost always manifest with neurological symptoms, including neurodevelopmental disturbances, intellectual disability, seizures, stroke-like episodes, migraine, dementia, ataxia, and many hyper-/and hypo-kinetic movement abnormalities [121].

An example of a mitochondrial disease frequently presenting with dystonia is Leigh syndrome, a progressive neurodegenerative disease with bilateral, symmetrical basal ganglia lesions. Around 60% of affected individuals manifest dystonia, usually in multifocal or generalized forms [122]. Implicated genes such as *MT-ATP6* (disruption of the oxidative phosphorylation subunit V), *SURF1* (impairment of cytochrome c assembly), and *PDHA1* (altering pyruvate metabolism) directly or indirectly alter the mitochondrial function [123].

Calcium

Calcium is a key intracellular second messenger controlling various functions of neurons [124]. Regulatory components include voltage-dependent ion channels in the cellular membrane, channels in the endoplasmic reticulum, cytosolic binding proteins, signaling molecules, transport proteins from cytosol to the extracellular space, and various additional up- and downstream effectors [125]. Dysregulation of calcium homeostasis in the central nervous system can result in multiple diseases ranging from cognitive impairment and seizures to variable movement-disorder phenotypes including dystonia [125]. Dystonia related monogenic, causal variants involve voltage-dependent calcium channels (e.g. CACNA1A, subunit of the presynaptic P/Q-type (Cav2.1) channel [126]), cytosolic proteins (e.g. HPCA, a calcium-sensor protein regulating intracellular homeostasis [127]), calcium exporters (e.g. ATP2B3, a P-type ATPase involved in calcium clearance [128]), calcium-dependent kinases (e.g. CAMK4, a Ca2+/calmodulin-dependent protein kinase [129]), and gene-transcription regulators (e.g. CAMTA1, a calcium-responsive, brain-specific transcription regulator [130]).

Gene-expression control

Continuously expanding the number of identified variants in developmental control genes, aberrant gene-expression regulation seems to play an important role

in dystonia pathogenesis [131]. Involved genes are often critical for cell-fate specification and motor-circuitry establishment during early brain development, leading to more complex developmental syndromes in which dystonia can be a more or less prominent clinical trait [131, 132].

Examples of disorders of transcriptional deregulation with a predominant dystonic phenotype are KMT2B- [133] and THAP1-related diseases [134]. *KMT2B* codes for the lysine-specific histone methyltransferase 2B that mediates the di- and tri-methylation of histone 3, lysine 4 (H3K4) [135]. Other KMT-type histone methyl-transferase-associated disorders (e.g., KMT2A-, KMT2C-, and KMT2D- related neurodevelopmental syndromes) are dominated by developmental-disorder manifestations without dystonia [136]. Interestingly, pathogenic *KMT2B* variants generate characteristic, non-random DNA hyper-methylation patterns in affected cells that selectively target regulatory domains that promote gene expression [137].

THAP1 encodes for the zinc finger-containing transcription factor THAP domain-containing protein 1. Impairments of THAP1 might lead to aberrant cytosolic localization of the encoded protein, impairing the interaction with nuclear DNA and resulting in transcriptional dysregulation [138]. Both *THAP1* and *KMT2B* have their highest expression in the cerebellum [96].

Endoplasmic reticulum homeostasis

In the common *TOR1A* associated dystonia disease causing mutations often result in mislocalization of the encoded protein Torsin 1A from the endoplasmic reticulum to the nuclear envelope, indicating abnormal function of these subcellular structures [139]. Torsin 1A deficient animal models confirmed defects in the homeostasis of the endoplasmic reticulum, resulting in enhanced stress sensitivity [140] and abnormalities in folding, assembly and trafficking of proteins [141]. However, it isn't fully understood how TOR1A-induced impairments of the endoplasmic reticulum translate into abnormal neuronal activity. Though further elaboration is needed, the translation initiation factor eIF2α could be a critical mediator of stress response and promote synaptic plasticity [142].

In conclusion, several biological pathways including endoplasmic reticulum and energy homeostasis, synaptic plasticity, neurodevelopment, neurotransmission, and transcriptional regulation might promote a disbalance of inhibition and excitability in several neuronal circuits, leading to an increased facilitation of movement (**Figure 6**).
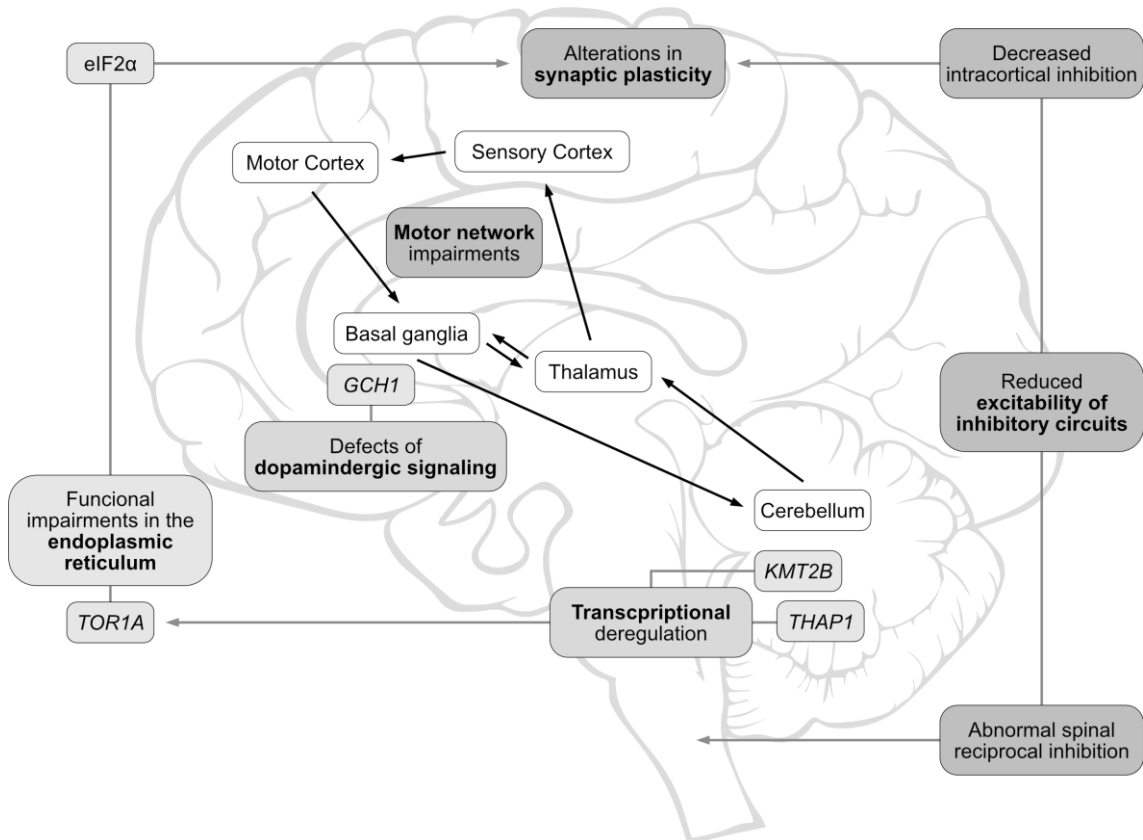

Figure 6: Possible pathogenetic model of dystonia

(modified from *Balint et al.* [96])

1.4.2 Restless legs syndrome (RLS)

RLS is a common, sleep-related sensorimotor disorder with an unclear pathophysiology, initially described by Sir Thomas Willis in the seventeenth century as an akathisiac psychiatric disorder [143]. Many years later the neurological nature of RLS was assessed by the Swedish neurologist Karl-Axel Ekbom with an extensive and detailed clinical connotation [144].

Despite the high prevalence of RLS in the general population, symptoms vary considerably regarding the frequency (from occurring less than once per year to daily) and severity (from mildly irritating to disabling with severe effects on sleep, mood, and overall quality of life) [145]. However, the awareness of RLS is comparatively low, with restricted recognition among physicians, and difficulties for patients in describing their symptoms and finding a specialist to be referred to. This often leads to frequent misdiagnosis, and severely delayed diagnosis and treatment, despite RLS being a lifelong disorder with severe impairments of the overall quality of life [146].

1.4.2.1   Epidemiology

Based on the methods of diagnosis, country, age, sex, comorbidities, differences in health behaviors, genetic risks, disease awareness and perception the prevalence of RLS varies considerably [147, 148]. While in most Asian populations the prevalence varies between 1% and 3% [149, 150], it ranges from 5% to 13% in most population-based studies of European ancestry [147]. Women are affected more frequently than men, with a 30-50% higher prevalence in the European population [151]. One in five pregnant women develops RLS, with a symptom peak in the third trimester [152].

RLS has been connected to a growing number of comorbid conditions, including renal failure, iron deficiency and pregnancy as major, and hypoxic conditions, diabetes mellitus, obesity, cardiovascular, rheumatological, psychological and neurological disorders as minor associations [153, 154]. Whether RLS or the comorbid condition had begun first, however, remains to be elucidated.

The age at onset of RLS has a bimodal distribution. One peak at around 20 years of age and a second peak at around 40 years of age, though the diagnosis of many patients is delayed up to the sixth decade of age [155, 156]. Individuals with early age at onset more often report a positive family history of RLS and a slowly progressing course of the disorder compared to patients with late-onset RLS [155, 156].

1.4.2.2   Clinical presentation and diagnosis

Due to the lack of a reliable biological diagnostic marker for RLS clinical evalua-
tion and assessment of the medical history are crucial for the diagnosis, distin-
guishing symptomatic from idiopathic forms, and disease management [157]. The
main diagnostic criteria are the so called RLS minimal diagnostic criteria by the
International RLS Study Group (IRLSSG) [156], consisting of five essential diag-
nostic criteria (**Table 3**).

Table 3: Diagnostic criteria by the International RLS Study Group (IRLSSG)
(modified from *Allen et al.* [156])

Essential Diagnostic Criteria (all must be met):	
1	An urge to move the legs usually accompanied by uncomfortable and unpleasant sensations in the legs (or sometimes other body parts).
2	The urge to move the legs and any accompanying unpleasant sensations begin or worsen during periods of rest or inactivity such as lying down or sitting.
3	The urge to move the legs and any accompanying unpleasant sensations are partially or totally relieved by movement, such as walking or stretching, at least if the activity continues.
4	The urge to move the legs and any accompanying unpleasant sensations during rest or inactivity only occur at night or are worse in the evening or night than during the day.
5	The occurrence of the above features is not solely accounted for as symptoms primary to another medical or a behavioral condition (such as myalgia, venous stasis, leg oedema, arthritis, leg cramps, positional discomfort, or habitual foot tapping).
Supportive Criteria	
1	Periodic limb movements: presence of periodic leg movements in sleep or resting at rates or intensity greater than expected for age or medical/medication status.
2	Dopaminergic treatment response: reduction in symptoms at least initially with dopaminergic treatment.
3	Family history of RLS among first-degree relatives.
4	Lack of profound daytime sleepiness.

The majority of RLS patients (around 60-70%) experiences sleep disruption, in-
cluding difficulties falling asleep, increased number of awakenings and reduced
total sleep time [158], with symptoms mainly occurring during the first part of the
night. Other signs of chronic sleep deprivation including fatigue, difficulties to

concentrate and depressive symptoms can occur, but the level of daytime sleepiness often seems disproportionate to the grade of night-time sleep impairment [159].

The severity of RLS can be addressed using the International RLS (IRLS) rating scale, consisting of ten questions in a face-to-face interview. It allows a classification in mild, moderate, severe and very severe [160].

1.4.2.3   Genetics

Genetic predisposition is an important contributor to RLS [157]. 20-60% of all RLS patients, up to 60% of individuals with idiopathic RLS, and up to 80% of monozygotic twins have a positive family history [161-163]. Using genome-wide association studies (GWAS), common genetic risk variants within 19 risk loci could be identified with an estimated contribution by common variants to RLS heritability of 19.6% [164]. Bioinformatic pathway and gene set enrichment analyses across the 19 risk loci identified processes of neurogenesis, axon guidance and synaptogenesis, which are crucial for the correct building and maintenance of functional neuronal circuitry [164].

Trying to identify putative causal genes for functional follow-up studies a transcriptome-wide association study integrating RLS GWAS summary statistics and publicly available genome-wide gene expression data yielded five candidate genes in known loci as well as six novel genes [165]. Testing for a significant burden of rare coding variants in 84 RLS candidate genes revealed 14 significant genes with a differential burden of low-frequency and rare variants, five of them not previously associated with RLS [166]. These genetic findings underline the involvement of both common and rare genetic variants in RLS susceptibility.

However, despite being a condition in which genetic background, environmental factors and gene-environment interactions predispose disease development and affect expression of the full clinical phenotype [153] large-scale EWAS in RLS are largely missing. The single previously published EWAS involved samples of only 15 cases and 15 controls, deriving an epigenetic diagnostic score based on DNA methylation at 49 CpG sites with a surprisingly high accuracy (area under the

curve [AUC]) of 87.5% [167]. However, also predicting iron deficiency anemia with a similar accuracy (83%) this diagnostic score may have focused on symptomatic RLS resulting from iron deficiency.

1.4.2.4   Pathophysiology

Though the exact location and extent of the anatomical substrate of RLS is not yet known, relevant brain regions implicated include the basal ganglia, thalamus, sensory-motor cortex, and spinal cord. Imaging studies indicated subtle white-matter changes, particularly in patients with severe RLS, in structures involved in sensory or motor control and sensorimotor integration [168]. Furthermore, decreased functional connectivity in the dopaminergic network has been demonstrated, leading to sensorimotor processing dysfunction involving nigrostriatal, mesolimbic, and mesocortical connections, accompanied by an altered functional connectivity in the thalamus [169].

The exact pathophysiology of RLS is not fully understood yet, but several mechanisms involving brain iron deficiency, dopaminergic dysfunction and alterations in neurotransmitter systems have been proposed.

Brain iron

Brain iron deficiency has been associated with an increased prevalence of RLS symptoms [170]. Despite hints of a decreased iron content of the substantia nigra [171, 172], the main iron store in the brain and the cerebrospinal fluid [173], systemic iron levels often remain unchanged [173]. Post-mortem protein expression studies identified alterations in iron management and regulation pathways in the brain, indicating altered brain iron acquisition [174]. However, it remains questionable whether all RLS patients show brain iron deficiency.

Dopamine

Hints for the role of dopaminergic pathology in RLS come from the initially positive pharmacological responses of RLS patients to the treatment with dopaminergic agonists [175]. Neuropathological studies have shown a decrease in dopamine 2 receptor expression in the putamen of RLS patients and an increase in tyrosine hydroxylase, which converts tyrosine to levodopa, a precursor of dopamine [176].

Despite these observations and the clinical suggestions of a "hypo-dopaminergic" state, however, subsequent research found increased synthesis and release of dopamine, in line with a concept of a presynaptic "hyper-dopaminergic" state [177]. The mechanism of action of dopaminergic medications in RLS is therefore unclear, though it may be partially explained by the circadian aspect of dopaminergic physiology [178]. The downregulation of postsynaptic dopamine 2 receptors may lead to a night-time dopaminergic deficit contributing to the development of symptoms at night and augmentation under long time dopaminergic medication [179, 180].

GABA and glutamate

Recently several studies have supported a role of altered glutamatergic and GABAergic balance in RLS [181, 182]. NMDA receptor inhibitors such as ketamine [183] and methadone [184] have shown to be effective for RLS. α2δ ligands such as pregabalin or gabapentin which are not only considered efficacious but also show less long-term treatment complications [185] may also derive their benefits from their actions on the glutamatergic system [186]. Iron deficiency might promote hypersensitivity of the glutamatergic cortico-striatal terminals [187, 188].

Adenosine

Downregulation of adenosine A1 receptors and an upregulation of striatal A2A receptors in relation with brain iron deficiency led to increased sensitivity of cortico-striatal glutamatergic terminals in rodents [189, 190]. In addition to effects on glutamate release, A1 and A2A receptors interact with dopamine D1 and D2 receptors forming different heteromers, that are highly expressed on the striatonigral and striatopallidal neurons, playing a role in modulation of the adenosine-dopamine-glutamate balance in the striatum [191]. In a crossover, placebo-controlled study the adenosine reuptake inhibitor dipyridamole improved sensory symptoms, periodic limb movements and sleep disturbances in RLS patients [192].
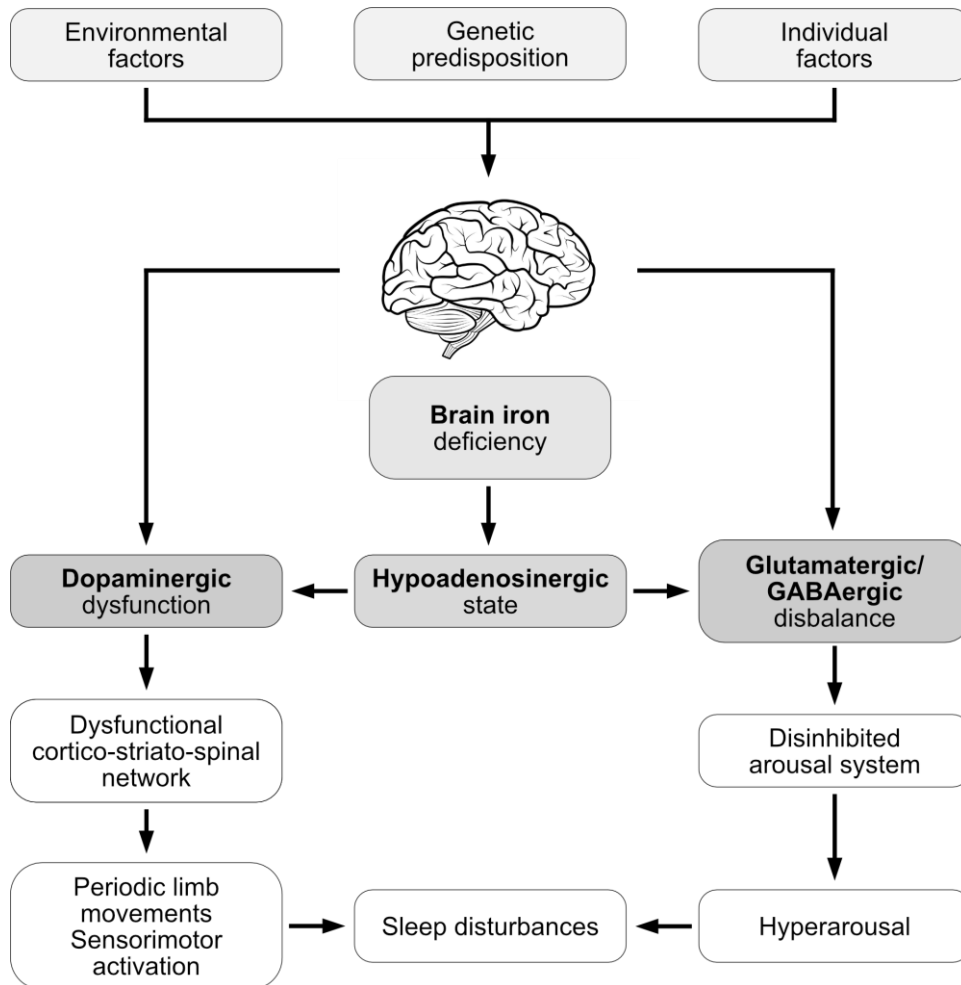
Figure 7: Possible pathogenetic model of RLS

(modified from *Manconi et al.* [157])

Overall, RLS might arise at different levels in the central nervous system influenced by genetic vulnerability, environmental, and individual factors, eventually resulting in modified, possibly temporary and circadian, neuronal networks [181]. While altered dopaminergic neurotransmission is considered as an essential mechanism underlying both akathisia and periodic limb movements, glutamatergic and GABAergic disbalances enhance alterations in the activity of neuronal circuits and promote a hyperarousal state [182]. The link between brain iron deficiency and the observed dopaminergic, GABAergic and glutamatergic disbalances could be a hypoadenosinergic state, modulating the neurotransmitter balance [191] (**Figure 7**).

1.5   Aims

Contributing to a better understanding of movement disorders is a challenging approach considering the complex nature of the phenomenology and pathophysiology of these disorders. The aim of this study was to employ diverse approaches on different omics levels to deepen our understanding of the molecular mechanisms involved in movement disorders.

Dystonia is often caused by monogenic variants with strong effects, directly affecting the function of encoded proteins (NUP54 [45], EIF4A2 [193]). This enabled targeted functional analyses on the protein level, aiming for information on altered expression profiles and subcellular locations, impaired protein-protein interactions, and resulting functional consequences.

In contrast, RLS, a common polygenic disease, arises from multiple predisposing genetic and non-genetic factors. Due to the comparably weak, but additive effects of multiple genetic variants to the disease, large case-control studies are needed to determine the underlying effect of those variants. While multiple GWAS already identified several common risk variants on the genomic level, epigenetic analyses and biomarkers are largely missing opposed to other neurobehavioral disorders. Therefore, we aimed to develop a biomarker based on DNA methylation in blood and to examine the methylation patterns in brain tissues for dissecting RLS pathophysiology [194].

2.   Publications

2.1   Recessive *NUP54* Variants Underlie Early-Onset Dystonia with Striatal Lesions

Contributions

The concept and design of the study were developed by me together with Audrey Schalk, Masaru Shimura, Juliane Winkelmann and Michael Zech. I planned the experimental procedure based on the pathophysiological hypothesis, coordinated the cultivation of patient fibroblasts together with Ivana Dzinovic, conducted the protein extraction, Western blots, and indirect immunofluorescent stainings together with Volker Kittke, and performed the data analyses of protein expression in Western blots. Drafting of the manuscript and figures was done by me, Ivo Melčák (in silico structural analysis of mutated NUP54) and Michael Zech.
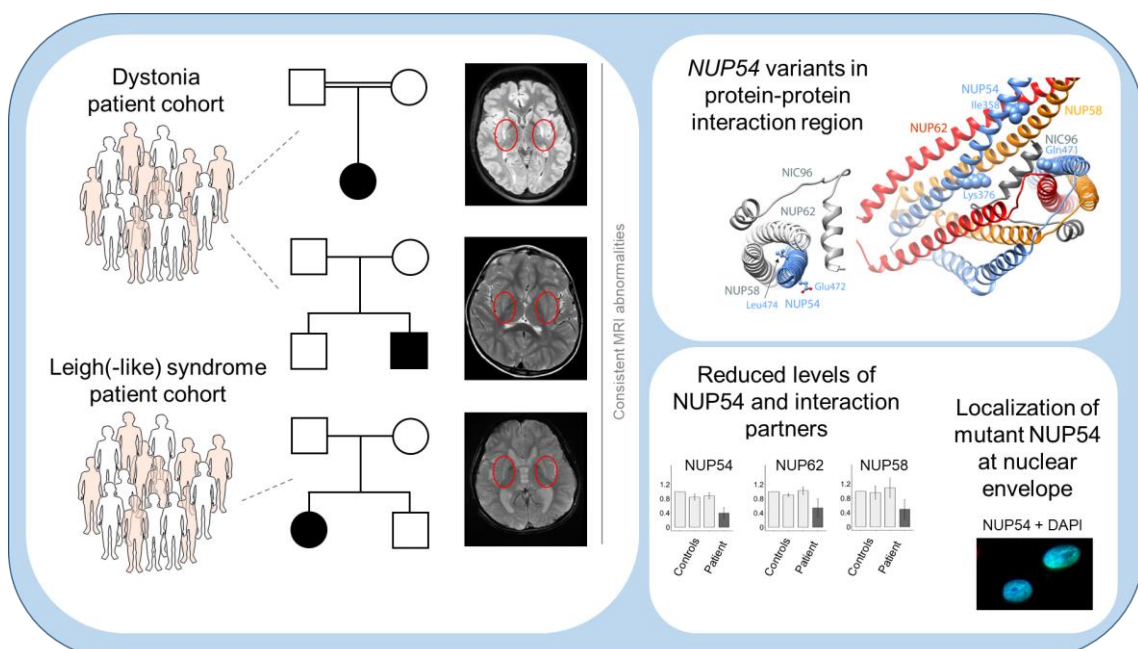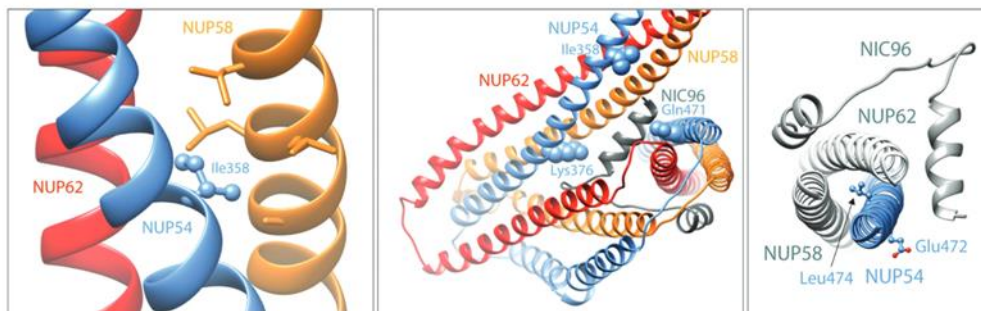
Graphical abstract

Figure 8: Graphical Abstract: Recessive *NUP54* Variants Underlie Early-Onset Dystonia

Annals of Neurology

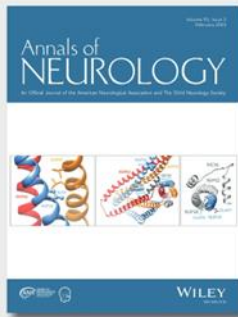AN OFFICIAL JOURNAL OF THE AMERICAN NEUROLOGICAL ASSOCIATION AND THE CHILD NEUROLOGICAL SOCIETY   Vol. 93 | No. 2 | February 2023

CONTENTS

ON THE COVER

In silico analysis of NUP54 variants identified in patients with dystonia. Structural analysis of mutated NUP54 residues within the NUP54-NUP62-NUP58 subcomplex is shown. Left panel: Schematic representation of wild-type "knob" NUP54 residue Ile358 that fits into hole formed by non-polar Ala and Leu side chains in NUP58 (in stick representation). The dystonia-associated Ile358Ser variant has apparently a deleterious effect on NUP54-NUP62-NUP58 trimerization processes as the polar Ser moiety is likely to destabilize the knob-into-hole contact between α-helices. NUP54, NUP62, and NUP58 are depicted in blue, red, and orange, respectively. Middle panel: Visualization of NUP54 Ile358, Lys376, and Gln471 within human hetero-trimeric NUP54-NUP62-NUP58 in complex with NIC96, a fungal homologue of the human nuclear pore complex (NPC) scaffold protein NUP93. Lys376 and Gln471 of NUP54 are surface residues in close vicinity of invaginated NIC96 α-helix and likely a part of the interacting surface. Dystonia-associated variants at these residues seem to affect the attachment of the heterotrimeric subcomplex to NIC96. In addition, the identified deletion of Gln471 would change the register of buried non-polar residues in the amphipathic

WILEY

NUP54 helix leading to gross changes in the stability of the C-terminal NUP54-NUP62-NUP58 coiled-coil region. Right panel: modeling of residues affected by two dystonia-associated missense changes in cis suggests that substitution of buried residue Leu474 to "bulkier" Phe most likely alters the packing of non-polar residues within the heterotrimeric NUP54-NUP62-NUP58 subcomplex. Glu472 is a surface residue, and the modelled human NUP54-NUP62-NUP58 with an interacting domain of NIC96 indicates that the variant Glu472Lys may perturb bridging interactions that anchor NUP54-NUP62-NUP58 to the NPC. See Harrer et al. (pp. 330–335).

WILEY

—— BRIEF COMMUNICATION ——

Recessive *NUP54* Variants Underlie Early-Onset Dystonia with Striatal Lesions

Philip Harrer, MD,[1,2] Audrey Schalk, MD,[3]
Masaru Shimura, MD, PhD,[1,4]
Sarah Baer, MD,[5,6]
Nadège Calmels, PharmD, PhD,[3,7]
Marie Aude Spitz, MD,[5]
Marie-Thérèse Abi Warde, MD,[5]
Elise Schaefer, MD, PhD,[8]
Volker M.Sc Kittke, MSc,[1,2]
Yasemin Dincer, MSc,[9,10]
Matias Wagner, MD [1,2]
Ivana Dzinovic, MSc,[1,2]
Riccardo Berutti, PhD,[1,2]
Tatsuharu Sato, MD,[11]
Toshihiko Shirakawa, MD, PhD,[11]
Yasushi Okazaki, MD, PhD [12]
Kei Murayama, MD, PhD [4,12]
Konrad Oexle, MD [1,2]
Holger Prokisch, PhD [1,2]
Volker Mall, MD,[9,13] Ivo Melčák, PhD,[14]
Juliane Winkelmann, MD,[1,2,15,16] and
Michael Zech, MD [1,2]

Infantile striatonigral degeneration is caused by a homozygous variant of the nuclear-pore complex (NPC) gene *NUP62*, involved in nucleo-cytoplasmic trafficking. By querying sequencing-datasets of patients with dystonia and/or Leigh(-like) syndromes, we identified 3 unrelated individuals with biallelic variants in *NUP54*. All variants clustered in the C-terminal protein region that interacts with NUP62. Associated phenotypes were similar to those of *NUP62*-related disease, including early-onset dystonia with dysphagia, choreoathetosis, and T2-hyperintense lesions in striatum. In silico and protein-biochemical studies gave further evidence for the argument that the variants were pathogenic. We expand the spectrum of NPC component-associated dystonic conditions with localized basal-ganglia abnormalities.

ANN NEUROL 2023;93:330–335

Introduction

In eukaryotes, protection of genome integrity and maintenance of the nuclear-transport machinery are mediated by the nuclear envelope, a physical-barrier system composed of the nuclear membranes and the nuclear-pore complexes (NPCs).[1] NPCs, formed by multiprotein assemblies that control trafficking between the nucleus and the cytoplasm, exhibit a strong degree of compositional conservation, and their functions are essential for tissue development and homeostasis.[2] Although the clinical importance of variants in most nuclear-envelope components remains unknown, those that have been implicated in Mendelian diseases frequently involve nervous-system pathology.[2, 3] Among the associated disorders, 2 produce movement disorder-predominant phenotypes: a dominantly inherited deletion-variant of the nuclear envelope-associated protein torsinA causes *TOR1A*-related dystonia, a childhood-onset dystonic syndrome with normal neuroimaging findings[3]; moreover, a missense variant in the gene encoding the NPC nucleoporin (NUP) protein NUP62 (*NUP62*) has been reported in

From the [1]Institute of Neurogenomics, Helmholtz Zentrum München, Munich, Germany; [2]Institute of Human Genetics, School of Medicine, Technical University of Munich, Munich, Germany; [3]Institut de génétique médicale d'Alsace (IGMA), Laboratoires de Diagnostic Génétique, Hôpitaux universitaires de Strasbourg, Strasbourg, France; [4]Center for Medical Genetics, Department of Metabolism, Chiba Children's Hospital, Chiba, Japan; [5]Department of Neuropediatrics, ERN EpiCare, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; [6]Institute for Genetics and Molecular and Cellular Biology (IGBMC), Illkirch, France; [7]Laboratoire de Génétique Médicale, INSERM U1112, Institut de génétique médicale d'Alsace, CRBS, Strasbourg, France; [8]Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Strasbourg, France; [9]Lehrstuhl für Sozialpädiatrie, Department of Pediatrics, Technische Universität München, Munich, Germany; [10]Zentrum für Humangenetik und Laboratoriumsdiagnostik (MVZ), Martinsried, Germany; [11]Department of Pediatrics, Nagasaki University Hospital, Nagasaki, Japan; [12]Diagnostics and Therapeutic of Intractable Diseases, Intractable Disease Research Center, Graduate School of Medicine, Juntendo University, Tokyo, Japan; [13]kbo-Kinderzentrum München, Munich, Germany; [14]Laboratory of Biochemical Pharmacology, Department of Pediatrics, Emory School of Medicine, Atlanta, Georgia, USA; [15]Lehrstuhl für Neurogenetik, Technische Universität München, Munich, Germany; and [16]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

Address correspondence to Dr Zech, Institute of Neurogenomics, Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany. E-mail: michael.zech@mri.tum.de

autosomal-recessive infantile striatonigral degeneration,[4] a multisymptomatic condition characterized by choreoathetosis, dystonia, and abnormal high T2-weighted MRI signals in the striatum. Here, we describe 3 families where biallelic variants in the gene encoding NUP62's direct interaction partner in the NPC,[1] *NUP54*, segregated with clinical presentations that showed striking similarities to *NUP62*-related disease.[4] Our study adds to the evidence that defects of nuclear-envelope components are associated with dystonia.[3]

Subjects and Methods

Case Identification

In 708 families with dystonia, we recently demonstrated that a disease-causing gene defect remains elusive in ~80% of cases.[5] With the goal of gene discovery, we reanalyzed genetic data from our cohort[5] resulting in identification of a *NUP54* candidate variant (family-A). A search for additional patients was undertaken using matchmaking nodes, which identified family-B in GeneMatcher[6] and family-C within the GENOMIT-project of individuals with suspected mitochondrial disorders including Leigh(-like) phenotypes. All families were enrolled in ethics review board-approved research protocols with informed consent. Families A-C were clinically evaluated in Munich, Germany; Strasbourg, France; and Tokyo, Japan.

Genetic Investigations

Whole-exome sequencing was performed on patient-parent trios using published procedures.[5, 7] Rare-variant interrogation was first done with virtual panels containing genes with described association with monogenic disorders to exclude known disease etiologies.[5, 7] New candidates were ranked based on established criteria.[5, 7] Sanger-verification and *in-silico* modeling[8, 9] were performed for *NUP54* variants; the structure of vertebrate NPC channel-NUP hetero-trimer NUP54-NUP62-NUP58 was utilized[10] (PDB:5C3L) and a human model of NUP54-NUP62-NUP58-NUP93 complex[11] was built using AlphaFold (https://alphafold.ebi.ac.uk/) from available templates (PDB:5CWS).

Western Blotting and Immunostaining

The following antibodies were used for Western-blot and/or immunocytochemistry experiments: anti-NUP54 (ab220890/HPA035929); anti-NUP62 (ab96134); anti-NUP58/NUP45 (HPA039360); and anti-NPC proteins/NUP98/NUP153 (mAb414).

Results

Clinical Cases

Phenotypic manifestations of 3 patients from 3 unrelated families (Fig 1A) are compared in Table; these individuals presented with shared features of progressive neurological deterioration, suggestive of underlying mixed neurodevelopmental-neurodegenerative pathologies. Findings common to all subjects were movement disorders with dystonia, which dominated the disease courses (Videos S1 and S2). Dystonic symptoms started in the legs between 12 months-5 years of age, followed by rapid involvement of craniocervical, trunk, and 4-limb muscles. Involuntary oro-bulbar spasms resulted in dysarthria and inability to swallow with need for tube feeding. Accompanying limb-choreoathetoid and/or ataxic movements were also seen in all patients. Neurodevelopmental symptoms, mainly motor delay and hypotonia, were documented, but only one patient had intellectual disability. Brain MRIs performed for patients of family-A (age 17 years) and family-B (age 5 years) revealed T2/FLAIR hyperintensities in the dorsal parts of both putamina (Fig 1B). In family-C's patient, MRI findings (age 7 years) were thought to resemble Leigh-syndrome, with symmetrical T2/FLAIR-hyperintense basal-ganglia lesions affecting the putamina (Fig 1B).

NUP54 Variants

The patients had homozygous or compound-heterozygous missense and in-frame deletion variants in *NUP54*, all located in close proximity toward the C-terminal end of the protein (Fig 1C; Table). An identical c.1073T > G (p.Ile358Ser) variant was identified in families A and B; the variant was homozygous in family-A's patient and carried in compound-heterozygosity with c.1126A > G (p.Lys376Glu) by family-B's patient. Family-C's patient harbored another set of compound-heterozygous alleles, a multi-nucleotide variation inducing 2 missense changes (c.1414G > A, p.Glu472Lys; c.1420C > T, p.Leu474Phe) and a c.1410_1412del (p.Gln471del) 1-amino acid deletion. The variants were extremely rare and predicted by CADD to be deleterious (Table). Additionally, all variants clustered at invariant residues within the evolutionarily conserved coiled-coil domains of NUP54[8, 9] (Fig 1D); these motifs are crucial for protein–protein interactions in the central NPC channel, supporting NUP54's complex formation with NUP62 and NUP58 and anchorage of the resultant triple-subcomplex to the NPC scaffold[8, 9, 11] (Fig 1E). Among these NPC-channel NUPs, NUP54 is the most critical for providing plasticity to multimeric assemblies of NUP54, NUP62, and NUP58[9] (Suppl Fig 1). Hence, mutational defects of NUP54 could result in destabilization of the interactions between functionally related NUPs and (partial) disassembly of the channel-forming triple-subcomplex. To test this hypothesis, we performed protein-modeling analyses revealing that indeed all variants were expected to perturb integrity of the NUP54-NUP62-NUP58 hetero-trimer and/or the structural stability of this complex in relation to the neighboring NPC scaffold-protein NUP93[8, 9, 11] (Fig 2A). Since destabilized protein-complexes, including those affecting correct NPC assembly, are often subject to cellular clearance mechanisms,[12] we investigated steady-state levels of different NUPs in patient-derived fibroblasts (families-B/C). We found significantly reduced amounts of NUP54,

ANNALS of Neurology
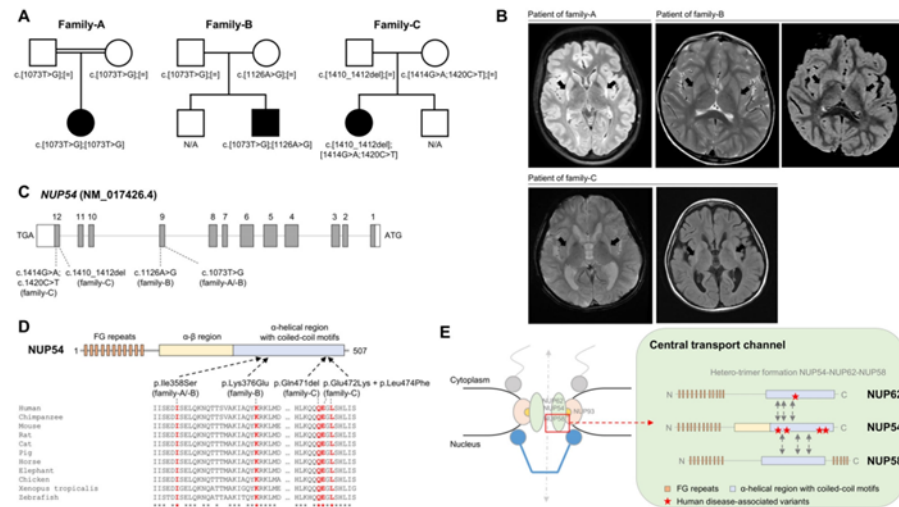
FIGURE 1: Domain-specific *NUP54* variants in 3 unrelated patients. (A) Pedigrees of studied families and segregation of *NUP54* variants supporting autosomal-recessive inheritance. N/A, no genotyping available. (B) MR axial images showing abnormally high signals (arrows) in the dorsal parts of both putamina (patient of family-A at 17 years; patient of family-B at 5 years; patient of family-C at 7 years). In addition, the patient from family-C had global loss of cerebral volume. (C) Schematic drawing of *NUP54* (canonical isoform: NM_017426.4) with location of the patient variants in the C-terminal exons. Intron-exon structure not drawn to exact scale. (D) Location of the variants on the NUP54 protein, mapping within the coiled-coil motif region and nearby residues. All affected amino-acid positions are highly conserved (fully conserved positions are highlighted with asterisks [*]). (E) Cartoon of the nuclear-pore complex (NPC; adapted from Guglielmi et al., 2020)[2] with the nucleoporins NUP54, NUP62, and NUP58 at its center (green). The coiled-coil motifs in α-helical regions of these nucleoporins are required to form triple-subcomplexes (indicated by double arrows) and maintain the structure of the central channel; the disease-associated variants in *NUP54* (this study) and *NUP62* (Basel-Vanagaite et al., 2006)[4] fall into this protein–protein interaction region. The central-channel hetero-trimer of NUP54-NUP62-NUP58 is anchored on the NPC via the scaffold protein NUP93 (yellow).

NUP62, and NUP58 in association to variants identified in this study (Fig 2B; Suppl Fig 2); by contrast, the non-associated NPC components NUP98 and NUP153 were intact (Fig 2B; Suppl Fig 3). Using immunofluorescent stainings, we were able to demonstrate that mutant NUP54-proteins were still localized at the nuclear envelope in affected fibroblasts (Fig 2C).

Discussion

We describe 3 independent patients with overlapping severe neurological phenotypes for whom trio-based whole-exome sequencing identified inherited variants of the nuclear-pore protein 54-encoding gene *NUP54* but no other pathogenic/or likely pathogenic gene-variations.

NUP54 belongs to the class of central-channel NUPs that represent integral components of the innermost layer of the NPC, controlling the transport of proteins, mRNAs, and other macromolecules into and out of the nucleus.[1] Channel NUPs are organized into 2 basic domain-structures[2]: (1) a phenylalanine-glycine-rich repeat region, located at the

N-terminus in NUP54, which binds directly to transport receptors of particular cargoes and mediates their nucleo-cytoplasmic exchange; and (2) a coiled-coil-motif segment in the C-terminal alpha-helical region, which is critically required for subcomplex formation with other NUPs, thereby maintaining the overall structural architecture of the central transport channel of the NPC. NUP54 and other transport-channel NUPs are highly expressed in the developing and adult brain,[13, 14] consistent with the key roles of the NPC in neurogenesis and promotion of neural maintenance.[2]

Collectively, we provide strong evidence that biallelic *NUP54* variants can cause a pediatric syndrome comprising progressive hyperkinetic movement abnormalities, striatal lesions, and variable neurodevelopmental disturbances. First, the identified variants, transmitted from asymptomatic carrier parents, segregated as expected for autosomal-recessive disease traits. No homozygous *NUP54* loss-of-function variants were represented in gnomAD, and we could neither find homozy-gous missense/in-frame deletion variants in the coiled-coil domain-encoding sequence of *NUP54* in gnomAD nor

TABLE. Molecular and Clinical Features of Patients with Biallelic *NUP54* Variants

	Patient of Family-A	Patient of Family-B	Patient of Family-C
Variant(s) (NM_017426.4)	c.1073T > G (p.Ile358Ser), homozygous	c.1073T > G (p.Ile358Ser); c.1126A > G (p.Lys376Glu), compound heterozygous	c.1410_1412del (p.Gln471del); c.1414G > A (p.Glu472Lys) + c.1420C > T (p.Leu474Phe), compound heterozygous
Variant frequency (gnomAD), CADD score	2/244430 (no homozygotes), 32	2/244430 (no homozygotes); not found, 32; 28	2/238818 (no homozygotes); not found + not found, 23; 24 + 27
ACMG classification (pathogenicity criteria)	pathogenic (PS3, PM1, PM2, PP3, PP4)	pathogenic (PS3, PM1, PM2, PP3, PP4); pathogenic (PS3, PM1, PM2, PP3, PP4)	pathogenic (PS3, PM1, PM2, PP3, PP4); pathogenic (PS3, PM1, PM2, PP3, PP4)
Gender, current age, geographical origin (ethnicity)	F, 22 yr, Germany (European)	M, 6 yr, France (European)	F, 18 yr, Japan (Asian)
Movement disorder(s); gross motor skills (last assessment)	Progressive generalized dystonia (onset 13 mo), dysarthria, dysphagia (PEG tube), choreoathetoid movements, ataxia; wheelchair use	Progressive, predominantly lower-limb dystonia (onset 5 yr), dysarthria, dysphagia (PEG tube), chorea, ataxia; wheelchair bound	Progressive generalized dystonia (onset 12 mo), dysarthria, dysphagia (PEG tube), ataxia; bedridden
Neurodevelopmental and other comorbidities	DD, hypotonia, microcephaly	Hypotonia, oculomotor apraxia, sleep apnea	DD, ID, aspiration pneumonia, congenital cataract, hypoparathyroidism, chronic nephritis
MRI	Symmetrical T2/FLAIR hyperintensities in dorsal putamina (age 17 yr)	Symmetrical T2/FLAIR hyperintensities in dorsal putamina (age 5 yr)	Symmetrical T2/FLAIR hyperintensities in dorsal putamina, progressive atrophy of the basal ganglia, global cerebral atrophy (age 7 yr)

Abbreviations: ACMG, American College of Medical Genetics and Genomics; CADD, combined annotation dependent depletion; DD, developmental delay; F, female; FLAIR, fluid-attenuated inversion recovery; ID, intellectual disability; M, male; MRI, magnetic resonance imaging; PEG, percutaneous endoscopic gastrostomy.

biallelic rare protein-altering *NUP54* variants in >25,000 clinical exomes of individuals with nonrelated conditions within our local databases. Bioinformatics analyses including frequency assessment, deleteriousness prediction, amino acid-conservation evaluation, and structural modeling demonstrated that the variants were ultra-rare and likely disruptive. Accordingly, all variants qualified as pathogenic alterations[15] (Table). Second, the observed C-terminal clustering of variants detected in different families was remarkable; in the 2 compound-heterozygous individuals, variants occurred on nearby residues within the same exons, a pattern that might reflect a pathophysiological mechanism. Notably, a homozygous p.Gln391Pro variant in the coiled-coil domain of NUP62 that directly interacts with NUP54's

coiled-coils has been shown to interfere with channel-NUP subomplex assembly,[8] suggesting a deleterious impact on NPC structure.[8] The functional importance of the coiled-coil motifs of channel NUPs has also been studied during Drosophila development, where C-terminal truncation of NUP54 was shown to produce neural-circuit architecture defects.[16] Given the locations of herein and previously reported[4] variants, in conjunction with support from the literature highlighting the role of NUP54's C-terminus in neurotypical development,[16] it seems plausible to hypothesize that perturbation of subcomplex interaction-domain residues of different channel NUPs could represent a general pathomechanism leading to related (neurodevelopmental) phenotypes. Similar to patients with *NUP62* variants,[4] our
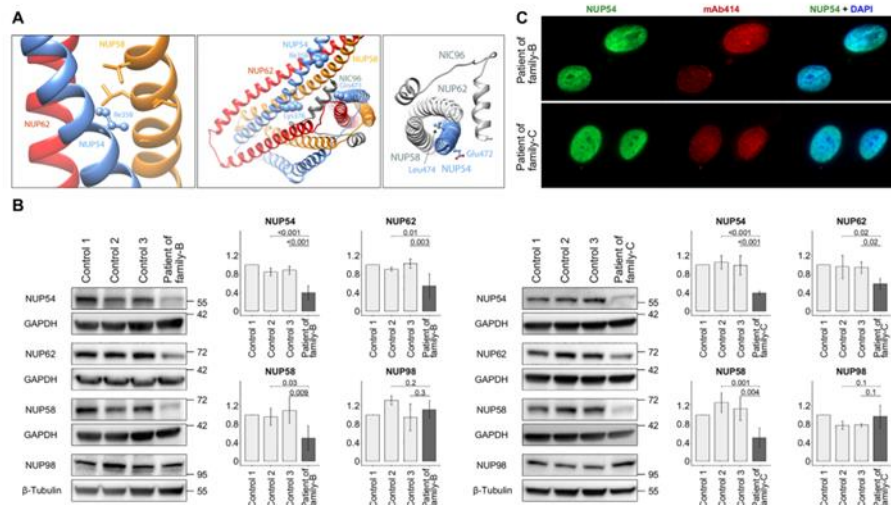
FIGURE 2: In silico and protein-biochemical studies of *NUP54* variants. (A) Structural analysis of mutated NUP54 residues within the NUP54-NUP62-NUP58 subcomplex. Left panel: Schematic representation of wild-type "knob"NUP54 residue Ile358 (in stick-and-ball representation) that fits into hole formed by non-polar Ala and Leu side chains in NUP58 (in stick representation). The Ile358Ser variant has apparently a deleterious effect on NUP54-NUP62-NUP58 trimerization processes as the polar Ser moiety is likely to destabilize the knob-into-hole contact between α-helices. The analyzed structure represents the partial vertebrate Xenopus laevis NUP54-NUP62-NUP58[10] (PDB:5C3L). The residues indicated are identical in the human NUP54 sequence and human Ile358 corresponds to Ile386 in Xenopus laevis NUP54. NUP54, NUP62, and NUP58 are depicted in blue, red, and orange, respectively. Middle panel: Visualization of NUP54 Ile358, Lys376, and Gln471 (in sphere representation) within human hetero-trimeric NUP54-NUP62-NUP58 in complex with NIC96, a fungal homologue of the human NPC scaffold protein NUP93. A human model of fungal NUP54-NUP62-NUP58 in complex with NIC96[11] was built using AlphaFold (https://alphafold.ebi.ac.uk/) from an available structure (PDB:5CWS). Lys376 and Gln471 of NUP54 are surface residues in close vicinity of invaginated NIC96 α-helix and likely a part of the interacting surface. Identified variants at these residues seem to affect the attachment of the heterotrimeric subcomplex to NIC96. In addition, the deletion of Gln471 would change the register of buried non-polar residues in the amphipathic NUP54 helix leading to gross changes in the stability of the C-terminal NUP54-NUP62-NUP58 coiled-coil region. Right panel: modeling of residues affected by the missense changes *in cis* suggests that substitution of buried residue Leu474 to "bulkier" Phe most likely alters the packing of non-polar residues within the heterotrimeric NUP54-NUP62-NUP58 subcomplex. Glu472 is a surface residue, and the modelled human NUP54-NUP62-NUP58 with an interacting domain of NIC96 (NUP93) indicates that the variant Glu472Lys may perturb bridging interactions that anchor NUP54-NUP62-NUP58 to the NPC. (B) Western blots showing reduced expression of NUP54 and its interaction partners NUP62/NUP58 in patient fibroblasts. Quantitative analyses of relative NUP protein levels are shown next to the Western blots; the intensities of total protein signals were normalized to GAPDH (G8795) or β-tubulin (11–13,002). Bars indicate the mean ± SD in lysates of separate fibroblast cultures. (C) Indirect immunofluorescent staining on fibroblasts of affected individuals. Expression of mutant NUP54 (green) is confined to the nucleus/nuclear envelope, similar to other NPC components (mAb414 staining; red). Nuclei were counterstained with DAPI (blue).

cases manifested early-onset dystonic and choreoathetoid movements with gradual progression, leading to significant impairments (1/3) or loss (2/3) of independent ambulation. Another important overlap was seen with regard to presence of marked oro-lingual-buccal dystonia with dysphagia and dysarthria, a clue that often points to strategic basal-ganglia lesions. Furthermore, patients with both *NUP54* and *NUP62* variants displayed distinctive neuroanatomical alterations with MRI abnormalities consisting primarily of a signal increase on T2/FLAIR-weighted images in the striatum. Third, we observed decreased amounts of NUP54 and its immediate subcomplex partner proteins in skin-fibroblasts

harboring the variants of our patients. Although this remains to be directly confirmed, the clustered variants are expected to induce expression changes of NUP54 and NUP62/NUP58 by impairing protein–protein interactions and multimerization processes, as has been shown for other NUP variant-alleles implicated in human disorders.[17] We speculate that mutant and/or mal-assembled patient proteins might be degraded by quality-control systems for NPC integrity; this would be consistent with studies in yeast and human cells demonstrating proteosomal clearance of defective NPC-assembly intermediates, e.g. via VPS4/VPS4A,[12] variants of which have also been associated with dystonia.[18] Absence of

Harrer et al: NUP54 Variants Cause Complex Dystonia

biallelic *NUP54* loss-of-function variants in our series and controls, as well as detectable levels of residual NUP54-protein in studied cells suggest that the identified variants are hypomorphic (i.e., partial loss-of-function) alleles; complete loss of NUP54 might be incompatible with life. Our immunocytochemistry analyses showing that patient fibroblasts had predominant localization of mutant NUP54 at the nuclear envelope indicate that the variants may exert their pathogenicity via disruption of nuclear pore-linked mechanisms such as macromolecular traffic-control. In line with this, an in vitro study of NUP-depleted cells identified significant alterations in nuclear-transport kinetics in association to a 50 to 75% decrease in total cellular amounts of NUP54.[19] However, a wide array of transport-independent or only indirectly transport-associated functions has been described for NPCs and nuclear pore-related factors including NUP54, ranging from regulation of chromatin states (similar to the dystonia-linked gene *KMT2B*)[2] to transposon silencing.[20] Future studies optimally involving iPSC-derived neuronal models are required to more precisely elucidate the impact of the NUP54 C-terminal variants on the channel-NUP subcomplex and their downstream effects in the context of dystonia pathogenesis.

Acknowledgments

Author Contributions

P.H., A.S., M.S., J.W., and M.Z. contributed to study concept and design. All authors contributed to data acquisition and analysis. P.H., I.M., and M.Z. contributed to drafting the manuscript and figures.

Potential Conflicts of Interest

None of the authors has any relevant conflict of interest to declare.

References

1. Strambio-De-Castillia C, Niepel M, Rout MP. The nuclear pore complex: bridging nuclear transport and gene regulation. Nat Rev Mol Cell Biol 2010;11:490–501.

2. Guglielmi V, Sakuma S, D'Angelo MA. Nuclear pore complexes in development and tissue homeostasis. Development 2020;147:dev183442.

3. Worman HJ, Dauer WT. The nuclear envelope: an intriguing focal point for neurogenetic disease. Neurotherapeutics 2014;11:764–772.

4. Basel-Vanagaite L, Muncher L, Straussberg R, et al. Mutated nup62 causes autosomal recessive infantile bilateral striatal necrosis. Ann Neurol 2006;60:214–222.

5. Zech M, Jech R, Boesch S, et al. Monogenic variants in dystonia: an exome-wide sequencing study. Lancet Neurol 2020;19:908–918.

6. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat 2015;36:928–930.

7. Schalk A, Cousin MA, Dsouza NR, et al. De novo coding variants in the AGO1 gene cause a neurodevelopmental disorder with intellectual disability. J Med Genet 2022;59:965–975.

8. Solmaz SR, Chauhan R, Blobel G, Melcak I. Molecular architecture of the transport channel of the nuclear pore complex. Cell 2011;147: 590–602.

9. Sharma A, Solmaz SR, Blobel G, Melcak I. Ordered regions of channel nucleoporins Nup62, Nup54, and Nup58 form dynamic complexes in solution. J Biol Chem 2015;290:18370–18378.

10. Chug H, Trakhanov S, Hulsmann BB, et al. Crystal structure of the metazoan Nup62*Nup58*Nup54 nucleoporin complex. Science 2015;350:106–110.

11. Stuwe T, Bley CJ, Thierbach K, et al. Architecture of the fungal nuclear pore inner ring complex. Science 2015;350:56–64.

12. Webster BM, Colombi P, Jager J, Lusk CP. Surveillance of nuclear pore complex assembly by ESCRT-III/Vps4. Cell 2014;159:388–401.

13. Thul PJ, Lindskog C. The human protein atlas: a spatial map of the human proteome. Protein Sci 2018;27:233–244.

14. Miller JA, Ding SL, Sunkin SM, et al. Transcriptional landscape of the prenatal human brain. Nature 2014;508:199–206.

15. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–424.

16. Nallasivan MP, Haussmann IU, Civetta A, Soller M. Channel nuclear pore protein 54 directs sexual differentiation and neuronal wiring of female reproductive behaviors in Drosophila. BMC Biol 2021;19:226.

17. Fichtman B, Harel T, Biran N, et al. Pathogenic variants in NUP214 cause "plugged" nuclear pore channels and acute febrile encephalopathy. Am J Hum Genet 2019;105:48–64.

18. Rodger C, Flex E, Allison RJ, et al. De novo VPS4A mutations cause multisystem disease with abnormal neurodevelopment. Am J Hum Genet 2020;107:1129–1148.

19. Yoo TY, Mitchison TJ. O-GlcNAc modification of nuclear pore complexes accelerates bidirectional transport. J Cell Biol 2021;220:e202010141.

20. Munafo M, Lawless VR, Passera A, et al. Channel nuclear pore complex subunits are required for transposon silencing in Drosophila. Elife 2021;10:e66321.

2.2 Dystonia linked to *EIF4A2* haploinsufficiency: a disorder of protein translation dysfunction

Contributions

The concept and design of the study were developed by me together with Matej Škorvánek, Volker Kittke, Juliane Winkelmann and Michael Zech. I planned the experimental procedure based on the pathophysiological hypothesis, coordinated the cultivation of patient fibroblasts together with Ivana Dzinovic, conducted the protein extraction, Western blots, and proximity ligation assays (PLA) together with Volker Kittke, and performed the data analyses of protein expression in Western blots and protein-protein-interactions in PLA. Drafting of the manuscript and figures was done by me and Michael Zech.

Graphical abstract
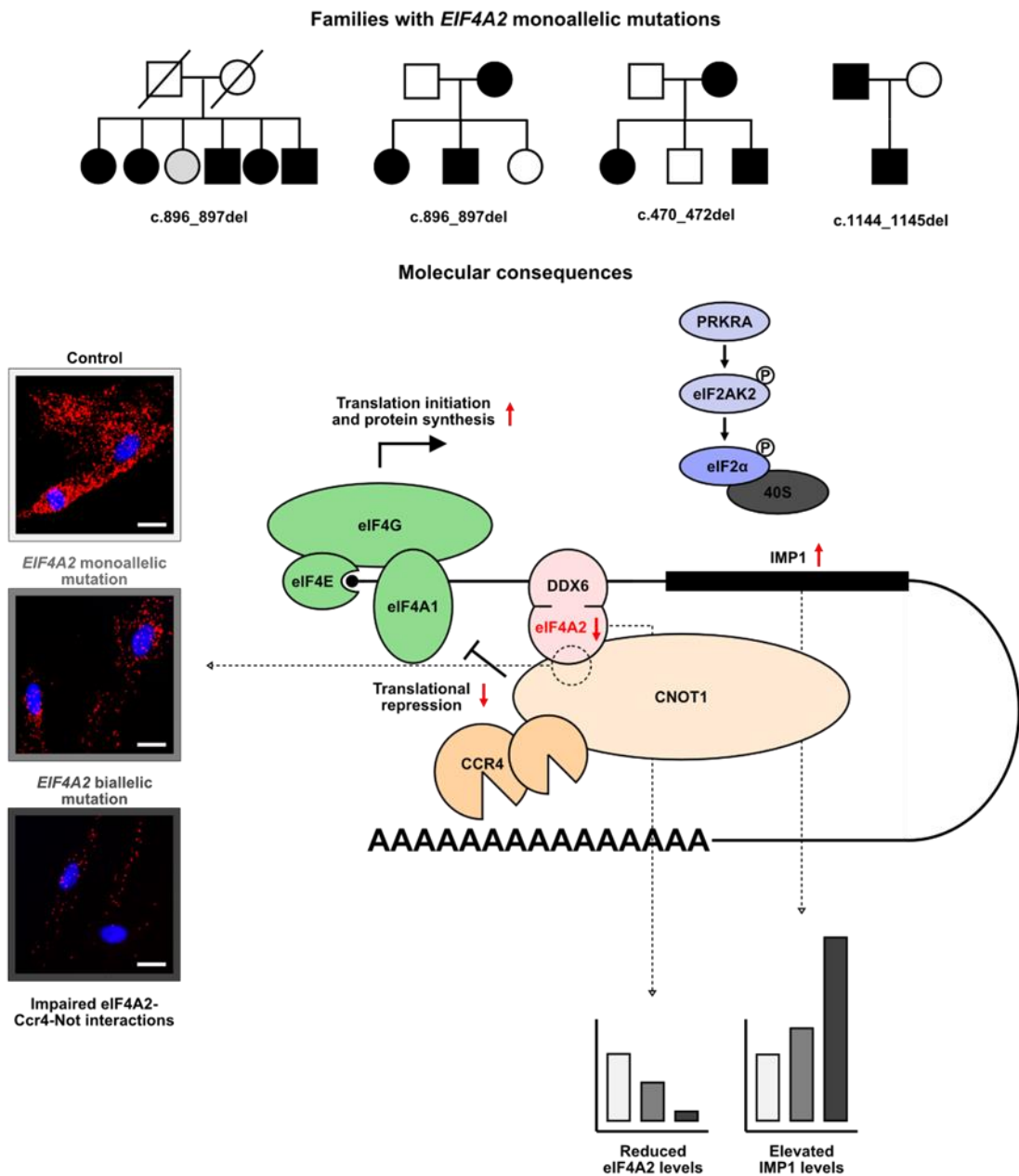
Figure 9: Graphical Abstract: Dystonia linked to *EIF4A2* haploinsufficiency: a disorder of protein translation dysfunction

RESEARCH ARTICLE

Dystonia Linked to *EIF4A2* Haploinsufficiency: A Disorder of Protein Translation Dysfunction

Philip Harrer, MD,[1,2] Matej Škorvánek, MD, PhD,[3,4] Volker Kittke, MSc,[1,2] Ivana Dzinovic, MSc,[1,2]
Friederike Borngräber, MD,[5] Mirja Thomsen, MSc,[6] Vanessa Mandel, MSc,[1] Tatiana Svorenova, MD,[3,4]
Miriam Ostrozovicova, MD,[3,4] Kristina Kulcsarova, MD,[3,4] Riccardo Berutti, PhD,[1,2] Hauke Busch, PhD,[7]
Fabian Ott, MSc,[7] Robert Kopajtich, MSc,[1,2] Holger Prokisch, PhD,[1,2] Kishore R. Kumar, MBBS, PhD, FRACP,[8,9]
Niccolo E. Mencacci, MD, PhD,[10] Manju A. Kurian, PhD,[11,12] Alessio Di Fonzo, MD, PhD,[13] Sylvia Boesch, MD,[14]
Andrea A. Kühn, MD,[5] Ulrike Blümlein, MD,[15] Katja Lohmann, PhD,[6] Bernhard Haslinger, MD,[16]
David Weise, MD,[17,18] Robert Jech, MD, PhD,[19] Juliane Winkelmann, MD,[1,2,20,21] and Michael Zech, MD[1,2*]

[1] Institute of Neurogenomics, Helmholtz Zentrum München, Munich, Germany
[2] Institute of Human Genetics, School of Medicine, Technical University of Munich, Munich, Germany
[3] Department of Neurology, P.J. Safarik University, Kosice, Slovak Republic
[4] Department of Neurology, University Hospital of L. Pasteur, Kosice, Slovak Republic
[5] Movement Disorder and Neuromodulation Unit, Department of Neurology, Charité—Universitätsmedizin Berlin, Berlin, Germany
[6] Institute of Neurogenetics, University of Lübeck, Lübeck, Germany
[7] Institute of Experimental Dermatology and Institute of Cardiogenetics, University of Lübeck, Lübeck, Germany
[8] Translational Neurogenomics Group, Molecular Medicine Laboratory and Neurology Department, Concord Clinical School, Concord
Repatriation General Hospital, The University of Sydney, Sydney, New South Wales, Australia
[9] Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia
[10] Ken and Ruth Davee Department of Neurology, Simpson Querrey Center for Neurogenetics, Northwestern University, Feinberg School of
Medicine, Chicago, Illinois, USA
[11] Department of Developmental Neurosciences, UCL Great Ormond Street Institute of Child Health, London, UK
[12] Department of Neurology, Great Ormond Street Hospital, London, UK
[13] Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Neurology Unit, Milan, Italy
[14] Department of Neurology, Medical University of Innsbruck, Innsbruck, Austria
[15] Department of Pediatrics, Carl-Thiem-Klinikum Cottbus, Cottbus, Germany
[16] Department of Neurology, Klinikum rechts der Isar, Technical University of Munich, School of Medicine, Munich, Germany
[17] Department of Neurology, Asklepios Fachklinikum Stadtroda, Stadtroda, Germany
[18] Department of Neurology, University of Leipzig, Leipzig, Germany
[19] Department of Neurology, Charles University in Prague, 1st Faculty of Medicine and General University Hospital in Prague,
Prague, Czech Republic
[20] Lehrstuhl für Neurogenetik, Technische Universität München, Munich, Germany
[21] Munich Cluster for Systems Neurology, SyNergy, Munich, Germany

ABSTRACT: Background: Protein synthesis is a tightly controlled process, involving a host of translation-initiation factors and microRNA-associated repressors. Variants in the translational regulator *EIF2AK2* were first linked to neurodevelopmental-delay phenotypes, followed by their implication in dystonia. Recently, de novo variants in *EIF4A2*, encoding eukaryotic translation initiation factor 4A isoform 2 (eIF4A2), have been described in pediatric cases with developmental delay and intellectual disability.
Objective: We sought to characterize the role of *EIF4A2* variants in dystonic conditions.

*Correspondence to: Dr. Michael Zech, Institute of Neurogenomics, Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany, E-mail: michael.zech@mri.tum.de

Philip Harrer, Matej Škorvánek, and Volker Kittke contributed equally to this work as co-first authors.

Juliane Winkelmann and Michael Zech contributed equally to this work as co-last authors.

HARRER ET AL

Methods: We undertook an unbiased search for likely deleterious variants in mutation-constrained genes among 1100 families studied with dystonia. Independent cohorts were screened for *EIF4A2* variants. Western blotting and immunocytochemical studies were performed in patient-derived fibroblasts.

Results: We report the discovery of a novel heterozygous *EIF4A2* frameshift deletion (c.896_897del) in seven patients from two unrelated families. The disease was characterized by adolescence- to adulthood-onset dystonia with tremor. In patient-derived fibroblasts, eIF4A2 production amounted to only 50% of the normal quantity. Reduction of eIF4A2 was associated with abnormally increased levels of IMP1, a target of Ccr4-Not, the complex that interacts with eIF4A2 to mediate microRNA-dependent translational repression. By complementing the analyses with fibroblasts bearing *EIF4A2* biallelic mutations, we established a correlation between IMP1 expression alterations and eIF4A2 functional dosage. Moreover, eIF4A2 and Ccr4-Not displayed significantly diminished colocalization in dystonia patient cells. Review of international databases identified *EIF4A2* deletion variants (c.470_472del, c.1144_1145del) in another two dystonia-affected pedigrees.

Conclusions: Our findings demonstrate that *EIF4A2* haploinsufficiency underlies a previously unrecognized dominant dystonia-tremor syndrome. The data imply that translational deregulation is more broadly linked to both early neurodevelopmental phenotypes and later-onset dystonic conditions. © 2023 The Authors. *Movement Disorders* published by Wiley Periodicals LLC on behalf of International Parkinson and Movement Disorder Society.

Key Words: *EIF4A2*; loss-of-function variants; translational dysfunction; dystonia; tremor

Introduction

Dystonia defines a phenotypically heterogeneous group of movement disorders that can be underlined by neurodegenerative lesions, neurodevelopmental abnormalities, or combinations of both.[1] Although genetic causes have been established for a growing number of syndromes involving dystonic features, many individuals with overlapping phenotypes remain undiagnosed, and the molecular mechanisms associated with known etiologies are diverse and incompletely understood.[2] A theme shared by several monogenic dystonias involves the maintenance of protein homeostasis,[3] which requires appropriate regulation of protein synthesis and turnover.[4] Defects of the translational machinery represent an important cause of human diseases related to proteome disturbance,[5,6] but only very few components of the translation apparatus have been shown to play a role in the pathogenesis of dystonic conditions. A neurodevelopmental disorder with infantile dystonia has been associated with biallelic variants in *SHQ1*, encoding a factor responsible for ribosome formation.[7] In addition, dystonia has been reported to result from pathologies in eukaryotic initiation factor 2α (eIF2α)-mediated processes,[8] known to be involved in neuronal development and survival.[9,10] In the early steps of protein synthesis, eIF2α functions in concert with the eukaryotic initiation factor 4F (eIF4F) complex consisting of eIF4E, eIF4G, and eIF4A, initiating or inhibiting the scanning of mRNAs.[11] Studies in animal models have demonstrated eIF2α perturbation in relation to mutations of the dystonia-linked genes *TOR1A*[8,12] and *THAP1*.[13] Furthermore, two upstream regulators of eIF2α have been implicated in hereditary dystonia: biallelic variants in *PRKRA* cause dystonia 16 (MIM: 612067),[14] whereas dominant *EIF2AK2* variants underlie dystonia 33 (MIM: 619687)[15]; both diseases are thought to arise as a consequence of translation-inhibition impairments.[15] Interestingly, the original descriptions of *EIF2AK2*-related disease were of individuals with neurodevelopmental phenotypes characterized by milestone delay and cognitive dysfunction,[16] followed by discovery of *EIF2AK2* variants in patients with dystonia.[15,17] Recently, variants in another component of the protein synthetic pathway, eukaryotic translation initiation factor 4A isoform 2 (eIF4A2, encoded by *EIF4A2*), were found to lead to neurodevelopmental disorders with developmental delay, intellectual disability, and epilepsy.[18] Especially de novo missense and deletion mutations (besides biallelic variants in two recessive pedigrees) were described, many of which were demonstrated to induce heterozygous loss of eIF4A2 function.[18] To date, no movement disorders have been reported with variants in *EIF4A2*. In this study, we mined large-scale genomic datasets of patients with dystonia[19,20] to identify two independent families with multiple affected individuals who segregated an identical unique frameshift *EIF4A2* variant. We found that the variant reduced eIF4A2 protein amounts consistent with haploinsufficiency, resulting in deregulation of translation control. We showed that this effect was likely related to an impaired ability of eIF4A2 to associate with Ccr4-Not, a master regulatory complex involved in microRNA-mediated translational inhibition.[21] Two additional rare *EIF4A2* deletion changes were prioritized from dystonia genetics consortia, suggesting a broader role for *EIF4A2* variants in causing dystonic phenotypes.

Subjects and Methods

Subjects and Molecular Methods

The study cohort used for primary analysis consisted of 1100 unrelated index patients with a diverse range of dystonic phenotypes, including isolated dystonia (59%) and dystonia with other neurological features

and/or extraneurological involvement; detailed demographics and clinical characteristics of the patients' conditions have been described elsewhere.[19,20] All subjects in the cohort had been recruited from movement disorders clinics through the practices of the investigators or by referral for dystonia genetics research from various international collaboration partners. The herein described patient II-4 from family A (A-II-4) and patient III-2 from family B (B-III-2) were part of this primary analysis cohort. Written informed consent had been obtained from the participating individuals or, in the case of children or those with intellectual impairment, from parents or legal representatives. Data collection and molecular studies were conducted in accordance with the standards of respective ethics institutional review boards. Each individual had undergone in-depth phenotypic evaluation with clinical examination, magnetic resonance imaging and routine laboratory studies when available, review of medical records, and assessment of affected family members. As part of an ongoing endeavor to uncover the genetic causes of dystonia, the individuals had received research whole-exome sequencing (WES) in different family-based analysis designs (sequencing of at least one additional affected or unaffected family member in 30%, including patients II-2 and II-5 from family A [A-II-2 and A-II-5] and patient II-2 from family B [B-II-2]). Our local WES protocols using Agilent enrichment kits and Illumina machines for generation of 100-bp paired-end reads have been reported previously.[19,20] Data were annotated and filtered according to established procedures with an in-house bioinformatics pipeline, as described previously.[19,20] Variant filtering included consideration of allele frequencies in population databases, expected impact on protein, gene constraint, pathogenicity predictions, and inheritance. Variants surviving the filtering steps were manually evaluated and prioritized.[22,23] In this study, we chose a prioritization strategy different from our previously applied methods designed for discovery of novel candidate genes,[22-24] combining the following lines of evidence for the variant(s) of interest: (1) protein-altering alteration absent from controls; (2) variant located in a mutation-constrained gene as determined by recommended statistical metrics[25]; (3) variant recurrent among unrelated patients; and (4) variant present in WES data of affected family members only. Pathogenic or likely pathogenic variants in established dystonia-associated genes were ignored. To identify additional putative disease-related variants in the selected candidate gene *EIF4A2*, we queried independent dystonia genomic sequencing datasets acquired in the context of multi-institutional consortia or center-specific research projects (Australian dystonia genomes; Lübeck dystonia exome project, Germany; dystonia exomes/genomes at UCL Great Ormond Street Institute, London, UK;

Fondazione Ca' Granda IRCCS, Milan, Italy; and Ken and Ruth Davee Department of Neurology, Chicago, IL, USA); respective cohorts and sequencing initiatives have been described before.[22,24,26] Candidate *EIF4A2* variants were confirmed and tested for cosegregation in all available family members by Sanger sequencing.

Human Cell Culture

Fibroblast lines were established from skin biopsies of a patient with heterozygous *EIF4A2* variant (patient A-II-4), healthy control subjects, as well as a previously reported pediatric patient with mixed neurodevelopmental-neurodegenerative disease and biallelic variants in *EIF4A2*.[18] The sample of the latter individual was included in this study to investigate the molecular effect of eIF4A2 protein loss in a dosage-dependent manner and to assess further a recently proposed correlation[18] between residual eIF4A2 amounts and differences in phenotypic outcomes. Cells were cultured according to established procedures.[23]

Western Blotting

Fibroblast protein extracts were prepared for Western blotting by standard methods.[23] Antibodies were used against the following proteins: eIF4A2 (1:20,000, ab31218; Abcam), eIF4A1 (1:20,000, ab31217; Abcam), IMP1 (1:500, 2852S; Cell Signaling), and DDX6 (1:1000, BLD-674402; BioLegend). All primary antibodies were used according to the manufacturer's instructions. Densitometric analyses were carried out with ImageJ, and statistical comparisons were performed with R; significances were calculated by unpaired 2-tailed t tests.

Proximity Ligation Assay

Proximity ligation assays (DUO92008; Sigma-Aldrich) were performed in accordance with the manufacturer's recommendations and by using a modified version of the previously published protocol for studying cellular interactions between eIF4A2 and CNOT1.[27] In short, primary antibody incubations were performed at 4°C overnight with antibodies against the following proteins: eIF4A2 (1:2000, sc-137,148; Santa Cruz) and CNOT1 (1:2000, 14,276-1-AP; Proteintech). Negative control reactions were performed using only one primary antibody on cells of control individuals. Nuclei were stained with DAPI. Cells from three biological replicates of each line were imaged on an Axio Imager Z1 (Zeiss) using an EC Plan-Neofluar 20×/0.50 M27 objective, recording 10–15 images per biological replicate. Images were evaluated using the image analysis software Definiens Developer XD 2 (Definiens AG, Munich, Germany). A specific rule set was defined to automatically detect nuclei, as well as fluorescent spots originating from the proximity ligation assay. The quantified parameter was the average number

HARRER ET AL

of spots per nucleus for each image (n = 35–39 per patient/control individual with an average of 10–15 cells per image), which was compared between patient and control cells using Student *t* test with Bonferroni correction for multiple testing.

Results

Genetics Data

Through integrated analysis of rare variants expected to be damaging to protein function and shared by independent dystonia-affected patients from 1100 exome-sequenced families,[19,20] we filtered out as top disease-causal candidate a single heterozygous 2-bp deletion (c.896_897del) in exon 8 of *EIF4A2* (NM_001967.4) (Fig. 1A,C, Table 1). Consistent with

the stringent prioritization scheme that we applied, this frameshifting allele was predicted to lead to a loss-of-function effect either by giving rise to generation of a deleteriously truncated polypeptide (p.-Thr299Serfs*7) or, more likely, by triggering nonsense-mediated mRNA decay with no protein production. The variant was unobserved in >140,000 control individuals from gnomAD (version 2.1/version 3.1 releases) and 40,000 in-house control chromosomes, and it affected a gene heavily depleted for loss-of-function variation in the general population (gnomAD probability of being loss-of-function intolerant (pLI) score = 1.0, loss-of-function variant observed vs. expected ratio = 0.04, confidence interval = 0.01–0.2).[25] As shown in Fig. 1A, c.896_897del was identically present in WES data of five dystonia-affected individuals from two separate pedigrees, including three siblings from a
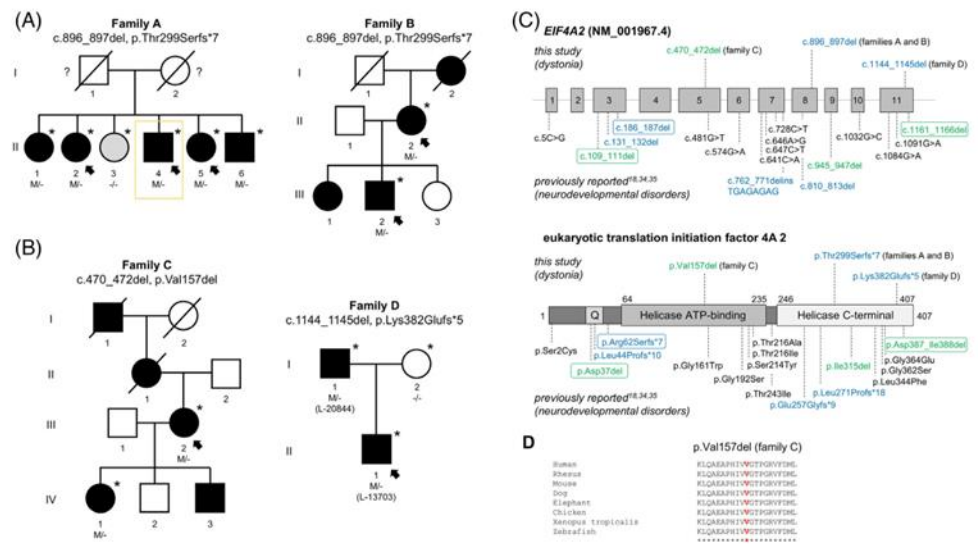
FIG. 1. Family trees and *EIF4A2* variants. (**A** and **B**) Pedigrees of the two unrelated families with the recurrent heterozygous c.896_897del (p.Thr299Serfs*7) frameshifting allele (**A**) and the additionally recruited families with *EIF4A2* deletion alterations (**B**). The dystonia-/tremor-affected individuals are represented by solid black symbols, white symbols indicate healthy family members, and a gray shaded shape in family A indicates a sister of the siblings with dystonia who displayed a clinically different neurological syndrome characterized by progressive dementia and ataxia. Diagonal lines denote deceased subjects. Phenotypic status was unknown for parents in family A (?). The family members marked by asterisks were clinically examined. Arrows point to individuals who were exome sequenced. Subjects with DNA available are labeled according to *EIF4A2* variant status: M/−, heterozygous carrier; −/−, homozygous wild-type. Patient A-II-4 from family A for whom a fibroblast line was obtained and characterized in functional studies is highlighted with an orange box. (**C**) Schematics of the *EIF4A2* gene and the encoded protein's primary structures. Functionally important protein regions are illustrated: the Q motif, the helicase ATP-binding domain, and the helicase C-terminal domain. Locations of herein identified and previously reported[18,28,29] variants are shown at the cDNA (upper panel) and the protein (lower panel) level. The dystonia-associated variants observed in families A–D are plotted above and variants described in patients with neurodevelopmental disease syndromes[18,28,29] below the gene/protein. Variants depicted in blue, green, and black represent loss of function (frameshift), in-frame deletion, and missense changes, respectively. All literature-reported variants were heterozygous de novo mutations, except for three biallelic variants that are marked with boxes; a newly acquired, previously undescribed fibroblast line established from a compound heterozygous carrier of c.186_187del (p.Arg62Serfs*7) and c.1161_1166del (p.Asp387_Ile388del)[18] was analyzed in the functional characterization studies performed as part of this work. (**D**) Multiple alignment of eukaryotic translation initiation factor 4A 2 orthologs showing evolutionary conservation of the amino acid residue deleted by the c.470_472del variant in family C. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Clinical characteristics of individuals harboring EIF4A2 heterozygous variants identified in this study

Patient	Gender/ Ethnicity	EIF4A2 variant	Age at last examination, y	Movement disorders at last examination	Involved areas (distribution)	Age at movement disorder onset, y (site of onset)	Additional neurological features (cognitive dysfunction and/or behavioral problems)	Brain MRI
A-II-4	M/European	c.896_897del (p.Thr299Serfs*7)	67	Dystonia, tremor, jerky movements, dyskinesia	Cranial, cervical, brachial, truncal (generalized)	55 (arms)	Yes	Normal
A-II-2	F/European	c.896_897del (p.Thr299Serfs*7)	70	Dystonia, tremor, jerky movements	Cervical, brachial (segmental)	60 (arms)	Yes (mild)	ND
A-II-1	F/European	c.896_897del (p.Thr299Serfs*7)	72	Dystonia, tremor, jerky movements	Cervical, brachial (segmental)	65 (arms)	Yes	ND
A-II-5	F/European	c.896_897del (p.Thr299Serfs*7)	66	Dystonia, tremor, jerky movements, dyskinesia	Cranial, cervical, brachial (segmental)	45 (arms)	Yes	ND
A-II-6	M/European	c.896_897del (p.Thr299Serfs*7)	59	Dystonia, tremor, jerky movements	Cervical, brachial (segmental)	51 (arms)	Yes (mild)	ND
B-III-2	M/European	c.896_897del (p.Thr299Serfs*7)	35	Dystonia, tremor, jerky movements, dyskinesia	Cranial, cervical, brachial, truncal (generalized)	16 (neck)	Yes (mild)	Normal
B-II-2	F/European	c.896_897del (p.Thr299Serfs*7)	61	Dystonia, tremor	Cervical, brachial (segmental)	13 (neck)	Not reported	ND
C-III-2	F/European	c.470_472del (p.Val157del)	45	Dystonia, tremor, jerky movements	Cranial, cervical, brachial, truncal (generalized)	3 (NK)	Not reported	Normal
C-IV-1	F/European	c.470_472del (p.Val157del)	23	Dystonia, tremor	Cervical, brachial, truncal (generalized)	10 (arms)	No	ND
D-II-1	M/European	c.1144_1145del (p.Lys382Glufs*5)	33	Dystonia, tremor	Cranial, cervical, brachial, truncal (generalized)	6 (arms)	Yes (mild)	Enlarged CSF spaces
D-I-1	M/European	c.1144_1145del (p.Lys382Glufs*5)	59	Tremor	Brachial (focal)	20–25 (arms)	Not reported	ND

Note: Variants are annotated according to genome build GRCh37/hg19 and EIF4A2 transcript NM_001967.4. None of the variants except for c.1144_1145del (patient D-II-1) are found in gnomAD or in-house control databases (>160,000 control datasets in total); c.1144_1145del is observed in one single gnomAD individual. pLI (probability of being loss-of-function intolerant) for heterozygous EIF4A2 loss is 1.00 (observed vs. expected ratio = 0.04).
Abbreviations: MRI, magnetic resonance imaging; F, female; M, male; ND, not performed; NK, not known; CSF, cerebrospinal fluid.

HARRER ET AL

Slovak family (family A, patient A-II-4 was part of the primary analysis cohort; see Subjects and Methods) and a mother–son pair of German descent (family B, patient B-III-2 from the primary analysis cohort; see Subjects and Methods). These patients' exomes contained no alternative rare variants considered to be responsible for their dystonic phenotypes. Sanger sequencing in additionally recruited members of family A detected c.896_897del in another two siblings with similar dystonic features, whereas the variant was not found in a sixth sister presenting a clinically distinct condition with progressive dementia and ataxia of suspected neurodegenerative origin (Fig. 1A). Genetic material of two further dystonia-affected relatives in family B was not available for segregation testing. Together, c.896_897del (p.Thr299Serfs*7) fulfilled criteria for classification as a "likely pathogenic" variant according to the American College of Medical Genetics and Genomics standards.[30] Our subsequent search for more *EIF4A2* candidate dystonia-associated variants singled out a heterozygous one-amino acid deletion, c.470_472del (p.Val157del), in a multigenerational pedigree with five affected individuals (family C) (Fig. 1B,C, Table 1); this solo WES-identified variant, absent from all aforementioned control databases, was predicted to disturb a phylogenetically highly conserved residue within the functional helicase ATP binding domain,[18] and cosegregation work demonstrated its presence in an affected offspring of family C's index case (Fig. 1B–D). Moreover, a fourth unrelated patient (family D) was identified who harbored a rare heterozygous frameshift variant, c.1144_1145del (p.-Lys382Glufs*5) (Fig. 1B,C, Table 1); c.1144_1145del was located in the last exon of *EIF4A2*, present in a single gnomAD control, and inherited from a tremor-affected father (Fig. 1B,C). The c.470_472del (p.-Val157del) and c.1144_1145del (p.Lys382Glufs*5) changes were formally classified as "variants of uncertain significance" according to American College of Medical Genetics and Genomics criteria.[30] Screening of available WES data for families C and D did not identify any other suspicious monogenic variant hits in the context of the observed dystonic presentations.

Clinical Findings

Five siblings in family A had overlapping phenotypes characterized by adult-onset dystonia associated with marked tremor and occasional myoclonic features (Table 1). Subject A-II-4 manifested involuntary tremulous movements of both arms at age 55 years, followed by appearance of constant head deviation, writing difficulties, and jerks with upper-body predominance at around age 60. Examination (age 67) indicated right torticollis, mild dystonic finger posturing, upper-limb postural tremor, irregular jerky movements of the

shoulder girdle musculature, and facial dyskinesia. His sister, A-II-2, reported impairments of fine motor skills and abnormal head postures since age 60; at age 70, she displayed jerky action and postural tremor of the hands, involuntary forearm pronation, and tremulous cervical dystonia. All other affected siblings (A-II-1, A-II-5, and A-II-6) developed similar signs of dystonia, tremor, and intermittent myoclonus-like jerks between 45 and 65 years of age; on assessment they had variably expressed combinations of jerky head and/or limb tremor, impaired finger dexterity, dystonia with craniocervical involvement, and perioral dyskinesia. In family B, the son (B-III-2) first noticed involuntary movements of his neck and right shoulder at age 16; over the following years, bibrachial tremor emerged, and symptoms spread to the trunk and face. During follow-up evaluations (age 30–35 years), he showed nonprogressive tremulous cervical dystonia with torti-retrocollis, trunk deviation, postural arm tremor, myoclonic jerks of the left hand, and orofacial abnormal movements. His mother (B-II-2) was diagnosed with adolescence-onset segmental dystonia; she presented with left torticollis and mild dystonic action tremor of both arms. Movement disorder features shared between seven individuals from two families with the c.896_897del (p.Thr299Serfs*7) variant are summarized in Table 1. Further clinical findings for some of these patients included relevant degrees of stable cognitive dysfunction and behavioral comorbidities (depressive-like behavior, anxiety, social withdrawal). The index patient in family C (C-III-2) demonstrated generalized dystonia, with pronounced craniocervical involvement, arm tremor, and intermittent hand jerky movements; her daughter (C-IV-1) experienced mild laterocollis, trunk dystonic movements, and postural tremor of the hands with involuntary finger cramps (Table 1). Finally, family D's patient (D-II-1) had bilateral arm tremor since childhood, followed by manifestation of generalized dystonia in adolescence; his father (D-I-1) presented upper-limb postural and action tremor since the age of 20 to 25 years (Table 1).

Functional Studies

To define the impact of the recurrent c.896_897del variant on eIF4A2, we performed immunoblotting on available fibroblasts from control individuals and family A patient A-II-4. In addition, cells from a published patient with biallelic *EIF4A2* variants,[18] for whom no cellular phenotypes have been described before, were included in the analysis. The abundance of eIF4A2 was reduced to ~50% in patient A-II-4 and to ~10% to 20% in the patient with biallelic variants relative to control individuals (Fig. 2A,B), confirming a variant zygosity-dependent loss of eIF4A2 levels in the mutation carriers. It is well appreciated that the DEAD-box
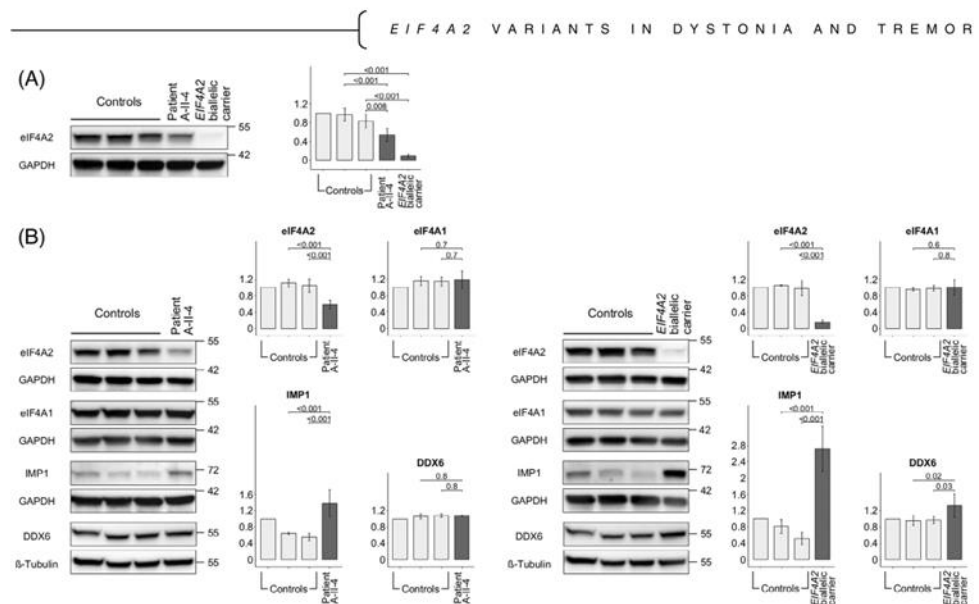
FIG. 2. Western blot analyses of fibroblasts obtained from patients with *EIF4A2* variants and control donors. (**A** and **B**) Primary skin fibroblasts derived from the dystonia-affected individual A-II-4 (family A with heterozygous *EIF4A2* variant c.896_897del), a compound heterozygous carrier of *EIF4A2* c.186_187del/c.1161_1166del[18] (*EIF4A2* biallelic carrier), and three unrelated control subjects were cultured under identical conditions. Total cell lysates were prepared for determination of protein steady-state levels. Western blot assays were performed using antibodies specific for eIF4A2 (**A**) and eIF4A2, eIF4A1, IMP1, as well as DDX6 (**B**). Anti-GAPDH and anti–β-tubulin antibodies were used as a loading reference. Bar charts represent densitometric measurements of eIF4A2, eIF4A1, IMP1, and DDX6 levels (normalized to GAPDH or β-tubulin); bars indicate the means ± standard deviation in lysates of three separate fibroblast cultures. Sizes of detected proteins are shown in kilodaltons, and *P* values are provided for the statistical analyses. An *EIF4A2* variant-specific decrease in eIF4A2 levels (**A** and **B**) and a corresponding, inversely correlated increase in the levels of the downstream translational target IMP1 (**B**) were observed.

RNA helicase eIF4A2, unlike its paralog eIF4A1, exerts dual functions in translational regulation.[21] Besides playing a role in the stimulation of translation initiation via interactions with other eukaryotic initiation factors in the eIF4F complex and eIF2α, eIF4A2 is known as a key effector in microRNA-mediated repression of translation through association with the Ccr4-Not complex.[21] Previous in vitro experiments have demonstrated that artificial knockdown of eIF4A2 critically altered protein levels of Ccr4-Not–related microRNA targets such as IMP1.[27] In light of these findings, we sought to assess whether IMP1 expression was deregulated in the presence of patient *EIF4A2* variants. As shown in Fig. 2B, basal IMP1 concentrations were significantly higher in both mutant fibroblast lines compared with controls, with a clear eIF4A2 protein dosage loss-dependent effect (~40%–50% and ~150%-170% IMP1 expression increase in cells of patient A-II-4 and the patient with biallelic variants, respectively). Remarkably, the abundance of eIF4A1 was not affected by the *EIF4A2* variants, whereas the expression of DDX6, another DEAD-box RNA helicase involved in Ccr4-Not-associated translational inhibition,[27] was

upregulated to ~20%–30% only in the cells with biallelic variants (Fig. 2B). This suggested that eIF4A1 and DDX6 were unable to compensate for the heterozygous loss of eIF4A2 in patient A-II-4. Collectively, these studies established that c.896_897del induced *EIF4A2* haploinsufficiency, and that the variant was associated with alterations of translational control suggestive of Ccr4-Not complex dysfunction. To validate a potential effect of the *EIF4A2* variants on eIF4A2-Ccr4-Not interactions, we performed proximity ligation assays[27] in patient and control fibroblasts. Again, we observed a correlation between the extent of eIF4A2 loss and cellular outcomes: compared with control cells, mutants harboring c.896_897del exhibited an ~35%–45% decrease (adjusted *P* < 0.001 for all comparisons) in colocalization of eIF4A2 and the Ccr4-Not component CNOT1, whereas this colocalization was almost completely lost in the patient line with biallelic variants (Fig. 3A,B). These experiments implied that because of the reduced eIF4A2 protein levels, the functionally important association with Ccr4-Not was impaired in patient A-II-4, although less significantly than in the pediatric case with recessive disease.
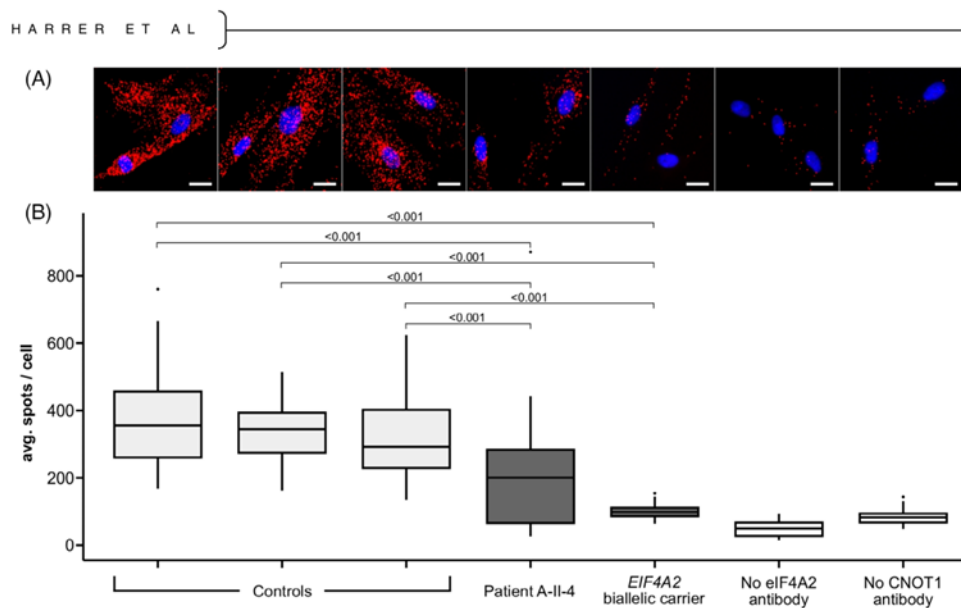
FIG. 3. Proximity ligation assay (PLA) on patient and control fibroblasts. (**A**) Epifluorescence microscopy images showing the association of eIF4A2 with the Ccr4-Not complex factor CNOT1 (from left to right): PLA reaction (red) in skin-derived primary fibroblasts from three unrelated control subjects, the dystonia-affected individual A-II-4 (family A with heterozygous *EIF4A2* variant c.896_897del), a compound heterozygous carrier of *EIF4A2* c.186_187del/c.1161_1166del (*EIF4A2* biallelic carrier),[18] and negative controls for each antibody on control subjects. Nuclei were stained with DAPI (blue). Images are representative of three independent experiments. Scale bars, 25 μm. (**B**) Quantification of PLA reaction dots, as assessed in three biological replicates. Raw data were analyzed using the Definiens Developer XD 2 software (Definiens AG, Munich, Germany); boxplots represent the median, first quartile, and third quartile, and whiskers extend to a maximum of 1.5× interquartile range. Statistically significant differences are indicated (unpaired 2-tailed *t* tests). The number of PLA dots was substantially diminished in cells bearing *EIF4A2* variants, indicating reduced but still functional interactions between eIF4A2 and the Ccr4-Not complex. [Color figure can be viewed at wileyonlinelibrary.com]

Discussion

By molecular and clinical characterization of individuals with heterozygous *EIF4A2* variants, we provide evidence for a previously unrecognized monogenic movement disorder. Our findings substantially broaden the clinical spectrum of *EIF4A2*-associated neurodevelopmental disorders to include dystonia-predominant manifestations, similar to observations in *EIF2AK2*-related disease, another condition linked to the protein translation machinery, which is characterized by presentations of both intellectual developmental syndromes[16] and isolated dystonia.[15,17] Our patients' phenotypes comprised dystonic features of variable severity, tremor, and jerky movements resembling myoclonus. The conditions bore some distinct similarities to presentations related to variants in *ANO3* (dystonia 24; MIM: 615034) and *KCTD17* (dystonia 26; MIM: 616398), with onset in adulthood or adolescence and leading involvement of the upper body (craniocervical region, arms).[31,32] The observed distribution of dystonia may help to distinguish patients with

EIF4A2 variants from those with variants in other recurrently mutated genes for dystonia, such as *TOR1A* and *KMT2B*, where prominent leg involvement is often seen.[33,34] The type of spreading of movement disorder features among patients from our four different families was variable, but initial manifestation in the upper extremities with secondary affection of neck and facial muscles was frequently noted (for details, see Table 1). Some patients also displayed nonprogressive cognitive impairments and behavioral/neuropsychiatric disturbances that might be regarded as signs of *EIF4A2*-associated developmental dysfunction[18]; there were, however, no reports of milestone delays or epileptic comorbidities, although we could not precisely assess early neurodevelopment because of advanced age of all subjects. In affected individuals of family A, the reported age at movement disorder onset was considerably later than in patients of family B (late adulthood vs. adolescence), although both families segregated the exact same *EIF4A2* variant. This difference might be explained by phenotypic heterogeneity related to modifying genetic, epigenetic, and/or environmental factors,

as commonly recognized in rare and more prevalent neurogenetic disease conditions.[35] In contrast, most affected members of family A did not actively seek medical attention for many decades in their life, whereas others never visited a neurologist before family-based movement disorder assessment as part of this study; therefore, we cannot exclude that milder undiagnosed dystonic and/or tremulous signs may have pre-existed during adolescence/younger adulthood in some of these individuals. Difficulties with cognition and abnormalities in behavior were more pronounced in some older persons from family A, an observation that could be associated with either incidental clinical variation or family-specific disease progression over a lifetime.

For the frameshift variant c.896_897del (p.-Thr299Serfs*7), we offer strong arguments for causal implication in the observed phenotypes, including demonstration of its rarity, segregation with disease in non-related pedigrees, and effect on protein and the downstream biological pathway. First, we demonstrated that c.896_897del led to ∼50% reduction of eIF4A2 protein amounts, indicative of degradation of the mutant transcript and/or the truncated polypeptide. *EIF4A2* haploinsufficiency has to be considered disease causing given that (1) several recently reported neurodevelopmental disorder–associated *EIF4A2* variants, including missense and frameshift alterations, were shown to represent dominant loss of eIF4A2 function mutations[18]; (2) *EIF4A2* heterozygous predicted loss-of-function variants exhibit significant enrichment in de novo variation catalogs derived from large neurodevelopmental disease cohorts[28,29] (Fig. 1C); and (3) loss of one *EIF4A2* copy is not tolerated among population controls.[25] Second, our studies in patient-derived cells uncovered a specific role for *EIF4A2* variants in producing perturbance of translational regulation, demonstrating an increased expression of the Ccr4-Not complex target IMP1 in association to c.896_897del; this finding strikingly recapitulated published in vitro observations from eIF4A2 knockdown systems.[27] In vivo work has established that *EIF4A2* variants identified in neurodevelopmental disorder cases compromised neuromotor function and morphological development,[18] but the underlying molecular mechanisms have not been examined. Our results thus represent the first evidence that impaired Ccr4-Not–dependent microRNA pathway function, as well as defects of protein-synthesis repression, may be primary contributors to *EIF4A2*-related phenotypes. This is supported by observations from our colocalization assays, indicating diminished direct interactions between eIF4A2 and Ccr4-Not. We further excluded compensatory upregulation of eIF4A2's paralog eIF4A1 in c.896_897del-bearing fibroblasts, consistent with their nonredundant functions in translational control.[21] Third, we analyzed molecular correlations

between heterozygous and biallelic loss-of-function effects in patient cells, generating experimental support for the recent proposition that phenotype severity in *EIF4A2*-associated disease may be determined by residual eIF4A2 functional dosage.[18] Our phenotypic and functional data align with the concept of severe encephalopathic recessive disease in biallelic *EIF4A2* mutation carriers and milder, more variable expressions with a strong neurodevelopment component, now also encompassing movement disorders, in heterozygous carrier individuals.[18]

For the additional herein identified variants, c.470_472del (p.Val157del) and c.1144_1145del (p.-Lys382Glufs*5), patient-derived fibroblasts were unobtainable for functional analyses. We highlight, though, that c.470_472del was located in a domain where pathogenic *EIF4A2* variants have previously been documented, and that deletions of single, highly conserved amino acids are part of the genotypic spectrum of *EIF4A2*-related conditions.[18] Further studies are required to firmly establish their pathogenicity, as are studies that help to understand the mechanisms contributing to the wide range of phenotypic expressions in disorders resulting from translational dysfunction, which appears also to include nonmanifestation in heterozygous parents from recessive families.[18]

A growing number of human disease genes, including genes implicated in movement disorders, have now been associated with both dominant and recessive inheritance patterns.[36,37] Our present study adds *EIF4A2* as another movement disorder–related gene to this catalog, which may be important to consider during clinical management and counseling of affected families. Even in heterozygous carriers of *EIF4A2* loss-of-function variants, the penetrance of movement disorder manifestations may be high, as demonstrated by the identification of the herein described pedigrees. However, it is also possible that the apparently complete penetrance in our families reflects an ascertainment bias, and it should be taken into account that genetic alterations linked to highly penetrant disease traits in patient families can have much lower effect sizes in the general population.[38] How heterozygous *EIF4A2* variants lead to predominant movement disorders on the one hand and neurodevelopmental syndromes on the other remains unknown, although this breadth of clinical variability is recognized for many developmentally important genes,[39] including the functionally related *EIF2AK2* locus.[15] It could be that there are specific phenotype-determining molecular effects of the individual variants that have yet to be identified. Another hypothesis might be that there is more generally a phenotypical continuum ranging from early neurodevelopmental features to later-onset dystonia, tremor, and other movement abnormalities, occurring in relation to similar or identical mutational

mechanisms, in which genotype–phenotype correlations are defined by modulation through environment, background (epi)genetic variation, or stochastic factors.[39,40] Identifying these mechanisms underlying variable expressivity for different developmental gene-related neurological diseases should be a priority of future research. ∎

Data Availability Statement

Data available on request from the authors.

References

1. Balint B, Mencacci NE, Valente EM, et al. Dystonia. Nat Rev Dis Primers 2018;4(1):25.

2. Keller Sarmiento IJ, Mencacci NE. Genetic Dystonias: update on classification and new genetic discoveries. Curr Neurol Neurosci Rep 2021;21(3):8.

3. Gonzalez-Latapi P, Marotta N, Mencacci NE. Emerging and converging molecular mechanisms in dystonia. J Neural Transm (Vienna) 2021;128(4):483–498.

4. Advani VM, Ivanov P. Translational control under stress: reshaping the Translatome. Bioessays 2019;41(5):e1900009.

5. Sossin WS, Costa-Mattioli M. Translational Control in the Brain in Health and Disease. Cold Spring Harb Perspect Biol 2019;11(8):a032912.

6. Wiebe S, Nagpal A, Sonenberg N. Dysregulated translational control in brain disorders: from genes to behavior. Curr Opin Genet Dev 2020;65:34–41.

7. Sleiman S, Marshall AE, Dong X, et al. Compound heterozygous variants in SHQ1 are associated with a spectrum of neurological features, including early-onset dystonia. Hum Mol Genet 2022;31(4):614–624.

8. Rittiner JE, Caffall ZF, Hernandez-Martinez R, et al. Functional genomic analyses of Mendelian and sporadic disease identify impaired eIF2alpha signaling as a generalizable mechanism for dystonia. Neuron 2016;92(6):1238–1251.

9. Cagnetta R, Wong HH, Frese CK, Mallucci GR, Krijgsveld J, Holt CE. Noncanonical modulation of the eIF2 pathway controls an increase in local translation during neural wiring. Mol Cell 2019;73(3):474–489e475.

10. Bellato HM, Hajj GN. Translational control by eIF2alpha in neurons: beyond the stress response. Cytoskeleton (Hoboken) 2016;73(10):551–565.

11. Hershey JW, Sonenberg N, Mathews MB. Principles of translational control: an overview. Cold Spring Harb Perspect Biol 2012;4(12):a011528.

12. Beauvais G, Watson JL, Aguirre JA, Tecedor L, Ehrlich ME, Gonzalez-Alegre P. Efficient RNA interference-based knockdown of mutant torsinA reveals reversibility of PERK-eIF2alpha pathway dysregulation in DYT1 transgenic rats in vivo. Brain Res 2019;1706:24–31.

13. Zakirova Z, Fanutza T, Bonet J, et al. Mutations in THAP1/DYT6 reveal that diverse dystonia genes disrupt similar neuronal pathways and functions. PLoS Genet 2018;14(1):e1007169.

14. Camargos S, Scholz S, Simon-Sanchez J, et al. DYT16, a novel young-onset dystonia-parkinsonism disorder: identification of a segregating mutation in the stress-response protein PRKRA. Lancet Neurol 2008;7(3):207–215.

15. Kuipers DJS, Mandemakers W, Lu CS, et al. EIF2AK2 missense variants associated with early onset generalized dystonia. Ann Neurol 2020;89(3):485–497.

16. Mao D, Reuter CM, Ruzhnikov MRZ, et al. De novo EIF2AK1 and EIF2AK2 variants are associated with developmental delay, leukoencephalopathy, and neurologic decompensation. Am J Hum Genet 2020;106(4):570–583.

17. Musacchio T, Zech M, Reich MM, Winkelmann J, Volkmann J. A recurrent EIF2AK2 missense variant causes autosomal-dominant isolated dystonia. Ann Neurol 2021;89(6):1257–1258.

18. Paul MS, Duncan AR, Genetti CA, et al. Rare EIF4A2 variants are associated with a neurodevelopmental disorder characterized by intellectual disability, hypotonia, and epilepsy. Am J Hum Genet 2023;110(1):120–145.

19. Zech M, Jech R, Boesch S, et al. Monogenic variants in dystonia: an exome-wide sequencing study. Lancet Neurol 2020;19(11):908–918.

20. Dzinovic I, Boesch S, Skorvanek M, et al. Genetic overlap between dystonia and other neurologic disorders: a study of 1,100 exomes. Parkinsonism Relat Disord 2022;102:1–6.

21. Wilczynska A, Gillen SL, Schmidt T, et al. eIF4A2 drives repression of translation at initiation by Ccr4-not through purine-rich motifs in the 5'UTR. Genome Biol 2019;20(1):262.

22. Zech M, Kumar KR, Reining S, et al. Biallelic AOPEP loss-of-function variants cause progressive dystonia with prominent limb involvement. Mov Disord 2022;37(1):137–147.

23. Zech M, Kopajtich R, Steinbrucker K, et al. Variants in mitochondrial ATP synthase cause variable neurologic phenotypes. Ann Neurol 2022;91(2):225–237.

24. Steel D, Zech M, Zhao C, et al. Loss-of-function variants in HOPS complex genes VPS16 and VPS41 cause early onset dystonia associated with lysosomal abnormalities. Ann Neurol 2020;88(5):867–877.

25. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581(7809):434–443.

26. Kumar KR, Davis RL, Tchan MC, et al. Whole genome sequencing for the genetic diagnosis of heterogenous dystonia phenotypes. Parkinsonism Relat Disord 2019;69:111–118.

27. Meijer HA, Schmidt T, Gillen SL, et al. DEAD-box helicase eIF4A2 inhibits CNOT7 deadenylation activity. Nucleic Acids Res 2019;47(15):8224–8238.

28. Jia X, Zhang S, Tan S, et al. De novo variants in genes regulating stress granule assembly associate with neurodevelopmental disorders. Sci Adv 2022;8(33):eabo7112.

29. Li K, Ling Z, Luo T, et al. Cross-disorder analysis of De novo variants increases the power of Prioritising candidate genes. Life (Basel) 2021;11(3):233.

30. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17(5):405–424.

31. Charlesworth G, Plagnol V, Holmstrom KM, et al. Mutations in ANO3 cause dominant craniocervical dystonia: ion channel implicated in pathogenesis. Am J Hum Genet 2012;91(6):1041–1050.

32. Mencacci NE, Rubio-Agusti I, Zdebik A, et al. A missense mutation in KCTD17 causes autosomal dominant myoclonus-dystonia. Am J Hum Genet 2015;96(6):938–947.

33. Ozelius LJ, Hewett JW, Page CE, et al. The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. Nat Genet 1997;17(1):40–48.

34. Meyer E, Carss KJ, Rankin J, et al. Mutations in the histone methyl-transferase gene KMT2B cause complex early-onset dystonia. Nat Genet 2017;49(2):223–237.

35. Burgunder JM. Mechanisms underlying phenotypic variation in neurogenetic disorders. Nat Rev Neurol 2023;19(6):363–370.

36. Harel T, Yesil G, Bayram Y, et al. Monoallelic and Biallelic variants in EMC1 identified in individuals with global developmental delay, Hypotonia, scoliosis, and cerebellar atrophy. Am J Hum Genet 2016;98(3):562–570.

37. Steel D, Vezyroglou A, Barwick K, et al. Both heterozygous and homozygous loss-of-function JPH3 variants are associated with a paroxysmal movement disorder. Mov Disord 2023;38(1):155–157.

38. Kingdom R, Tuke M, Wood A, et al. Rare genetic variants in genes and loci linked to dominant monogenic developmental disorders cause milder related phenotypes in the general population. Am J Hum Genet 2022;109(7):1308–1316.

39. Dzinovic I, Winkelmann J, Zech M. Genetic intersection between dystonia and neurodevelopmental disorders: insights from genomic sequencing. Parkinsonism Relat Disord 2022;102:131–140.

40. von Scheibler E, van Eeghen AM, de Koning TJ, et al. Parkinsonism in genetic neurodevelopmental disorders: a systematic review. Mov Disord Clin Pract 2023;10(1):17–31.

SGML and CITI Use Only
DO NOT PRINT

Author Roles

P.H.: study design and concept, acquisition of data, analysis and interpretation of data, and writing of the manuscript.

M.Š.: study design and concept, acquisition of data, analysis and interpretation of data, and revision of manuscript for critical intellectual content.

V.K.: study design and concept, acquisition of data, analysis and interpretation of data, and revision of manuscript for critical intellectual content.

I.D.: acquisition of data and revision of manuscript for critical intellectual content.

F.B.: acquisition of data and revision of manuscript for critical intellectual content.

M.T.: acquisition of data and revision of manuscript for critical intellectual content.

V.M.: acquisition of data and revision of manuscript for critical intellectual content.

T.S.: acquisition of data and revision of manuscript for critical intellectual content.

M.O.: acquisition of data and revision of manuscript for critical intellectual content.

K.K.: acquisition of data and revision of manuscript for critical intellectual content.

R.B.: acquisition of data and revision of manuscript for critical intellectual content.

H.B.: acquisition of data and revision of manuscript for critical intellectual content.

F.O.: acquisition of data and revision of manuscript for critical intellectual content.

R.K.: acquisition of data and revision of manuscript for critical intellectual content.

H.P.: acquisition of data and revision of manuscript for critical intellectual content.

K.R.K.: acquisition of data and revision of manuscript for critical intellectual content.

N.E.M.: acquisition of data and revision of manuscript for critical intellectual content.

M.A.K.: acquisition of data and revision of manuscript for critical intellectual content.

A.D.F.: acquisition of data and revision of manuscript for critical intellectual content.

S.B.: acquisition of data and revision of manuscript for critical intellectual content.

A.A.K.: acquisition of data and revision of manuscript for critical intellectual content.

U.B.: acquisition of data and revision of manuscript for critical intellectual content.

K.L.: acquisition of data and revision of manuscript for critical intellectual content.

B.H.: acquisition of data and revision of manuscript for critical intellectual content.

D.W.: acquisition of data and revision of manuscript for critical intellectual content.

R.J.: acquisition of data and revision of manuscript for critical intellectual content.

J.W.: study design and concept, study supervision, analysis and interpretation of data, and revision of manuscript for critical intellectual content.

M.Z.: study design and concept, study supervision, analysis and interpretation of data, and writing of the manuscript.

Financial Disclosures

2.3 Epigenetic Association Analyses and Risk Prediction of RLS

Contributions

Concept and design of the study were developed by me together with Nazanin Mirza-Schreiber, Barbara Schormair, Juliane Winkelmann and Konrad Oexle. I coordinated and executed the extraction of the examined brain regions in collaboration with the Center for Neuropathology and Prion Research (Ludwig-Maximilians-Universität, Munich, Germany), tested and optimized different DNA/RNA and protein extraction protocols and conducted the established extraction protocol for DNA/RNA and proteins (RNA and protein analyses not published yet). I was involved in quality control of the data from blood and brain, and reviewed the data analyses (including EWAS, epigenetic clock calculations, risk profile extraction, and pathway analyses) conducted by Nazanin Mirza-Schreiber, Vanessa Mandel, Chen Zhao and Konrad Oexle, together with Chen Zhao and Konrad Oexle. Drafting of the manuscript and figures was done by me in exchange with Nazanin Mirza-Schreiber, and Konrad Oexle.

Graphical Abstract
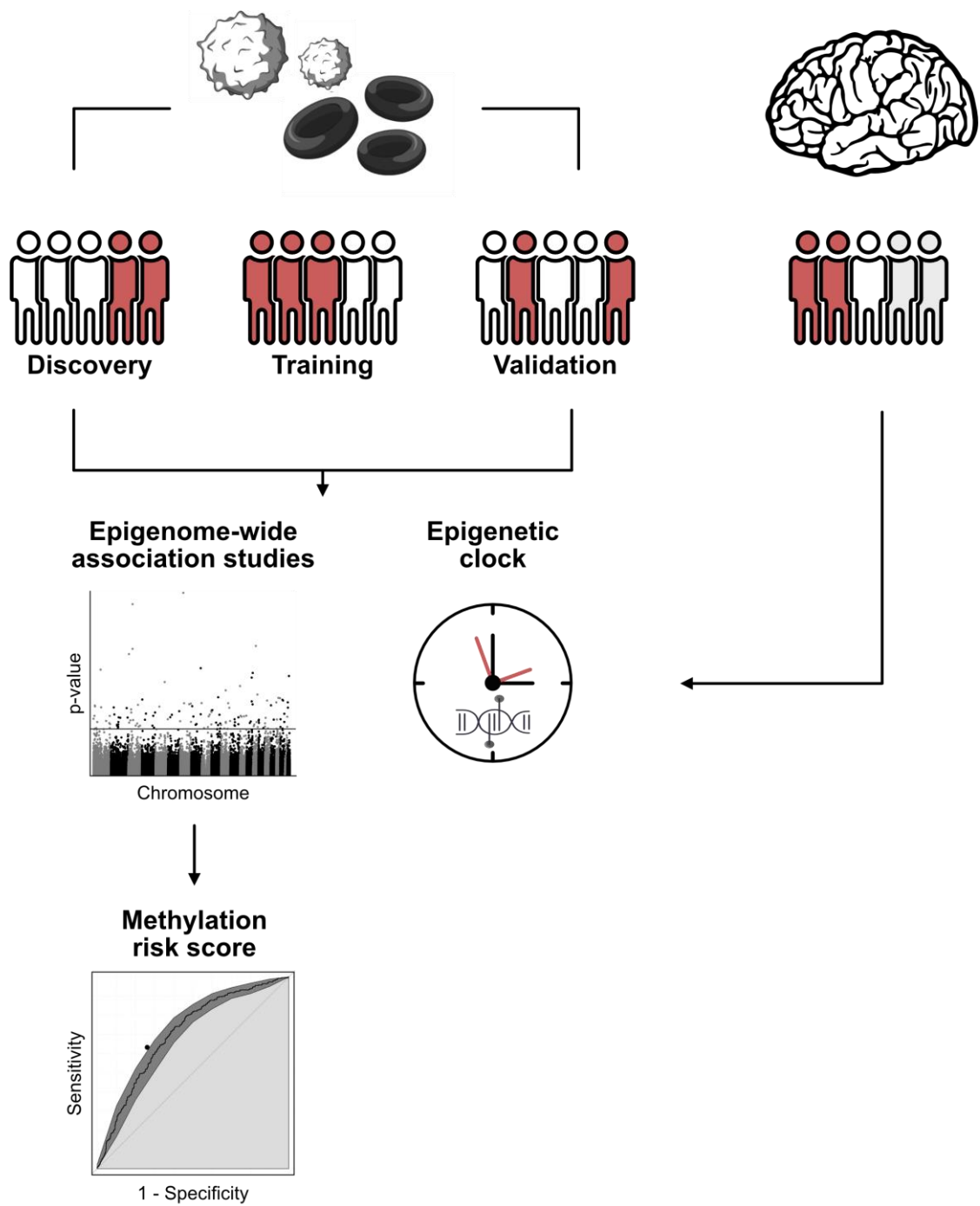
Figure 10: Graphical Abstract: Epigenetic Association Analyses and Risk Prediction of RLS

Epigenetic Association Analyses and Risk Prediction of RLS

Philip Harrer, MD,[1,2] Nazanin Mirza-Schreiber, PhD,[1,3] Vanessa Mandel, MSc,[1,2,3] Sigrun Roeber, MD,[4] Ambra Stefani, MD,[5] Shamsun Naher, MSc,[1,3] Matias Wagner, MD,[1,2] Christian Gieger, PhD,[6,7] Melanie Waldenberger, PhD,[6,7] Annette Peters, PhD,[7,8,9] Birgit Högl, MD,[5] Jochen Herms, MD,[4,10] Barbara Schormair, PhD,[1,2] Chen Zhao, PhD,[1,3] Juliane Winkelmann, MD,[1,2,10,11] and Konrad Oexle, MD[1,2,3*]

[1]Institute of Neurogenomics, Helmholtz Zentrum München, German Research Centre for Environmental Health, Munich, Germany
[2]Institute of Human Genetics, School of Medicine, Technical University of Munich, Munich, Germany
[3]Neurogenetic Systems Analysis Group, Institute of Neurogenomics, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany
[4]Center for Neuropathology and Prion Research, Ludwig-Maximilians-Universität, Munich, Germany
[5]Department of Neurology, Medical University of Innsbruck, Innsbruck, Austria
[6]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[7]Institute of Epidemiology, Helmholtz Zentrum München, German Research Centre for Environmental Health, Munich, Germany
[8]German Center for Diabetes Research (DZD), Neuherberg, Germany
[9]Chair of Epidemiology, Institute for Medical Information Processing, Biometry and Epidemiology, Medical Faculty, Ludwig-Maximilians-Universität München, Munich, Germany
[10]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
[11]Chair of Neurogenetics, School of Medicine, Technical University of Munich, Munich, Germany

ABSTRACT: Background: As opposed to other neuro-behavioral disorders, epigenetic analyses and biomarkers are largely missing in the case of idiopathic restless legs syndrome (RLS).
Objectives: Our aims were to develop a biomarker for RLS based on DNA methylation in blood and to examine DNA methylation in brain tissues for dissecting RLS pathophysiology.
Methods: Methylation of blood DNA from three independent cohorts (n = 2283) and post-mortem brain DNA from two cohorts (n = 61) was assessed by Infinium EPIC 850 K BeadChip. Epigenome-wide association study (EWAS) results of individual cohorts were combined by random-effect meta-analysis. A three-stage selection procedure (discovery, n = 884; testing, n = 520; validation, n = 879) established an epigenetic risk score including 30 CpG sites. Epigenetic age was assessed by Horvath's multi-tissue clock and Shireby's cortical clock.
Results: EWAS meta-analysis revealed 149 CpG sites linked to 136 genes (P < 0.05 after Bonferroni correction) in blood and 23 CpG linked to 18 genes in brain (false discovery rate [FDR] < 5%). Gene-set analyses of blood EWAS results suggested enrichments in brain tissue types and in subunits of the kainate-selective glutamate receptor complex. Individual candidate genes of the brain EWAS could be assigned to neurodevelopmental or metabolic traits. The blood epigenetic risk score achieved an area under the curve (AUC) of 0.70 (0.67–0.73) in the validation set, comparable to analogous scores in other neurobehavioral disorders. A significant

*Correspondence to: Prof. Dr. Konrad Oexle, Institute of Neurogenomics, Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany; E-mail: konrad.oexle@helmholtz-muenchen.de

Philip Harrer, Nazanin Mirza-Schreiber, Vanessa Mandel these authors contributed equally to this work as first authors.

Juliane Winkelmann, Konrad Oexle these authors contributed equally to this work as last authors.

difference in biological age in blood or brain of RLS patients was not detectable.

Conclusions: DNA methylation supports the notion of altered neurodevelopment in RLS. Epigenetic risk scores are reliably associated with RLS but require even higher accuracy to be useful as biomarkers. © 2023 The Authors. *Movement Disorders* published by Wiley

Introduction

Restless legs syndrome (RLS) is a circadian movement disorder with impaired sleep, depression, anxiety, and potentially increased cardiovascular risk.[1] Its prevalence is up to 10% in elderly European ancestry. The pathophysiology of RLS still has not been elucidated sufficiently. Long-term treatment is frequently unsatisfactory, indicating the importance of dissecting the pathogenesis of RLS in order to identify new entry points for treatment.[2] Genome-wide association studies (GWAS) identified common genetic risk variants within 19 risk loci accounting for 60% of the single nucleotide polymorphism (SNP)-based heritability.[3] In contrast, large-scale epigenome-wide studies (EWAS) have not been performed yet. The single previously published EWAS involved blood DNA samples of only 15 cases and 15 controls.[4] Using DNA methylation at 49 CpG dinucleotides (CpG sites), that study derived an epigenetic diagnostic score for RLS with high accuracy (area under the curve [AUC] of 87.5%). However, this diagnostic score also predicted iron deficiency anemia with a similar accuracy (83%), thus indicating that the study may have focused on symptomatic RLS resulting from iron deficiency.

We have examined the epigenome-wide DNA methylation in idiopathic RLS, performing EWAS on large sets of blood samples (1133 RLS cases and 1150 population-based controls) and on tissue samples from several brain regions (40 RLS cases and 21 controls). Thereby, we aimed for a deeper understanding of the epigenetic pathophysiology of RLS and for a biomarker that might serve in diagnostics and prediction of the disease.

Patients and Methods

RLS Patients and KORA Controls

Peripheral blood DNA was available from three independent batches (Table 1) of German or Austrian ancestry comprising 1133 mutually unrelated patients with RLS and 1150 mutually unrelated population controls with similar age and sex distribution that participated in the Cooperative Health Research in the Augsburg Region (KORA) study on the Bavarian population.[5] RLS patients were recruited in specialist clinics for movement disorders and in sleep units. RLS was diagnosed in face-to-face interviews by an expert neurologist, based on the International Restless Legs Syndrome Study Group diagnostic criteria.[6] Individuals with secondary RLS were excluded. The overall response rate of RLS patients for study participation was 90%. In KORA, the response rate among the randomly invited individuals was 75%. Shared European ancestry of RLS cases and controls was confirmed by principal component analysis and admixture supervised ancestry composition analysis of pruned genotype data (Supplementary Fig. 1 in Appendix S1). The study protocols have been approved by the responsible ethics committees. All study participants gave informed consent.

Post-Mortem Brain Samples

Brain tissue was available from 40 RLS cases and 21 controls, obtained from the Neurobiobank Munich (NBM) in two batches (Table 1) and in accordance with protocols approved by the LMU Munich Ethics Committee. Members of the German Restless Legs Syndrome Foundation were informed of the possibility of donating their brain, but no specific selection was made and no compensation or other preferential treatment was given. Written informed consent was obtained from all donors. All procedures were performed in accordance with the 1964 Declaration of Helsinki or comparable ethical standards. The brains were taken at different German hospitals, immediately frozen, and sent to the NBM for storage. All RLS cases had been diagnosed and confirmed in follow-up visits by clinical experts. Controls were selected from the NBM registry by matching for age, sex, post-mortem interval, and comorbidities. Four different brain regions—cerebellum (CB), parietal cortex (PC), caudate nucleus (NC), and putamen (PU)—were examined in each individual if available. After mechanical disruption and homogenization (Precellys® 24), DNA was extracted by Qiagens AllPrep DNA/RNA/miRNA protocol. Extracted DNA did not show any signs of degradation.

Quality Control, EWAS, and Meta-Analyses

Methylation profiling was performed by Illumina MethylationEPIC BeadChip (Illumina, San Diego, CA) according to the manufacturer's protocol. Determination

TABLE 1 Sample overview of peripheral blood and brain tissue study cohorts

Batch		N	Age (years) (mean [range])	Males (n [%])	Females (n [%])	PMI (mean [range])
Combined (blood)	Total	2283	60.45 [6–95]	786 [34.43]	1497 [65.57]	–
	Cases	1133	60.79 [6–95]	388 [34.25]	745 [65.75]	–
	Controls	1150	60.34 [34–88]	398 [34.61]	752 [65.39]	–
Batch 1 (blood)	Total	879	62.87 [35–90]	271 [30.83]	608 [69.17]	–
	Cases	426	63.54 [35–90]	129 [30.28]	297 [69.72]	–
	Controls	453	62.23 [38–88]	142 [31.35]	311 [68.65]	–
Batch 2 (blood)	Total	520	59.51 [38–83]	170 [32.69]	350 [67.31]	–
	Cases	257	60.53 [40–75]	84 [32.68]	173 [67.32]	–
	Controls	263	58.51 [38–83]	86 [32.69]	177 [67.31]	–
Batch 3 (blood)	Total	884	58.59 [6–95]	345 [39.03]	539 [60.97]	–
	Cases	450	58.33 [6–95]	175 [38.89]	275 [61.11]	–
	Controls	434	58.86 [34–83]	170 [39.17]	264 [60.83]	–
Combined (brain)	Total	61	79.16 [19–96]	24 [39.34]	37 [60.66]	48.93 [7.82–168]
	Cases	40	85.24 [63–96]	11 [27.5]	29 [72.5]	52.61 [7.82–133.5]
	Controls	21	68.17 [19–89]	13 [61.9]	8 [38.1]	42.63 [8–168]
Batch 1 (brain)	Total	44	77.86 [19–96]	18 [40.91]	26 [59.09]	53.09 [7.82–168]
	Cases	29	83.5 [63–96]	9 [31.03]	20 [68.97]	55.13 [7.82–133.5]
	Controls	15	67.33 [19–89]	9 [60]	6 [40]	49.57 [8–168]
Batch 2 (brain)	Total	17	82.68 [53.33–95.67]	6 [35.29]	11 [64.71]	38.26 [11.83–88]
	Cases	11	90.13 [81.5–95.67]	2 [18.18]	9 [81.82]	46.06 [11.83–88]
	Controls	6	70.25 [53.33–81.08]	4 [66.67]	2 [33.33%]	25.27 [12–46.7]

Abbreviation: PMI, post-mortem interval.

of methylation intensities, including background correction and normalization (*preprocessQuantile*), were done using the *minfi* package[7] running on R version 4.0.2 (R Core Team 2020). Probes with detection *P*-value>0.01, on sex chromosomes or at SNPs, with cross-reactivity, or a call rate < 0.95% were excluded, as well as samples with mean detection *P*-value>0.05 or call rate < 95%. The methylation level of each CpG site was assessed as beta-value (β) from which the *M*-value[8] was calculated as log2 ($\beta/(1-\beta)$). For epigenome-wide association study (EWAS) the *limma* package[9] was applied to three sets of peripheral blood samples from cases and controls (Table 1) with adjustment for sex, age, and Houseman-estimates of white blood cells. Covariates in the brain EWAS were post-mortem interval, sex, and age. Further, we corrected for potential technical biases by including the first 30 principal components (PCs) of control probe intensities.[10]

EWAS summary statistics of the three peripheral blood sets were subjected to meta-analysis using the R meta package[11] with a random effects model. The pooled data of the three sets contained 590,431 common CpG sites

and N = 2283 individuals. The epigenome-wide significance threshold was 8.47×10^{-8}.

In order to leverage information on the correlation between DNA methylation in blood and in brain we uploaded all CpG sites included in the blood EWAS meta-analysis to BECon (https://redgar598.shinyapps.io/BECon/). Of 28,298 sites above the 75th percentile of correlation, we selected those that reached Bonferroni-corrected significance ($P < 0.05/28,298$) in the blood EWAS meta-analysis.

Brain EWASs were performed analogously. Meta-analyses included 763,264 CpGs and 58 individuals for CB, 764,556 and 54 for NC, 761,209 and 54 for PC, and 761,069 and 58 for PU. As there was no serious inflation, we used the less conservative 5% false discovery rate (FDR) significance threshold.

CpG sites were mapped to the nearest genes according to the *IlluminaHumanMethylationEPICanno.ilm10b4.hg19* annotation file.

Phillips et al.[12] reported associations of RLS with alcohol consumption, body mass index (BMI), diabetes,

income, physical activity, and smoking. We excluded these potential confounders as described in the extended online methods in Appendix S1.

Differentially methylated regions (DMRs) were derived from the EWAS meta-analyses on blood and on the four brain regions using the ipdmr software[13] with default settings.

Methylation Risk Score

For constructing a methylation risk score in blood DNA of RLS patients, the largest dataset (batch 3) was used for discovery, batch 2 for testing, and batch 1 for validation (Table 1). Candidate CpG sites for the weighted risk score were selected from the EWAS on the discovery batch (see earlier) if they passed an FDR-threshold of 5% and an abs(logFC)-threshold of 0.3. To avoid redundant information, we calculated the pairwise correlation of all selected CpG sites. If a CpG pair had a correlation larger than 0.3, only the site with the lower P-value was kept in the risk score. 308 CpG sites passed the threshold criteria, four of them were excluded due to the correlation criterion, resulting in 304 CpGs for the testing phase. In the testing phase, we performed logistic regression of the disease state in the testing batch, using a model that included CpG sites, age sex, and the Houseman-estimates of white blood cell type composition. In order to optimize the number of CpG sites, we analyzed the receiver operating characteristic (ROC) curve of the prediction of RLS and control status in the testing batch. To do so, we increased the set of CpG sites stepwise by 10, starting with the 10 most significant, each time performing a logistic regression and calculating the AUC of the ROC. The model with the highest AUC value was selected and applied to the validation batch. For comparison, we trained a support vector machine (SVM) classifier with linear kernel (R package e1017, https://CRAN.R-project.org/package=e1071) on the selected model parameters in the testing batch and predicted the disease states in the validation batch. Sensitivity and specificity of this prediction were then entered in the ROC diagram. In order to assess the potential relevance of age, sex, and blood cell composition estimates on the accuracy of the prediction, training and application of the logistic regression model and of the SVM were also done without the CpG sites. For prediction of the brain samples' disease states using blood EWAS results, CpG sites, age, and sex were used, but not the white blood cell type estimates.

Epigenetic Age

We applied Horvath's epigenetic clock,[14] a multi-tissue estimator of DNA methylation age (DNAmAge). We calculated the difference between DNAmAge and chronological age and conducted linear regression

analysis of that difference on RLS status, combining batch results by random-effect meta-analysis. In addition, we used Shireby's cortical clock[15] to estimate the DNAmAge in the brain DNA samples. For a detailed description see extended online methods in Appendix S1.

Gene-Set Enrichment Analysis

Tissue and pathway enrichment analyses were carried out using the tools GENE2FUNC within FUMA (v1.3.8, https://fuma.ctglab.nl/gene2func/),[16] PANTHER (v.17.0, http://www.pantherdb.org/),[17] and missMethyl (v.1.32.1, https://bioconductor.org/packages/missMethyl/).[18] With these tools we accessed the databases GO_BP, GTEx (v.8),[19] GWAScatalog, KEGG, and PANTHER-pathways. Enrichment tests compare an input set (eg, the significant CpG sites or the genes associated with them) to the representations of a background set in the tissue and pathway genes sets of the analyzed databases. As background set, we used the CpG sites (or linked genes) covered by the respective EWAS which had generated the input set. In case of the EWAS meta-analysis on blood data, for instance, the background set consisted of 590,431 CpG sites (or 21,084 linked genes with unique Entrez ID). For the enrichment analyses of CpG sites with blood–brain correlation, we restricted the background to the 307,651 sites that could be assessed for such correlation at BECon (see earlier). We set the significance cut-off at FDR of 5% and a required overlap of at least 10 genes in order to match the settings of Czamara et al.[20] However, we also considered smaller overlaps if the analyzed EWAS results comprised fewer than 10 genes with Entrez ID. Analogous gene-based enrichment analyses were performed with the DMR-related gene sets.

Results

RLS EWAS Meta-Analysis of Peripheral Blood DNA

Meta-analysis of the three EWAS on peripheral blood DNA methylation in RLS patients (batches 1–3; Table 1) versus controls resulted in 149 differentially methylated CpG sites (Fig. 1), at a Bonferroni significance level of $P < 8.47 \times 10^{-8}$, that were linked to 136 genes (Supplementary Table 1 in Appendix S1). Effect sizes (logFC) ranged between −0.6 and 0.3 (Fig. 2). 146 (98%) of the significant CpG sites had a negative effect direction. Across all CpG sites, the results of the EWAS were balanced with 50.76% negative association beta-coefficients and 49.24% positive beta-coefficients.

We specifically analyzed 28,298 CpG sites whose correlation between blood DNA methylation and brain DNA methylation is above the 75th percentile according
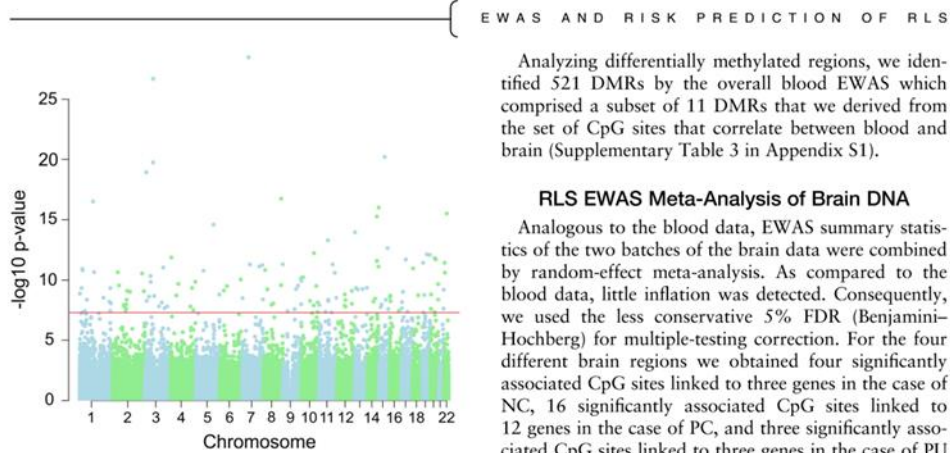
FIG. 1. Manhattan plot of the meta-analysis of restless legs syndrome (RLS) epigenome-wide association study (EWAS) on blood DNA. One hundred and forty-nine differentially methylated CpG sites were associated with RLS at epigenome-wide significance. Linked genes are listed in Supplementary Table 1 in Appendix S1. [Color figure can be viewed at wileyonlinelibrary.com]
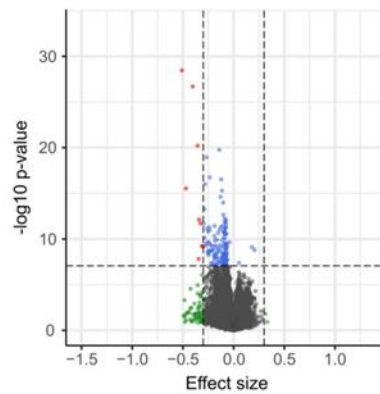
FIG. 2. Volcano plot of restless legs syndrome (RLS) epigenome-wide association study (EWAS) on blood DNA. CpG sites are represented by dots indicating effect sizes E (logFC, x-axis) and significances (y-axis) of the association between methylation and RLS. Red: Bonferroni-significant, E > 0.3; blue: Bonferroni-significant, E < 0.3; green: non-significant, E > 0.3. [Color figure can be viewed at wileyonlinelibrary.com]

to BECon (https://redgar598.shinyapps.io/BECon/). Nine of them reached Bonferroni-corrected significance ($P < 0.05/28,298$), including six that had not already been significant in the general blood EWAS meta-analysis (Supplementary Table 2 in Appendix S1). However, when the nine CpG sites were looked up in the brain EWAS results (see later), none of them reached nominal significance ($P \geq 0.05$).

Analyzing differentially methylated regions, we identified 521 DMRs by the overall blood EWAS which comprised a subset of 11 DMRs that we derived from the set of CpG sites that correlate between blood and brain (Supplementary Table 3 in Appendix S1).

RLS EWAS Meta-Analysis of Brain DNA

Analogous to the blood data, EWAS summary statistics of the two batches of the brain data were combined by random-effect meta-analysis. As compared to the blood data, little inflation was detected. Consequently, we used the less conservative 5% FDR (Benjamini–Hochberg) for multiple-testing correction. For the four different brain regions we obtained four significantly associated CpG sites linked to three genes in the case of NC, 16 significantly associated CpG sites linked to 12 genes in the case of PC, and three significantly associated CpG sites linked to three genes in the case of PU (Supplementary Table 4 in Appendix S1). CB did not yield any significant CpG sites after FDR correction. Significant CpG sites from the blood EWAS did not reach nominal significance ($P > 0.05$) in any of the brain tissues (Supplementary Table 1 and Supplementary Fig. 2 in Appendix S1).

DMRs were found in NC (4), PC (6), and CB (3). They did not overlap with the DMRs derived from the blood EWAS (Supplementary Table 3 in Appendix S1).

Gene-Set Enrichment Analysis of Blood and Brain EWAS Results

Tissue enrichment analyses using the blood EWAS meta-analysis results were significant (adjusted P-value $= 3.5 \times 10^{-2}$) for brain tissue when compared to GTEx v8 gene sets with differential overexpression ("up-regulated DEG"; Supplementary Fig. 3 in Appendix S1) by FUMA's GENE2FUNC. The restriction to upregulated DEG was made because the direction of effect was negative for most of the significant CpG sides in blood (Fig. 2), indicating hypomethylation in RLS cases with hypomethylation usually implying increased expression.

Moreover, when considering 54 more specific tissue types available in GTEx, the input genes were overrepresented in multiple CNS tissue types, including cortex ($P = 1.1 \times 10^{-3}$), frontal cortex BA9 ($P = 4.3 \times 10^{-3}$), anterior cingulate cortex BA24 ($P = 6.5 \times 10^{-3}$), substantia nigra ($P = 4.5 \times 10^{-2}$), and spinal cord cervical c-1 ($P = 1.6 \times 10^{-2}$) (Supplementary Fig. 3B in Appendix S1). However, when we tried to replicate this result using the small set of CpG sites that correlate between blood and brain, the enrichment P-values of central nervous system (CNS) regions were still on top but did not pass the Bonferroni-corrected threshold anymore, likely due to insufficient power.

In the pathway enrichment analyses, GENE2FUNC enrichment analyses of the DMRs derived from CpG sites that correlate between blood and brain identified the kainate-selective glutamate receptor complex among GO-cellular components as well as GO-molecular functions (adjusted *P*-value<0.004). This was due to differential methylation of the genes of subunits GluK2 (*GRIK2*) and GluK4 (*GRIK4*). In *GRIK4* we determined hypomethylation in patients while the DMR in *GRIK2* was hypermethylated, potentially indicating a shift in the subunit expression spectrum.

All further enrichment analyses in blood and brain EWAS did not provide significant results. However, literature-based analyses of the specific functions of individual candidate genes that came up in the brain EWASs revealed that almost all of the candidates could be assigned to neurodevelopmental or metabolic traits (Supplementary Table 5 in Appendix S1).

Methylation Risk Score

As a discovery cohort for RLS methylation risk score construction we used the largest batch (ie, batch 3) (Table 1). With thresholds on significance, effect size, and correlation of $q = 0.05$, abs(logFC) = 0.3, and

$r = 0.3$, respectively, as described in the Methods section, we selected 304 CpG sites. From them we derived an optimized subset of 30 CpG sites to be used in the risk score by stepwise adding CpG sites and performing ROC analysis in the test set (batch 2). This set of 30 CpG-sites (Supplementary Table 6 in Appendix S1) predicted RLS with the maximal area under the curve (AUC) of 70% (90% confidence interval [CI] 67%–73%; Fig. 3A) in the blood validation set (batch 1). In comparison, a model with just age, sex, and cell composition reached an AUC of only 51% (47%–54%; Fig. 3B) and did not significantly differ from chance prediction.

The RLS methylation risk score did not reliably distinguish between RLS and control status when applied to post-mortem brain tissues. Here, the 90% CIs of the AUCs always included the 50% level (Fig. 3C). Sensitivities and specificities of the predictions by an SVM resided within or at the upper border of the ROCs' 90% CI areas (Fig. 3).

Epigenetic Age

Horvath's biological age (DNAmAge) estimates in peripheral blood samples were highly correlated with
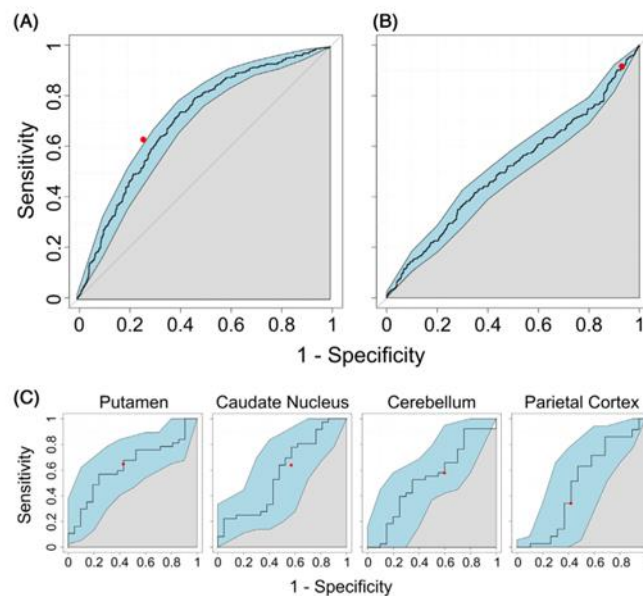
FIG. 3. Receiver operating characteristic (ROC) curves of risk score in the blood validation batch (**A**, area under the curve [AUC] = 0.70, 90% confidence interval [CI] = 0. 67–0.73) compared to a score trained without CpGs (**B**, AUC = 0.51, 90% CI = 0.47–0.54) and applied to brain tissues (**C**, AUCs not significantly different from 0.5). Red dots indicate support vector machine (SVM) prediction. [Color figure can be viewed at wileyonlinelibrary.com]

the chronological age of the individuals with $r = 0.86$ (95% CI = 0.85–0.87). Correcting for sex and chronological age, DNAmAge appeared to be 2.0 years advanced in RLS patients ($P = 2 \times 10^{-9}$) of batch 1. However, this did not replicate in batch 2 (0.3 years, $P = 0.5$) and seemingly was inverted in batch 3 (−0.5 years, $P = 0.09$). Accordingly, random-effect meta-analysis of the three batches did not produce a significant result ($P = 0.4$). Analogous analyses of the brain samples also did not produce significant results, neither in individual batches, nor in meta-analyses of the batches of each of the brain regions ($P > 0.1$), nor in meta-analysis across the two batches and the four regions (1.1 years, $P = 0.6$), nor in regression of the average difference of the four brain regions and meta-analysis of the two batches (1.4 years, $P = 0.36$). Performing the same individual batch- and meta-analyses with Shireby's cortical clock, which previously was shown to have the highest prediction accuracy for brain age assessment,[15,21] also no significant evidence of advanced age in RLS brain tissue was found ($P > 0.1$).

Discussion

To elucidate the underlying pathophysiology and for deriving diagnostic biomarkers, EWASs have been performed on several common diseases during the last decade.[22] Recent studies aimed to establish DNA methylation signatures for different neurobehavioral disorders such as depression[23] or schizophrenia[24,25] to provide supportive diagnostic tools.

The tissue-specific analysis of significant candidate genes of our blood EWAS showed an enrichment of genes differentially overexpressed in the brain. This involvement in neurological processes is in line with genetic[3] and imaging[26,27] studies which identified various CNS structures as relevant for the pathogenesis of RLS. However, we could not replicate the significant blood EWAS results by brain EWAS. This non-replication likely relates to the unavoidably low number of brain samples, that is, to the low replication power, and to the fact that blood methylation is not a good predictor of brain methylation for the vast majority of CpG sites. For less than 6.5% of the CpG sites blood DNA methylation can explain more than 20% of their methylation variance in the brain, and for less than 1.5% more than 50%.[28]

Current concepts on RLS pathophysiology focus on dopaminergic, GABAergic, and glutamatergic neurotransmitter systems[29] and brain iron metabolism.[30,31] Several genes highlighted by our EWAS on cortex, putamen, and caudate nucleus fit into these concepts, such as BRD2 (Bromodomain-containing protein 2), which is assumed to influence the neurogenesis of GABAergic neurons in the basal ganglia,[32] or

SLC25A28 (Mitoferrin 2), a mediator of mitochondrial iron ion uptake.[33] Notably, enrichment analysis of DMRs derived from CpG sites that correlate between blood and brain identified the kainate-selective glutamate receptor complex driven by the genes of the subunits GluK2 (GRIK2) and GluK4 (GRIK4). GRIK4 was hypomethylated in patients while GRIK2 was hypermethylated, suggesting opposed changes in expression. GluK2 can form homomeric or heteromeric receptors, while GluK4 is always associated with any of the GluK1-GluK3 subunits.[34] Therefore, the opposed changes in expression would change the spectrum of these receptors, potentially alter the balance of inhibitory and excitatory activity, and thus affect behavior.[35] The genes described here may be considered as candidates for functional studies. Some of them have previously been related to other neurobehavioral disorders such as schizophrenia, depression, or bipolar disorder, matching the increased rate of depression and anxiety in RLS.[36]

In order to derive a biomarker for clinical diagnosis and prediction of RLS we constructed a blood DNA methylation signature of 30 CpG sites which we then tested and validated in two independent batches. The signature was significantly associated with RLS with an AUC of 0.70 (0.67–0.73) in the validation set. Similar accuracies have been achieved by methylation risk scores in other neurobehavioral disorders such as depression (validation AUC of 0.68)[23] or schizophrenia (validation AUC of 0.69).[25] Although being significant, scores of that level of prediction accuracy need to be improved to enable clinical application. Combination with other omics levels may help to achieve that goal.[23]

Recently, Roy et al.[4] published an RLS methylation score based on 49 CpG sites. Although it was derived from only 15 cases and 15 controls, this score achieved an AUC of 87.5% in a validation set of 20 blood samples and of 73.4% in neural tissue. 39 of the 49 CpG sites could be retrieved in our blood DNA methylation dataset after quality control. Only four sites passed the nominal significance threshold of 0.05 and after Bonferroni correction for 39 tests none were significant. Besides differences in discovery sample size (15/15 vs. 450/434) and in the method of risk score construction, Roy et al.[4] may have addressed specifically iron-deficient RLS since their score also identified iron-deficient anemia with an AUC of 83% whereas the present study addressed idiopathic RLS in a large heterogeneous set of individuals (n = 2283).

RLS is an age-related disease. Therefore, we also studied the epigenetic age in our RLS patients. Epigenetic age has been proven to be valuable as a biomarker of biological aging and for prediction of morbidity and mortality.[37] Associations with markers of physical and mental fitness, and with age-accelerating effects of diet/obesity, infections, and diseases have been identified.[38]

Using Horvath's multi-tissue clock[14] we did not find any evidence of advanced epigenetic age in blood or brain samples of RLS patients. Horvarth's multi-tissue clock may not be optimally suited for assessing brain aging, however. Therefore, we applied Shireby's cortical clock[15] which has been trained on brain tissues and found to indicate advanced epigenetic aging in cortical tissue of brain samples with neurodegeneration[21] while other clocks have not been proven to detect dementia and cognitive aging sufficiently.[39] Shireby's cortical clock also did not show advanced epigenetic aging in RLS. Indeed, there also is no other evidence for an increased risk of RLS patients to develop neurodegenerative diseases.[40-42] Thus, in line with the GWAS and EWAS results, the epigenetic age analysis also underlines the concept that RLS is rather a neurodevelopmental disorder than a neurodegenerative disease.

Two relevant limitations of our study need to be considered. The design is cross-sectional and thus susceptible to bias which we cannot entirely exclude although we focused on idiopathic RLS and matched cases and controls for age, sex, and ancestry. Once population cohorts such as the UK Biobank and the German NAKO provide a sufficient number of samples with both RLS phenotypes and epigenome-wide methylation data, RLS EWAS should be conducted in these datasets. Moreover, our major results originated in the EWAS on blood DNA samples although blood most likely is not the tissue of origin of RLS. They did not replicate in the brain samples available to us, but the number of brain samples was small and they were selected from regions which may not be relevant for RLS. Nevertheless, we consider our blood results valuable. As other groups who work on the epigenetics of brain diseases such as schizophrenia[24,25] or depression,[23] we were able to derive functional epigenetic risk scores from the blood data. Blood is the tissue of choice for clinical diagnosis and prediction while brain will never be used for this purpose. Moreover, genetic causes of neurological disease may leave epigenetic traces in blood cell DNA as has been shown for various monogenic neurodevelopmental disorders.[43]

In summary, our EWAS and epigenetic aging results support the notion of altered neurodevelopment but not neurodegeneration playing a role in RLS. The epigenetic risk scores derived in blood were reliably associated with RLS but require even higher accuracy to be useful as biomarkers. ●

who are responsible for the design and conduct of that study. Open Access funding enabled and organized by Projekt DEAL.

Data Availability Statement

Data available on request due to privacy/ethical restrictions.

References

1. Manconi M, Garcia-Borreguero D, Schormair B, et al. Restless legs syndrome. Nat Rev Dis Primers 2021;7(1):80.

2. Trenkwalder C, Allen R, Högl B, et al. Comorbidities, treatment, and pathophysiology in restless legs syndrome. Lancet Neurol 2018; 17(11):994–1005.

3. Schormair B, Zhao C, Bell S, et al. Identification of novel risk loci for restless legs syndrome in genome-wide association studies in individuals of European ancestry: a meta-analysis. Lancet Neurol 2017; 16(11):898–907.

4. Roy A, Earley CJ, Allen RP, Kaminsky ZA. Developing a biomarker for restless leg syndrome using genome wide DNA methylation data. Sleep Med 2021;78:120–127.

5. Holle R, Happich M, Löwel H, Wichmann HE. KORA—a research platform for population based health research. Gesundheitswesen 2005;67(Suppl. 1):S19–S25.

6. Allen RP, Picchietti DL, Garcia-Borreguero D, et al. Restless legs syndrome/Willis-Ekbom disease diagnostic criteria: updated International Restless Legs Syndrome Study Group (IRLSSG) consensus criteria—history, rationale, description, and significance. Sleep Med 2014;15(8):860–873.

7. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 2014;30(10):1363–1369.

8. Du P, Zhang X, Huang CC, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics 2010;11:587.

9. Barbera A. Limma. Oxford: Oxford University Press; 2001.

10. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol 2015;16(1):37.

11. Schwarzer G, Carpenter J, Rücker G. Multivariate meta-analysis. Meta-Analysis with R. Berlin: Springer; 2015.

12. Phillips B, Young T, Finn L, Asher K, Hening WA, Purvis C. Epidemiology of restless legs symptoms in adults. Arch Intern Med 2000; 160(14):2137–2141.

13. Xu Z, Xie C, Taylor JA, Niu L. ipDMR: identification of differentially methylated regions with interval P-values. Bioinformatics 2021;37(5):711–713.

14. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14(10):R115.

15. Shireby GL, Davies JP, Francis PT, et al. Recalibrating the epigenetic clock: implications for assessing biological age in the human cortex. Brain 2020;143(12):3763–3775.

16. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun 2017;8(1):1826.

17. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou LP, Mi H. PANTHER: making genome-scale phylogenetics accessible to all. Protein Sci 2022;31(1):8–22.

18. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics 2016;32(2):286–288.

19. Lonsdale J, Thomas J, Salvatore M, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45(6):580–585.

20. Czamara D, Neufang A, Dieterle R, et al. Effects of stressful life-events on DNA methylation in panic disorder and major depressive disorder. Clin Epigenetics 2022;14(1):55.

21. Grodstein F, Lemos B, Yu L, et al. The association of epigenetic clocks in brain tissue with brain pathologies and common aging phenotypes. Neurobiol Dis 2021;157:105428.

22. Wei S, Tao J, Xu J, et al. Ten Years of EWAS. Adv Sci 2021;8(20):e2100727.

23. Xie Y, Xiao L, Chen L, Zheng Y, Zhang C, Wang G. Integrated analysis of methylomic and transcriptomic data to identify potential diagnostic biomarkers for major depressive disorder. Genes (Basel) 2021;12(2):178.

24. Gunasekara CJ, Hannon E, MacKay H, et al. A machine learning case-control classifier for schizophrenia based on DNA methylation in blood. Transl Psychiatry 2021;11(1):412.

25. Chen J, Zang Z, Braun U, et al. Association of a reproducible epigenetic risk profile for schizophrenia with brain methylation and function. JAMA Psychiat 2020;77(6):628–636.

26. Rizzo G, Li X, Galantucci S, Filippi M, Cho YW. Brain imaging and networks in restless legs syndrome. Sleep Med 2017;31:39–48.

27. Stefani A, Mitterling T, Heidbreder A, et al. Multimodal magnetic resonance imaging reveals alterations of sensorimotor circuits in restless legs syndrome. Sleep 2019;42(12):zsz171.

28. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. Epigenetics 2015;10(11):1024–1032.

29. Lanza G, Ferri R. The neurophysiology of hyperarousal in restless legs syndrome: hints for a role of glutamate/GABA. Adv Pharmacol 2019;84:101–119.

30. Earley CJ, Connor J, Garcia-Borreguero D, et al. Altered brain iron homeostasis and dopaminergic function in restless legs syndrome (Willis-Ekbom disease). Sleep Med 2014;15(11):1288–1301.

31. Connor JR, Patton SM, Oexle K, Allen RP. Iron and restless legs syndrome: treatment, genetics and pathophysiology. Sleep Med 2017;31:61–70.

32. Velíšek L, Shang E, Velíšková J, et al. GABAergic neuron deficit as an idiopathic generalized epilepsy mechanism: the role of BRD2 haploinsufficiency in juvenile myoclonic epilepsy. PLoS One 2011;6(8):e23656.

33. Baldauf L, Endres T, Scholz J, et al. Mitoferrin-1 is required for brain energy metabolism and hippocampus-dependent memory. Neurosci Lett 2019;713:134521.

34. Negrete-Díaz JV, Falcón-Moya R, Rodríguez-Moreno A. Kainate receptors: from synaptic activity to disease. FEBS J 2022;289(17):5074–5088.

35. Arora V, Pecoraro V, Aller MI, Román C, Paternain AV, Lerma J. Increased Grik4 gene dosage causes imbalanced circuit output and human disease-related behaviors. Cell Rep 2018;23(13):3827–3838.

36. Berger K, Kurth T. RLS epidemiology—frequencies, risk factors and methods in population studies. Mov Disord 2007;22(Suppl 18):S420–S423.

37. Gibson J, Russ TC, Clarke TK, et al. A meta-analysis of genome-wide association studies of epigenetic age acceleration. PLoS Genet 2019;15(11):e1008104.

38. Hüls A, Czamara D. Methodological challenges in constructing DNA methylation risk scores. Epigenetics 2020;15(1–2):1–11.

39. Zhou A, Wu Z, Zaw Phyo AZ, Torres D, Vishwanath S, Ryan J. Epigenetic aging as a biomarker of dementia and related outcomes: a systematic review. Epigenomics 2022;14(18):1125–1138.

40. Iranzo A, Comella CL, Santamaria J, Oertel W. Restless legs syndrome in Parkinson's disease and other neurodegenerative diseases of the central nervous system. Mov Disord 2007;22(Suppl 18):S424–S430.

41. Gan-Or Z, Alcalay RN, Rouleau GA, Postuma RB. Sleep disorders and Parkinson disease; lessons from genetics. Sleep Med Rev 2018;41:101–112.

42. Postuma RB, Iranzo A, Hu M, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. Brain 2019;142(3):744–759.

43. Levy MA, Relator R, McConkey H, et al. Functional correlation of genome-wide DNA methylation profiles in genetic neurodevelopmental disorders. Hum Mutat 2022;43(11):1609–1628.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Author Roles

(1) Research Project: A. Conception and Design, B. Data Acquisition and Analysis; (2) Statistical Analysis: A. Design, B. Execution, C. Review and Critique; (3) Manuscript Preparation: A. Drafting the Manuscript and Figures, B. Editing the Final Version; (4) Other: A. EWAS, Epigenetic Clock Calculations, Risk Profile Extraction, and Pathway Analyses.

P.H.: 1A, 1B, 2C, 3A, 3B, 4A.
N.M.S.: 1A, 1B, 2A, 2B, 3A, 3B, 4A.
V.M.: 1B, 2B, 3B, 4A.
S.R.: 1B, 3B.
A.S.: 1B, 3B.
S.N.: 1B, 3B.
Ma.W.: 1B, 3B.
C.G.: 1B, 3B.
Me.W.: 1B, 3B
A.P.: 1B, 3B.
B.H.: 1B, 3B.
J.H.: 1B, 3B.
B.S.: 1A, 1B, 3B.
C.Z.: 1B, 2A, 2B, 2C, 3B, 4A.
J.W.: 1A, 1B, 3B.
K.O.: 1A, 1B, 2A, 2B, 2C, 3A, 3B, 4A.

3.   Discussion

Genetic testing plays a crucial role for various aspects of medical decision-making, that is, diagnosing affected individuals, evaluating disease risks, predicting disease onset, providing guidance in reproductive genetic counseling, and establishing pharmacogenetic profiles for informed treatment choices. While these tasks may be similar in monogenic and polygenic traits, there exist disparities in terms of feasibility and methodological procedures between these types of disorders. Rare deleterious variants often carry more pronounced effects compared to more prevalent variants [25] [41], and most rare variants of large effect may lie in coding regions [195]. Conversely, common variants mostly reside in between genes and frequently lack direct implications on protein sequences [23], which makes it difficult to derive models that link common variants to (patho-) physiological effects.

For rare genetic variants in genes that haven't been previously described as disease-causing for the present phenotype, a conclusive diagnosis often remains absent. To obtain certainty about the cause of the disease, provide optimal care to the patient, predict the disease course, and to extend the pathophysiological understanding, functional follow-up investigations of genetic variants with unknown clinical significance are crucial [196]. In case of rare variants these follow-ups often concentrate on the disease gene and the encoded protein, while conditions in which the genetic background, environmental factors and gene-environment interactions significantly predispose the disease development [153] require to look at different omics levels as a whole, such as the epigenome to better understand the gene-environment interactions and the underlying disease mechanisms.

In this work, we studied both types of diseases, two monogenic forms of dystonia being the use case for molecular diagnostics of monogenic diseases and RLS as a common disease. In addition to genetic and epigenetic analyses we applied different molecular approaches to provide functional evidence of potential pathophysiological mechanisms and to support the diagnostic procedure.

3.1 Dystonia – identification of *NUP54* und *EIF4A2* as novel disease genes

As outlined in the respective publications, we successfully incorporated two novel genes, *NUP54* and *EIF4A2*, into the compendium of genes associated with movement disorders. Through functional follow-up experiments, we produced additional evidence in support of two new concepts on pathomechanisms in dystonia, encompassing aberrations at the nuclear envelope and facilitated shifts in the regulation of protein translation.

3.1.1 Consequences for our understanding of the pathophysiology of dystonia

Changes at the nuclear envelope

The nuclear envelope is a physical-barrier system composed of the nuclear membranes and the nuclear-pore complexes (NPCs), mediating the protection of genome integrity and maintenance of the nuclear-transport machinery [197]. Transport of large molecules through the nuclear envelope occurs prominently via NPCs, formed by multiprotein assemblies referred to as nucleoporins (NUPs) [198]. NPCs exhibit a strong degree of compositional conservation, and their functions are essential for tissue development and homeostasis [199].

In addition to the identified disease-causing variant in *NUP54* [45], the impairment of NUP62 has been previously linked to hereditary forms of early-onset dystonia with striatal lesions [200]. Interestingly, individuals with complex neurological syndromes due to impairments of other NUPs and proteins associated to the nuclear envelope may present dystonia as a prominent clinical feature [201]. Based on recent observations all major components of the nuclear envelope including individual NUPs and NPCs, are degraded by the ubiquitin-proteasome system and through lysosomes/vacuoles, triggered by protein damage, misfolding or cellular stress [202]. Integral membrane proteins require additional factors to be degraded, achieved by engaging specific adenosine triphosphatases (ATPases) of the "AAA" type ("ATPases associated with diverse cellular activities") [203]. The protein encoded by *TOR1A*, associated with the most frequent form of childhood-onset isolated dystonia [96], also belongs to the family of "AAA" ATPases.

Dysfunction leads to impairments of the nuclear export of large ribonucleoprotein granules [204], NUP delocalization and dysregulation [205], and alterations of the nuclear envelope morphology [206].

In total there are more than 30 mostly highly conserved NUPs, providing the basis for multiple different NPCs [207] [208]. The central channel of the NPC is lined by NUPs that harbor phenylalanine- and glycine-rich repeats (FG-NUPs). Despite very high transport rates FG-NUPs in the central channel facilitate a highly specific transport [209]. The impairment of a specific NPC, as present in *NUP54* [45] and *NUP62* [200] associated early onset dystonia could therefore alter the nuclear import and export of a highly specific subset of proteins and RNAs, dependent on transport-receptor mediated import and export routes connected to the diminished NPC. Studies on *KMT2B* associated dystonia outlined that deficiency of the histone methyltransferase KMT2B led to an episignature of 113 increasingly methylated DNA CpG sites in blood, that showed significant epigenome-wide association, allowing a definitive evaluation of variants of uncertain significance [137]. Following a similar approach an extension of the studies on *NUP54* and *NUP62* associated dystonia from the present targeted approach to full consideration of the epigenomic, transcriptomic or proteomic level could reveal distinct signatures based on the affected subset of proteins, improving the evaluation of potentially causal variants in these genes, and providing a perspective function in disease monitoring.

Regulation of protein translation

Considering examples of disorders of transcriptional deregulation with a predominant dystonic phenotype such as KMT2B- [170] and THAP1-related diseases [171], aberrant gene-expression regulation seems to play an important role in dystonia pathogenesis [168]. In addition to transcriptional control, regulating the availability of mRNA, the cellular rates of protein synthesis are determined through the global or mRNA specific translation regulation. The translation pathway in cells can be broken down into four stages: initiation, elongation, termination, and ribosome recycling. The synthesis rate of a specific protein is determined by the number of ribosomes translating its mRNA, the number of active mRNAs, coding length of the mRNA, rate of ribosome attachment to the mRNA (initiation), and rate of elongation [210].

Recently the eukaryotic translation initiation factor 2 alpha kinase 2 (eIF2AK2) [211, 212] and the Interferon-induced, double-stranded RNA-activated protein kinase (PRKRA) [213], a regulator of eIF2AK2 activity, have uncovered a novel link between translational regulation and dystonic phenotypes. eIF2AK2 is responsible for regulating protein synthesis by phosphorylating translation initiation factor eIF2α, which is one of the best-characterized mechanisms for regulation of both global and specific mRNA translation in response to a wide variety of different stimuli, especially in immune response [214].

Adding to the evidence of translational regulation as a potential disease mechanism in dystonia we were able to demonstrate that *EIF4A2* haploinsufficiency, encoding eukaryotic translation-initiation factor-4A-2 (eIF4A2), underlies a previously unrecognized dominant dystonia-tremor syndrome [193]. The described *EIF4A2* variants led to perturbance of translational regulation, revealing an increased expression of a target of the Ccr4-Not complex, a key regulator of eukaryotic translation [215]. This provided first evidence that impaired Ccr4-Not-dependent microRNA pathway function and defects of protein-synthesis repression might be primary contributors to eIF4A2-related phenotypes [193]. To understand the full consequence of the impaired interaction of eIF4A2 with the Ccr4-Not complex, future studies need to address the resulting changes on the transcriptome and proteome to identify a comprehensive list of proteins and pathways influenced by eIF4A2 driven translational regulation.

3.1.2 Implications for molecular diagnostic approaches of rare variants

With modern genomics more and more dystonia-related variants and associated potential disease-causing genes were identified. As inherited or acquired protein-coding variants represent the majority of disease-causing variants, accounting for upwards of 60% of all known causative genomic variants [35, 36], we used WES as a rapid and effective diagnostic tool for variant discovery. WES enabled a huge step forward in genomic testing, providing a far higher diagnostic and clinical utility in comparison to previously used chromosomal microarrays [216]. However, as WGS becomes increasingly accessible across healthcare systems (e.g., the 100.000 genomes project [217]), it might supersede other modalities in the future.

Based on the used WGS technique it can comprehensively assess multiple variant types, including structural and copy-number variants, short tandem repeats, and mitochondrial variants in a single test with shorter sample preparation times [218]. Advances in long read sequencing techniques such as PacBio HiFi sequencing provide read lengths averaging between 10 and 25kb in combination with accuracies greater than 99.5% [219], enabling the identification of larger structural variants that couldn't be identified by short-read sequencing techniques such as Illuma sequencing before [220]. Despite those advantages of different WGS approaches, however, it still relates to higher costs, partly unoptimized analysis tools and a lack of clear evidence for a significant increase in diagnostic yield [221, 222]. Interpretation and prediction of the functional effects of WGS-identified variants, in particular non-coding variants, remain difficult [223, 224]. While variant annotations are useful for identifying the relevant variants and prioritizing them for further investigation [225, 226], functional follow-up experiments are still needed to clearly determine the contribution of potential disease-causing variants to the observed phenotype and to potentially identify effects that have remained unknown so far.

3.2   RLS – development of an epigenetic risk score

To elucidate the underlying pathophysiology and derive diagnostic biomarkers, EWAS have already been conducted on various common diseases over the last decade. On one hand, analyses of altered epigenetic methylation patterns at specific sites and their associated biological pathways can aid in identifying biological correlates. On the other hand, utilization of these patterns irrespective of their biological functions can assist in constructing risk scores, offering a valuable tool for disease diagnosis and management.

3.2.1   Impact on RLS pathophysiology

Genetic and imaging studies [227, 228] mostly identified CNS structures as relevant for pathogenetic processes in RLS, including dopaminergic, GABAergic and glutamatergic neurotransmitter systems and brain iron metabolism. We could confirm these observations on the epigenetic level showing an enrichment of

differentially expressed genes in the brain in a tissue-specific analysis of signifi-
cant candidate genes in blood EWAS [194]. Several genes identified in the EWAS
on different brain tissues highlighted a possible impairment of neurodevelopment,
neurotransmission, and metabolic functions in the disease development of RLS.
Interestingly epigenetic age estimation using different mathematical models un-
derlined the concept that RLS is rather a neurodevelopmental disorder than a
neurodegenerative disease, in line with missing evidence for an increased risk of
RLS patients to develop neurodegenerative diseases [229-231], despite RLS be-
ing an age-related disease. The strongest RLS association signal in GWAS stud-
ies, MEIS1, also plays an important role in the regulation of neurogenesis [232],
further strengthening the observed concept of RLS being a neurodevelopmental
disorder that manifests through the exposure to individual and environmental fac-
tors during lifetime.

3.2.2 Diagnostic usability of epigenetic risk scores

Epigenetic modifications, particularly DNA methylation, play a crucial role in the
regulation of gene expression and have been extensively studied in the context
of various traits and diseases. Methylation risk scores have emerged as promis-
ing tools to assess disease risk and prognosis based on DNA methylation pat-
terns. However, the implementation of methylation risk scores in clinical settings
comes with several challenges and risks that must be carefully considered.

In the case of RLS, CNS structures are likely to harbor pathogenetic processes.
Blood methylation is not a good predictor of brain methylation for the vast majority
of CpG sites [233]. For less than 1.5% of the CpG sites the interindividual varia-
tion of methylation in blood can explain more than 50% of their methylation vari-
ance in brain [233]. However, we still tried to establish an epigenetic risk score
for RLS in blood, in the hope that those findings can help in the study and risk
prediction of RLS. As other groups that work on the epigenetics of brain diseases
such as schizophrenia [234, 235] or depression [236] we were able to derive a
functional epigenetic risk score from the blood data. While brain tissue is difficult
to assess and will never be used for this purpose, blood is the tissue of choice for
standardized clinical diagnosis and prediction. Moreover, it has been previously

shown for several monogenic neurodevelopmental disorders that the genetic causes may leave epigenetic traces in blood cell DNA [237].

To develop reliable risk scores, large and diverse cohorts are required for validation. Insufficient sample sizes can lead to overfitting and hinder the reproducibility of methylation risk scores across different populations [238]. In our efforts to address this challenge, we employed a three-stage selection procedure utilizing three large and independent batches, including a testing set (n = 520) in conjunction with discovery (n = 884) and validation (n = 879) sets [194]. Despite notable advancements in recent years, the performance of methylation risk scores in common genetic diseases such as schizophrenia [234, 235], depression [236], or RLS [194] still requires improvement to facilitate clinical implementation. This progress may be achieved by enhancing the scale, precision, and diversity of genetic and phenotypic data or through the integration of other omics data, such as genetic, transcriptomic, or proteomic information.

Even if epigenetic risk scores attain an accuracy suitable for clinical use, several challenges will persist. Apart from considerations regarding cost-effectiveness and affordability, risk scores often rely on intricate algorithms and mathematical models, posing challenges in interpretation for untrained healthcare providers and patients, and making a broadly assessable visualization of the results mandatory [248]. Moreover, it is crucial to consider ethical conflicts related to privacy, informed consent, and the potential stigmatization of individuals with elevated risk scores [248].

Methylation risk scores demonstrate considerable potential in the prediction of disease risk and prognosis. However, to fully capitalize on their benefits, researchers and clinicians must diligently confront the challenges and risks intrinsic to their development, validation, and clinical implementation. Ongoing research, interdisciplinary collaboration, and ethical considerations are imperative to optimize their utility.

3.3   Conclusion and perspective

A growing number of human disease genes implicated in movement disorders have been identified. We were able to add the two new genes *NUP54* and *EIF4A2* as movement disorder-related genes to this catalog and to provide first functional evidence of two potential disease mechanisms, which may be important to consider during clinical management and counseling of affected families. Regarding the common movement disorder RLS we have been able to support the notion of altered neurodevelopment but not neurodegeneration playing a role in RLS on an epigenetic level and to derive the first methylation risk score reliably associated with idiopathic RLS.

Looking ahead, the aim of molecular genetic diagnostics will be to assess a person's entire set of genetic and epigenetic risk factors – both common and rare – to optimize and individualize therapy, prognosis, and prevention. While rare variants contribute to the heritability of quantitative phenotypes such as height [24] and common traits such as breast cancer [26], common variants can influence the clinical presentations of monogenic disorders [23].

Since completion of the human genome sequencing [33], the bottleneck in human genetics shifted from the discovery of genetic variants to the identification of causal mechanisms by which genetic variants lead to a specific trait. Of more than 200,000 genetic variants that have been related to human traits by GWAS, the majority remain mechanistically uncharacterized [239]. Similarly, of nearly 1 million entries in the ClinVar database [47] of variants identified in patients with severe genetic disease, about 50% are classified as having uncertain or conflicting annotations [240].

Considering our focus on individual omics levels, application of a variety of omics technologies in parallel could complement clinical characterization and sequencing approaches for rare and common diseases. Integration of multi-omics data might promote a more comprehensive understanding of the molecular landscape, strengthening the identification of disease-causing variants by combining several lines of evidence [241, 242]. Transcriptomic or proteomic analyses allow the re-classification of potential splice or ambiguous missense variants based on

aberrantly expressed gene products [243]. Metabolomics and DNA methylation data can be used to identify molecular biomarkers with applications in cohort stratification, disease classification or disease gene prediction [244].

While these approaches improve the identification of disease-causing variants, some level of uncertainty will always remain, making a functional follow-up in model systems necessary to clearly prove the disease-causing effect of a variant. Taking the thousands of links between individual gene variants and various traits and diseases into account, new methods of high-throughput experimental assays are urgently needed to address the variant-to-function problem [240]. Recent consortia such as the "Atlas of Variant Effects" [245] aim to address this problem and upcoming methods like pooled single-cell CRISPR screens to identify target genes in cis and trans, measure dosage effects, and decipher gene-regulatory networks provide first promising approaches [246].

A systematic collection of phenotypic data is the prerequisite of any such investigation. Highlighting the significance of well-defined subgroups determined by robust and comprehensive phenotypic data, recent insights have challenged the notion of common disorders as homogenous disease entities. In case of type 2 diabetes, for instance [247], phenotypic cluster analysis of independent patient cohorts identified and replicated subgroups with significant differences in patient characteristics and risks of diabetic complications [248]. Notably, these clusters also exhibited discernible epigenetic patterns, underlining the distinctiveness of type 2 diabetes subgroups and their specific types of diabetic complications [249]. A more comprehensive, standardized phenotyping of movement disorders, such as RLS and dystonia, might similarly hold the potential to identify subgroups with group-specific molecular architectures. To achieve this, it is necessary to establish structured databases with detailed and systematic recording of phenotypic features.

4.   Supplementary material

4.1   Supplement to: Recessive *NUP54* Variants Underlie Early-Onset Dystonia with Striatal Lesions

Supplementary Online Data

Recessive *NUP54* variants underlie early-onset dystonia with striatal lesions

Table of Contents

Supplementary Online Figures

Video Legends

Supplementary Online Figures

Supplementary Online Figure 1 NUP54 plays a key role in structural stabilization of the central transport channel of the nuclear-pore complex
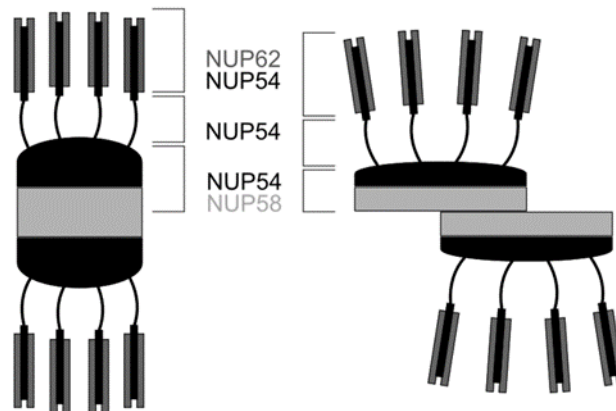
Schematic diagram adapted from Sharma et al., 2015[1]. Within the central channel of the nuclear-pore complex formed by NUP54, NUP62, and NUP58, NUP54 has a crucial function in providing plasticity to various multimerization processes. More specifically, the C-terminal coiled-coil regions of NUP54 form modular assemblies with the coiled-coils of NUP62 ("NUP54-NUP62 interactome") and NUP58 ("NUP54-NUP58 interactome"), which in turn are critically required to build "finger" (NUP54-NUP62) and "ring" (NUP54-NUP58) structures that organize the transport channel of the nuclear-pore complex. Hence, mutational changes of the coiled-coil motifs of NUP54 may deleteriously affect protein-protein interactions, thereby impairing the protein`s ability to maintain structurally important assemblies with NUP62/NUP58 within the nuclear-pore complex.

1.   Sharma A, Solmaz SR, Blobel G, Melcak I. Ordered Regions of Channel Nucleoporins Nup62, Nup54, and Nup58 Form Dynamic Complexes in Solution. J Biol Chem. 2015 Jul 24;290(30):18370-8.

Supplementary Online Figure 2 Confirmation of reduced NUP54 protein levels in patient fibroblasts by use of an independent antibody
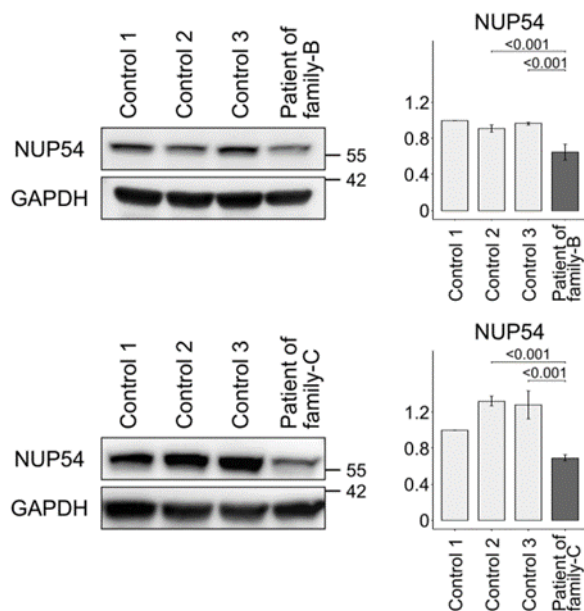
NUP54 protein levels were detected in patient fibroblasts (families B and C) by use of the NUP54-specific antibody HPA035929 (SIGMA-ALDRICH). Consistent with the data obtained in western-blot studies with the NUP54-directed antibody ab220890 (abcam; Figure 2B), a significantly decreased abundance of NUP54 was seen.

Supplementary Online Figure 3 Normal levels of the non-channel-related nuclear-pore complex component NUP153 in patient fibroblasts
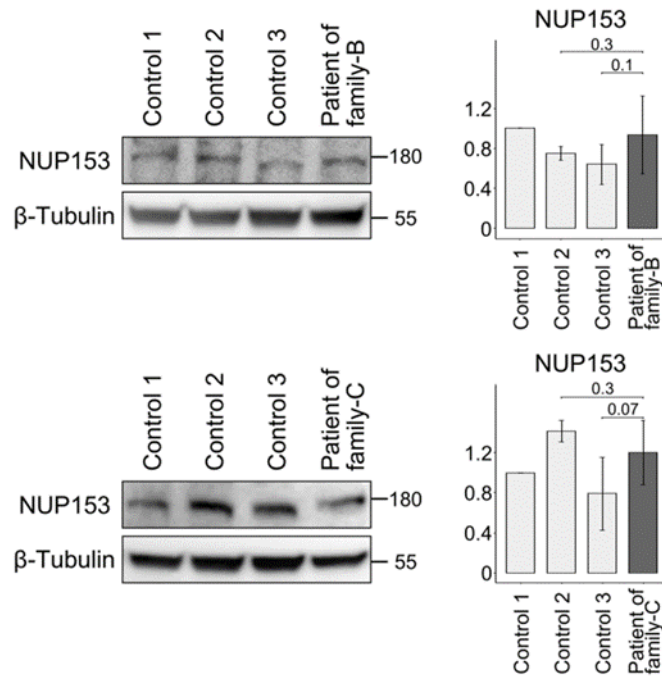
NUP153 protein levels were unaltered in patient fibroblasts (families B and C) as compared to controls, confirming a specific decrease in channel-NUPs in relation to the *NUP54* variants identified in this study.

Video Legends

Video S1: This clip demonstrates severe dystonic symptoms in the patient of family-A at last examination (age 22 years): dystonia affects the upper limbs bilaterally with strongly impaired motor performance, the trunk, and the cranio-cervical district with latero- and retrocollis; note massive involvement of the oro-bulbar region with dysarthria, dysphagia, and drooling. The patient uses preferentially a wheelchair but is able to walk short distances (see next clip Video S2).

Video S2: This clip demonstrates gait impairment in the patient of family-A at last examination (age 22 years): generalized dystonia with bilateral lower limb dystonic postures resulting in significant gait instability; an ataxic component is also seen.

Videos are available online as MPEG-4 video files (Video S1 and Video S2):

https://onlinelibrary.wiley.com/doi/full/10.1002/ana.26544

4.2 Supplement to: Epigenetic Association Analyses and Risk Prediction of RLS

Supplementary Online Data

Epigenetic association analyses and risk prediction of RLS

Table of Contents:

Supplementary figure 1

Shared genetic ancestry of RLS cases and controls.

Supplementary figure 2

Scatter plots comparing effect sizes of RLS EWASs on the 4 brain regions to the effect sizes of the RLS EWAS on blood DNA for the 149 CpG sites with epigenome-wide significance in the blood EWAS meta-analysis.

Supplementary figure 3

Gene set enrichment analysis of blood EWAS results.

Supplementary table 1

Summary statistics of meta-analysis of RLS EWASs on blood DNA.

Supplementary table 2

Summary statistics of the meta-analysis of RLS EWASs on blood DNA with the restricted set of CpG sites that show correlation between peripheral blood DNA methylation and brain tissue methylation according to BECon.

Supplementary table 3

Differentially methylated regions in blood, cerebellum, caudate nucleus, and parietal cortex.

Supplementary table 4

RLS EWASs on cerebellum, caudate nucleus, putamen, and parietal cortex.

Supplementary table 5

Functions of genes located at differentially methylated CpG sites (FDR < 5%) in the

analyzed brain regions.

Supplementary table 6

CpG sites included in the optimized methylation risk score.

Extended online methods:

Quality control, EWAS, and meta-analyses: Exclusion of potential confounders

Epigenetic age

Supplementary references
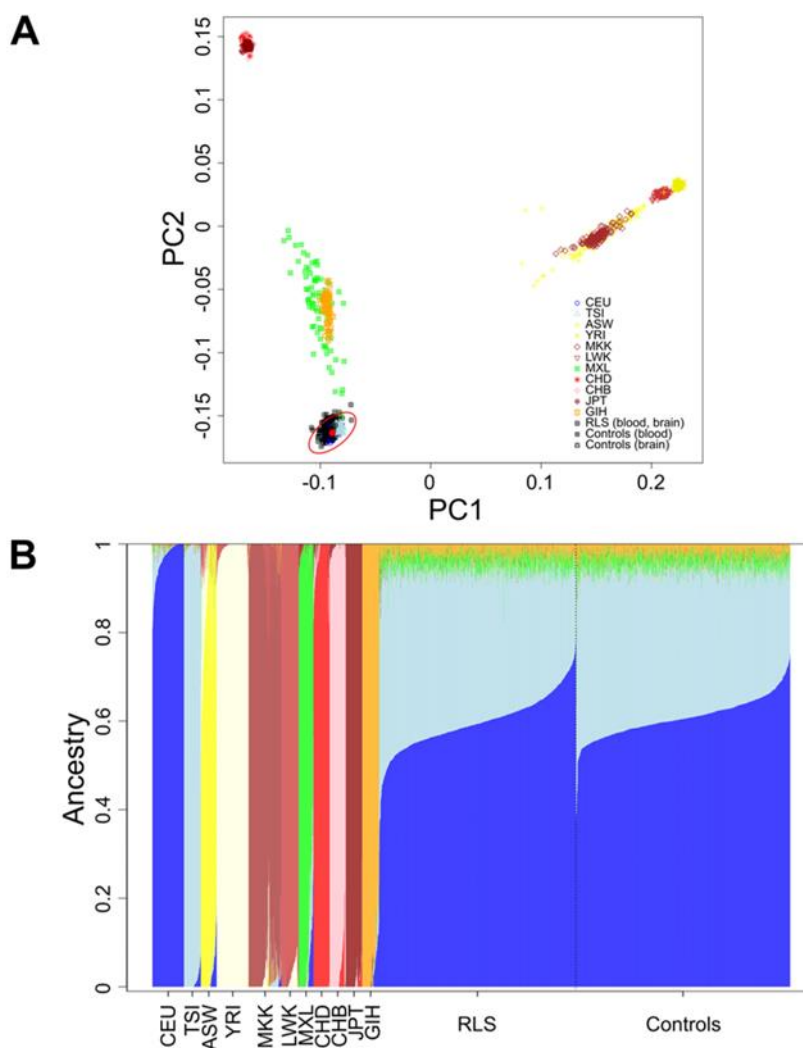
Supplementary figure 1: Shared genetic ancestry of RLS cases and KORA controls was confirmed by principal component analysis (A) and admixture supervised ancestry composition analysis (B) of pruned genotype data. 98.89% of the cases are expected to be located within the 3 SD ellipse of the 2-dimensional standard deviation[1], explaining the small number of outliers in the principal component analysis (A). The

1000 genomes project was used as reference dataset: Utah residents with Northern and Western European ancestry from the CEPH collection (CEU); Toscani in Italia (TSI); African ancestry in Southwest USA (ASW); Yoruba in Ibadan, Nigeria (YRI); Maasai in Kinyawa, Kenya (MKK); Luhya in Webuye, Kenya (LWK); Mexican ancestry in Los Angeles, California (MXL); Chinese in Metropolitan Denver, Colorado (CHD); Han Chinese in Beijing, China (CHB); Japanese in Tokyo, Japan (JPT); Gujarati Indians in Houston, Texas (GIH).
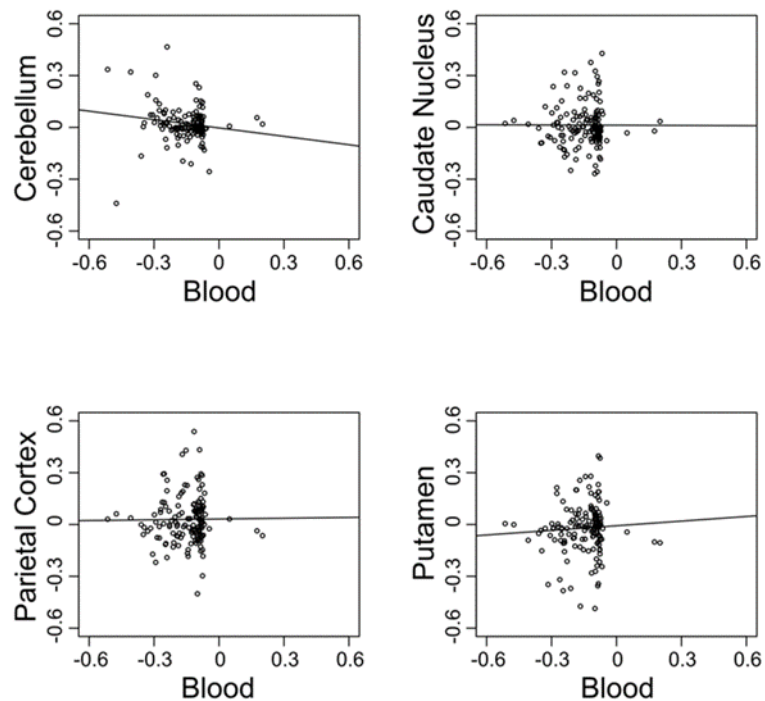
Supplementary figure 2: Scatter plots comparing effect sizes of RLS EWASs on the 4 brain regions to the effect sizes of the RLS EWAS on blood DNA for the 149 CpG sites with epigenome-wide significance in the blood EWAS meta-analysis. Lines represent the result of linear regressions which were non-significant (p > 0.05) for all brain regions.

Supplementary figure 3: Tissue specificity plots of differentially expressed gene (DEG) sets from GTEx v8 (general) tissue types generated in FUMA for the set of genes annotated to the significant CpG sites of the RLS EWAS meta-analysis on blood data. Fuma allows to detect and compare sets of genes more ("up-regulated DEG") or less expressed in a specific tissue as compared to other tissue types. The GTEx DEG sets are derived from log2-transformed and normalized (zero-mean) expression values (TPM) by t-test of each tissue type against all others. Genes are regarded as differentially expressed if they have a $|\log2\text{-fold change}| \geq 0.58$ and a significance of

p < 0.05 after Bonferroni correction. When applied to input genes such as our set of genes indicated by significant EWAS CpGs, FUMA compares them to the various DEG sets by using the hypergeometric test, thus enabling the detection of possible overrepresentation in a tissue type. Background genes in the hypergeometric test have been filtered to match the CpGs present in the EWAS meta-analysis and to have an average TPM > 1 in at least one GTEx tissue. The required minimum of input genes overlapping with a tested gene set to be reported as significant has been set to 10. P-values as indicated on the y-axis are Bonferroni corrected on the number of DEG sets tested. Red bars indicate significantly enriched DEG sets for the specific tissue indicated on the x-axis, with a Bonferroni corrected significance of less than 0.05. The enrichment analysis was limited to up-regulated DEGs because the direction of effect for most of the significant CpG sides was negative, indicating hypomethylation in RLS cases with the likely consequence of differential overexpression in RLS. **(A)** Among the general tissue types of GTEx the brain DEGs showed significant enrichment. **(B)** When looking at more specific tissue types, multiple brain tissues were significantly enriched.

Supplementary table 1: Meta-analysis of RLS EWASs on blood DNA. 149 differentially methylated CpG-sites across the genome associated with RLS with a significance below the Bonferroni threshold of 8.47×10^{-8}. RLS EWASs on the 4 brain regions (cerebellum (CB), caudate nucleus (NC), parietal cortex (PC), putamen (PU)) for these 149 sites had a minimal p-value in all brain regions of 0.003 (cg17733353 in parietal cortex), which did not pass the Bonferroni-corrected significance threshold (0.05/149 = 0.0003).

CpG	Chr	Pos	Gene	beta	SE	p.blood	p.CB	p.NC	p.PC	p.PU
cg07960450	7	45614300	ADCY1	-0.51	0.05	3.40E-29	0.962	0.785	0.112	0.309
cg12916346	3	63849521	ATXN7	-0.40	0.04	2.00E-27	0.411	0.364	0.405	0.913
cg25803961	15	63340702	TPM1	-0.36	0.04	6.30E-21	0.541	0.498	0.906	0.697
cg07753241	3	63849527	ATXN7	-0.14	0.02	1.80E-20	0.195	0.658	0.377	0.009
cg10676442	3	12525808	TSEN2	-0.27	0.03	1.20E-19	0.51	0.542	0.209	0.958
cg02898212	8	133787670	PHF20L1	-0.24	0.03	1.80E-17	0.115	0.955	0.192	0.603
cg07706463	1	101005451	GPR88	-0.12	0.02	3.00E-17	0.577	0.99	0.11	0.944
cg01097031	14	105766980	BRF1	-0.28	0.03	9.90E-17	0.229	0.352	0.094	0.741
cg14864852	22	39102110	GTPBP1	-0.47	0.06	3.00E-16	0.124	0.682	0.35	0.717
cg05039463	14	91282525	TTC7B	-0.11	0.01	5.30E-16	0.804	0.68	0.167	0.701
cg15637972	5	132112917	SEPT8	-0.13	0.02	2.60E-15	0.945	0.126	0.291	0.819
cg08195842	13	21278118	IL17D	-0.10	0.01	1.10E-14	0.898	0.033	0.976	0.753
cg07582047	11	65658459	CCDC85B	-0.29	0.04	5.00E-14	0.219	0.847	0.712	0.409
cg02839273	15	85525387	PDE8A	-0.10	0.01	2.20E-13	0.511	0.737	0.455	0.555
cg12298562	10	95361132	RBP4	-0.09	0.01	6.00E-13	0.261	0.955	0.087	0.916
cg26224223	19	31840737	TSHZ3	-0.34	0.05	7.50E-13	0.016	0.137	0.938	0.379
cg03113572	19	54057415	ZNF331	-0.07	0.01	8.20E-13	0.03	0.132	0.266	0.888
cg12094402	4	3767402	ADRA2C	-0.09	0.01	1.40E-12	0.213	0.865	0.602	0.486
cg09835867	20	31173147	NA	-0.08	0.01	1.50E-12	0.751	0.874	0.836	0.419
cg23791813	20	39317833	MAFB	-0.29	0.04	1.90E-12	0.139	0.247	0.029	0.561
cg01805124	17	30814031	CDK5R1	-0.32	0.05	1.90E-12	0.191	0.073	0.999	0.166
cg26802049	14	91884045	CCDC88C	-0.08	0.01	2.60E-12	0.983	0.474	0.172	0.937
cg14821761	22	22652529	LOC96610	-0.11	0.02	3.70E-12	0.269	0.926	0.541	0.491
cg08070578	19	2456734	LMNB2	-0.19	0.03	3.70E-12	0.682	0.131	0.684	0.04
cg13470032	9	214612	C9orf66	-0.13	0.02	4.90E-12	0.153	0.382	0.895	0.991
cg17630771	11	101981004	YAP1	-0.08	0.01	5.00E-12	0.133	0.879	0.899	0.929

cg12927203	7	56032122	GBAS	-0.24	0.04	5.00E-12	0.324	0.695	0.28	0.862
cg00871390	7	150716072	ATG9B	-0.08	0.01	5.70E-12	0.669	0.609	0.205	0.938
cg15740366	7	128049868	IMPDH1	-0.26	0.04	7.30E-12	0.914	0.945	0.182	0.061
cg02198017	14	103801143	EIF5	-0.09	0.01	7.70E-12	0.937	0.363	0.425	0.414
cg06767010	3	140770193	SPSB4	-0.14	0.02	9.10E-12	0.857	0.039	0.315	0.543
cg05179908	18	5296338	ZBTB14	-0.24	0.04	1.10E-11	0.453	0.736	0.099	0.658
cg02856606	1	19282682	IFFO2	-0.09	0.01	1.20E-11	0.148	0.232	0.927	0.031
cg04573959	8	12612631	LONRF1	-0.10	0.02	1.50E-11	0.668	0.156	0.143	0.147
cg06319359	1	22109595	USP48	-0.25	0.04	1.60E-11	0.409	0.97	0.183	0.871
cg16971831	5	56110935	MAP3K1	-0.09	0.01	1.70E-11	0.462	0.955	0.441	0.618
cg15737840	1	108507483	VAV3	-0.15	0.02	2.20E-11	0.673	0.04	0.178	0.017
cg27074297	2	39348117	SOS1	-0.09	0.01	2.30E-11	0.897	0.882	0.467	0.023
cg02706110	22	24552132	CABIN1	-0.09	0.01	2.50E-11	0.466	0.759	0.323	0.619
cg20997773	11	117667841	DSCAML1	-0.07	0.01	2.50E-11	0.865	0.184	0.441	0.629
cg16975959	2	223184510	NA	-0.07	0.01	2.90E-11	0.164	0.294	0.614	0.636
cg24822001	3	43732329	ABHD5	-0.19	0.03	4.60E-11	0.975	0.568	0.961	0.549
cg00753739	10	11653489	USP6NL	-0.10	0.02	5.90E-11	0.14	0.989	0.963	0.543
cg17733353	9	136857387	VAV2	-0.29	0.04	7.20E-11	0.056	0.958	0.003	0.097
cg01112643	6	157802216	ZDHHC14	-0.09	0.01	9.00E-11	0.79	0.731	0.517	0.685
cg13716829	9	140196785	NRARP	-0.16	0.03	1.20E-10	0.587	0.679	0.521	0.502
cg16092956	4	185747098	ACSL1	-0.11	0.02	1.40E-10	0.006	0.4	0.291	0.075
cg24652001	3	45635702	LIMD1	-0.18	0.03	1.40E-10	0.998	0.551	0.521	0.894
cg22105146	4	56262445	TMEM165	-0.25	0.04	2.10E-10	0.31	0.521	0.293	0.988
cg17334845	20	57463572	GNAS	-0.04	0.01	2.20E-10	0.615	0.965	0.21	0.893
cg22676654	8	133787679	PHF20L1	-0.24	0.04	2.60E-10	0.116	0.593	0.534	0.764
cg14556391	2	47748407	KCNK12	-0.07	0.01	2.80E-10	0.229	0.043	0.906	0.006
cg23303764	9	115248973	KIAA1958	-0.27	0.04	2.90E-10	0.557	0.405	0.811	0.571
cg05238375	10	112836780	ADRA2A	-0.08	0.01	3.10E-10	0.774	0.601	0.08	0.873
cg21271198	6	53213658	ELOVL5	-0.09	0.02	3.20E-10	0.42	0.501	0.896	0.399
cg06825833	13	111566802	ANKRD10	-0.07	0.01	4.40E-10	0.684	0.7	0.558	0.649
cg22574825	13	29068987	FLT1	-0.22	0.04	4.60E-10	0.581	0.335	0.029	0.642
cg11382291	4	160024674	NA	-0.13	0.02	4.60E-10	0.063	0.776	0.452	0.632
cg10503751	1	14075866	PRDM2	-0.18	0.03	4.60E-10	0.912	0.898	0.774	0.66
cg11637191	1	231557525	EGLN1	-0.08	0.01	4.70E-10	0.513	0.365	0.14	0.23
cg11361121	13	78272390	SLAIN1	-0.27	0.04	5.60E-10	0.999	0.448	0.437	0.114
cg12862706	9	100396088	NCBP1	-0.31	0.05	5.60E-10	0.437	0.487	0.325	0.417
cg04039925	8	29952812	LEPROTL1	-0.22	0.04	5.90E-10	0.723	0.279	0.511	0.489

cg17921034	13	42846375	AKAP11	-0.10	0.02	6.10E-10	0.487	0.367	0.462	0.379
cg21945459	21	47706156	MCM3AP	-0.30	0.05	6.50E-10	0.683	0.741	0.894	0.66
cg13583911	6	30614788	C6orf136	0.18	0.03	7.50E-10	0.361	0.527	0.639	0.291
cg09246016	12	112819890	HECTD4	-0.18	0.03	8.00E-10	0.512	0.756	0.974	0.093
cg27005794	2	109746251	LOC100287216	-0.06	0.01	8.00E-10	0.502	0.796	0.864	0.451
cg07232033	2	236403031	AGAP1	-0.14	0.02	8.50E-10	0.324	0.189	0.319	0.508
cg04420083	2	131792524	ARHGEF4	-0.10	0.02	9.20E-10	0.643	0.96	0.168	0.897
cg14589466	17	9548894	USP43	-0.07	0.01	9.50E-10	0.943	0.784	0.617	0.525
cg04672450	1	860185	SAMD11	-0.08	0.01	1.30E-09	0.998	0.558	0.774	0.864
cg03607359	5	176057465	SNCB	-0.20	0.03	1.60E-09	0.794	0.869	0.502	0.541
cg06708432	12	57916558	MBD6	-0.15	0.03	1.60E-09	0.028	0.766	0.132	0.974
cg23948071	20	3767605	CENPB	0.21	0.03	1.60E-09	0.847	0.971	0.258	0.513
cg27020941	19	10679777	CDKN2D	-0.21	0.04	1.70E-09	0.576	0.259	0.494	0.025
cg11913496	18	43914226	RNF165	-0.14	0.02	1.80E-09	0.377	0.756	0.601	0.351
cg24657817	4	42153708	BEND4	-0.25	0.04	1.90E-09	0.326	0.211	0.84	0.235
cg04996852	16	3930112	CREBBP	-0.24	0.04	1.90E-09	0.592	0.223	0.31	0.611
cg03024619	3	71803308	GPR27	-0.23	0.04	2.10E-09	0.667	0.88	0.653	0.511
cg17029062	3	32859445	TRIM71	-0.07	0.01	2.30E-09	0.539	0.351	0.496	0.307
cg10554340	16	639207	RAB40C	-0.09	0.01	2.60E-09	0.883	0.369	0.566	0.646
cg27449836	1	179846920	TOR1AIP2	-0.16	0.03	2.90E-09	0.528	0.817	0.032	0.123
cg04910677	11	129245689	BARX2	-0.26	0.04	2.90E-09	0.225	0.837	0.262	0.719
cg16839083	7	193229	FAM20C	-0.11	0.02	3.20E-09	0.691	0.958	0.031	0.948
cg06932612	14	77965123	ISM2	-0.24	0.04	3.40E-09	0.492	0.217	0.552	0.061
cg24452135	7	154794643	PAXIP1	-0.17	0.03	4.00E-09	0.174	0.903	0.748	0.707
cg10011623	20	57463527	GNAS	-0.10	0.02	4.10E-09	0.701	0.519	0.131	0.465
cg11343870	11	67888858	CHKA	-0.07	0.01	4.20E-09	0.998	0.453	0.545	0.891
cg18663513	1	53793486	LRP8	-0.07	0.01	4.20E-09	0.295	0.91	0.8	0.169
cg17078686	2	105471965	POU3F3	-0.07	0.01	4.30E-09	0.584	0.707	0.354	0.463
cg08294267	15	72565024	PARP6	-0.09	0.02	4.40E-09	0.927	0.471	0.826	0.992
cg14221884	12	63328524	PPM1H	-0.09	0.02	4.50E-09	0.974	0.713	0.296	0.645
cg27430682	10	65225552	JMJD1C	-0.18	0.03	5.30E-09	0.608	0.432	0.819	0.038
cg01428769	3	15248316	DVWA	-0.10	0.02	5.30E-09	0.355	0.309	0.752	0.502
cg27431247	11	67888603	CHKA	-0.20	0.03	5.50E-09	0.028	0.19	0.445	0.366
cg02454264	16	4897716	GLYR1	-0.25	0.04	5.70E-09	0.769	0.104	0.302	0.952
cg22940211	14	64971470	ZBTB25	-0.11	0.02	6.10E-09	0.652	0.464	0.339	0.996
cg00034433	1	224803837	CNIH3	-0.08	0.01	6.20E-09	0.657	0.752	0.757	0.899
cg08299976	17	1552302	RILP	-0.10	0.02	6.80E-09	0.839	0.787	0.461	0.966

cg22310653	10	123734262	NSMCE4A	-0.08	0.01	7.20E-09	0.551	0.622	0.67	0.831
cg24770850	3	129325130	PLXND1	-0.10	0.02	7.40E-09	0.792	0.914	0.909	0.665
cg04703844	8	74207197	RPL7	-0.08	0.01	8.70E-09	0.924	0.956	0.632	0.9
cg15955291	1	236445184	ERO1LB	-0.10	0.02	8.70E-09	0.896	0.556	0.87	0.371
cg03159671	12	112123840	ACAD10	-0.17	0.03	9.40E-09	0.025	0.577	0.826	0.917
cg13742648	9	138987085	NACC2	-0.19	0.03	1.00E-08	0.128	0.573	0.431	0.014
cg24751648	2	103236124	SLC9A2	-0.08	0.01	1.00E-08	0.959	0.703	0.694	0.547
cg01683505	15	40650336	DISP2	-0.11	0.02	1.10E-08	0.713	0.582	0.392	0.969
cg21798926	8	61429570	RAB2A	-0.19	0.03	1.10E-08	0.827	0.723	0.45	0.452
cg07442657	2	85198334	KCMF1	-0.08	0.01	1.10E-08	0.342	0.7	0.935	0.421
cg11945871	11	1411012	BRSK2	-0.10	0.02	1.20E-08	0.31	0.936	0.556	0.85
cg25390656	10	92980550	PCGF5	-0.14	0.02	1.30E-08	0.777	0.301	0.836	0.104
cg05013783	12	56652058	ANKRD52	-0.09	0.02	1.40E-08	0.999	0.421	0.564	0.732
cg09643587	3	107809710	CD47	-0.35	0.06	1.50E-08	0.571	0.857	0.954	0.467
cg12897942	3	170626357	EIF5A2	-0.08	0.01	1.50E-08	0.729	0.229	0.259	0.25
cg11722766	2	86564644	REEP1	-0.08	0.01	1.50E-08	0.891	0.32	0.99	0.274
cg06802374	4	148653449	ARHGAP10	-0.11	0.02	1.60E-08	0.78	0.79	0.884	0.843
cg26964202	12	63328896	PPM1H	-0.17	0.03	1.60E-08	0.164	0.168	0.77	0.36
cg26732930	15	32162819	NA	-0.06	0.01	1.80E-08	0.112	0.622	0.708	0.866
cg06569284	3	150128235	TSC22D2	-0.08	0.02	2.20E-08	0.586	0.48	0.86	0.238
cg08785922	17	882831	NXN	-0.07	0.01	2.20E-08	0.906	0.007	0.688	0.198
cg02114786	21	46293672	PTTG1IP	-0.07	0.01	2.40E-08	0.699	0.495	0.609	0.838
cg22334142	2	102314462	MAP4K4	-0.11	0.02	2.40E-08	0.615	0.923	0.98	0.34
cg10649494	20	2673604	EBF4	-0.07	0.01	2.50E-08	0.715	0.355	0.941	0.264
cg06991829	15	90293883	MESP1	-0.09	0.02	2.60E-08	0.58	0.415	0.618	0.14
cg00636641	9	89561390	GAS1	-0.09	0.02	2.60E-08	0.331	0.76	0.65	0.438
cg04821107	4	170947779	MFAP3L	-0.15	0.03	3.20E-08	0.613	0.781	0.453	0.883
cg03974985	2	98612359	TMEM131	-0.20	0.04	3.40E-08	0.572	0.418	0.982	0.482
cg02988046	1	231557915	EGLN1	-0.07	0.01	3.70E-08	0.725	0.032	0.648	0.53
cg07487786	16	68119375	NFATC3	-0.10	0.02	3.70E-08	0.631	0.389	0.822	0.156
cg06380157	1	39875049	KIAA0754	-0.09	0.02	3.70E-08	0.874	0.29	0.721	0.95
cg19325947	9	35665192	ARHGEF39	-0.14	0.03	3.80E-08	0.66	0.547	0.268	0.923
cg06759215	17	42786055	DBF4B	0.05	0.01	4.20E-08	0.353	0.657	0.949	0.446
cg24317637	11	47429873	SLC39A13	-0.09	0.02	4.30E-08	0.278	0.789	0.518	0.689
cg16953106	20	31350405	DNMT3B	-0.22	0.04	4.40E-08	0.762	0.996	0.521	0.96
cg16030006	10	123357754	FGFR2	-0.12	0.02	4.50E-08	0.41	0.205	0.032	0.905
cg20596348	1	20960202	PINK1	-0.16	0.03	4.60E-08	0.877	0.579	0.808	0.524

cg06505271	19	5623257	SAFB2	-0.08	0.02	4.80E-08	0.448	0.57	0.952	0.291
cg00796728	11	94276898	FUT4	-0.08	0.02	4.80E-08	0.024	0.564	0.042	0.361
cg25907916	10	81205166	ZCCHC24	-0.09	0.02	5.50E-08	0.74	0.54	0.096	0.486
cg10010376	1	144932311	PDE4DIP	-0.09	0.02	5.60E-08	0.015	0.891	0.642	0.658
cg24592364	5	14871736	ANKH	-0.09	0.02	5.80E-08	0.822	0.907	0.389	0.734
cg19306864	1	10270515	KIF1B	-0.10	0.02	5.80E-08	0.522	0.991	0.25	0.712
cg16076997	1	47905067	FOXD2	-0.08	0.02	6.30E-08	0.255	0.33	0.321	0.635
cg03214094	14	93581708	ITPK1	-0.21	0.04	6.50E-08	0.451	0.518	0.624	0.682
cg26911804	13	111806221	ARHGEF7	-0.18	0.03	6.70E-08	0.719	0.056	0.933	0.314
cg08083016	6	144471643	STX11	-0.08	0.02	6.80E-08	0.677	0.926	0.122	0.992
cg13931559	20	33146515	MAP1LC3A	-0.06	0.01	7.70E-08	0.714	0.904	0.309	0.319
cg23262020	14	93651423	C14orf109	-0.17	0.03	8.50E-08	0.508	0.749	0.525	0.193

Supplementary table 2: Using BECon (https://redgar598.shinyapps.io/BECon/), 9 out of 28,298 CpG sites correlated between blood and brain at the level above the 75th percentile of all sites examined at BECon and showed Bonferroni-corrected significance (p < 0.05/28,298) in the RLS EWAS on DNA methylation in blood. In the EWASs on the 4 brain regions (cerebellum, caudate nucleus, parietal cortex, putamen), none of the 9 CpG sites showed significant association.

CpG	Chr	Pos	Gene	beta	SE	p.blood	p.CB	p.NC	p.PC	p.PU
cg13559259	18	30050338	GAREM1	-0.10	0.02	1.7E-06	0.76	0.71	0.41	0.81
cg15174564	11	120856801	GRIK4	-0.26	0.05	1.1E-06	0.85	0.71	0.39	0.1
cg23506135	6	31922463	NELFE	0.06	0.01	6.0E-07	0.94	0.14	0.85	0.37
cg16975959	2	223184510	NA	-0.07	0.01	2.9E-11	0.16	0.29	0.61	0.64
cg24452135	7	154794643	PAXIP1-AS1; PAXIP1	-0.17	0.03	4.0E-09	0.17	0.9	0.75	0.71
cg25390656	10	92980550	PCGF5	-0.14	0.02	1.3E-08	0.78	0.3	0.84	0.1
cg01656470	19	57350096	PEG3; ZIM2	-0.06	0.01	4.2E-07	0.56	0.05	0.17	0.64
cg05884032	18	76740088	SALL3	-0.14	0.03	7.8E-07	0.67	0.92	0.55	0.44
cg04266460	16	11348956	SOCS1	-0.20	0.04	4.0E-07	0.52	0.37	0.79	0.8

Supplementary table 3: Differentially methylated regions (DMRs) in blood, cerebellum, caudate nucleus, and parietal cortex as derived from the respective RLS EWASs using the ipdmr tool[2] with default settings. At the end of the table, DMRs with correlation between blood and brain above the 75th percentile are displayed.

Chr	Pos (start)	Pos (end)	p	FDR
		Blood		
20	57463526	57463615	6.60E-20	2.81E-17
3	63849239	63849756	1.08E-19	2.81E-17
1	101005120	101005509	1.03E-18	1.80E-16
7	45614289	45614720	1.82E-16	2.37E-14
5	110062383	110062780	4.21E-16	4.39E-14
14	91884027	91884049	1.03E-14	8.92E-13
1	231556681	231557917	9.95E-14	7.41E-12
1	108507339	108507764	1.28E-13	7.63E-12
14	103801127	103801183	1.45E-13	7.63E-12
17	30813980	30814514	1.46E-13	7.63E-12
18	43914225	43914264	2.04E-13	9.67E-12
9	100396068	100396098	2.60E-13	1.13E-11
8	133787542	133787679	1.06E-12	4.24E-11
7	150715830	150716072	1.16E-12	4.30E-11
9	140196250	140196794	2.57E-12	8.94E-11
16	2390351	2391081	4.95E-12	1.61E-10
3	129325119	129325130	5.53E-12	1.70E-10
2	103236104	103236124	6.17E-12	1.79E-10
10	11653475	11653489	6.54E-12	1.79E-10
9	90341040	90341158	8.14E-12	2.12E-10
5	176730609	176730787	1.02E-11	2.52E-10
4	42153694	42153708	1.80E-11	4.26E-10
19	31840736	31840740	3.08E-11	6.93E-10
21	45138794	45138843	3.19E-11	6.93E-10
2	236402328	236403704	3.42E-11	7.13E-10
22	22652039	22652537	3.68E-11	7.37E-10
1	14075790	14076293	5.06E-11	9.76E-10
12	63328523	63328896	5.89E-11	1.10E-09
12	129308175	129308273	9.87E-11	1.77E-09
19	10828927	10828972	1.29E-10	2.24E-09
19	16738997	16739131	1.73E-10	2.91E-09
2	86564635	86564659	1.86E-10	3.03E-09
2	10588824	10588877	2.09E-10	3.31E-09
14	64971469	64971476	2.29E-10	3.51E-09
15	66995984	66996086	2.45E-10	3.64E-09
10	92980408	92980627	3.17E-10	4.59E-09
3	52479104	52479108	3.26E-10	4.59E-09

19	50979745	50979755	3.51E-10	4.78E-09
4	148653442	148653539	3.58E-10	4.78E-09
3	58477621	58477628	4.00E-10	5.13E-09
12	57916557	57916571	4.04E-10	5.13E-09
16	2040016	2040023	4.22E-10	5.24E-09
12	107712209	107713740	4.49E-10	5.44E-09
17	79885239	79885904	4.87E-10	5.77E-09
15	40650268	40650635	5.44E-10	6.30E-09
15	63340699	63340840	5.67E-10	6.42E-09
1	22109592	22109595	5.83E-10	6.46E-09
1	231114437	231114534	7.40E-10	8.03E-09
4	90228966	90229039	8.52E-10	9.06E-09
15	85524777	85525673	8.74E-10	9.10E-09
16	11348955	11349023	1.28E-09	1.30E-08
15	72565015	72565024	1.37E-09	1.37E-08
6	144329908	144329962	1.47E-09	1.42E-08
3	12525801	12525846	1.47E-09	1.42E-08
9	136857243	136857387	1.55E-09	1.46E-08
5	176057464	176057465	1.57E-09	1.46E-08
13	78271875	78272393	1.61E-09	1.47E-08
16	3930111	3930112	1.88E-09	1.69E-08
10	65225543	65225552	2.41E-09	2.13E-08
11	43964097	43964451	2.56E-09	2.22E-08
20	34287143	34287216	3.27E-09	2.79E-08
14	93581707	93581717	3.34E-09	2.81E-08
11	65657894	65658826	4.09E-09	3.38E-08
1	41327640	41327924	4.31E-09	3.51E-08
1	53793466	53793486	4.52E-09	3.62E-08
6	157801995	157802429	4.62E-09	3.64E-08
3	15248315	15248316	5.32E-09	4.14E-08
2	39347985	39348152	5.50E-09	4.21E-08
16	639203	639209	6.00E-09	4.53E-08
4	159689900	159690372	6.11E-09	4.54E-08
1	6662477	6662645	6.31E-09	4.63E-08
7	154794624	154795251	6.46E-09	4.67E-08
9	140083167	140083171	7.28E-09	5.20E-08
5	139028148	139028156	8.41E-09	5.92E-08
11	9595190	9595200	9.20E-09	6.39E-08
12	112123839	112123840	9.35E-09	6.39E-08
16	4897377	4897918	9.45E-09	6.39E-08
10	79397345	79397421	1.11E-08	7.41E-08
18	76740087	76740091	1.20E-08	7.87E-08
5	14871735	14871845	1.21E-08	7.87E-08
5	132112890	132112923	1.45E-08	9.36E-08
3	170626356	170626357	1.54E-08	9.77E-08
15	65067695	65068326	1.61E-08	1.01E-07

1	19282681	19283102	1.76E-08	1.09E-07
1	85930808	85930863	1.84E-08	1.13E-07
17	48585852	48585858	1.89E-08	1.15E-07
5	6714654	6714664	2.06E-08	1.24E-07
3	150128234	150128235	2.15E-08	1.28E-07
17	882830	882831	2.20E-08	1.29E-07
15	68871404	68871409	2.25E-08	1.30E-07
2	98612358	98612370	2.46E-08	1.40E-07
11	1410749	1411018	2.47E-08	1.40E-07
4	185747097	185747409	2.52E-08	1.40E-07
10	81204568	81205166	2.55E-08	1.40E-07
15	90293882	90293883	2.56E-08	1.40E-07
1	94375051	94375064	2.66E-08	1.44E-07
6	53213605	53213754	2.95E-08	1.59E-07
4	160024474	160024674	3.00E-08	1.59E-07
2	38152298	38152302	3.08E-08	1.60E-07
9	35665142	35665306	3.08E-08	1.60E-07
8	29952760	29952812	3.16E-08	1.63E-07
11	2905495	2905931	3.29E-08	1.68E-07
7	56032111	56032364	3.32E-08	1.68E-07
13	25254703	25254742	3.67E-08	1.83E-07
16	68119374	68119375	3.70E-08	1.83E-07
14	101291135	101291180	3.99E-08	1.96E-07
1	151810585	151810887	4.10E-08	2.00E-07
20	31350404	31350405	4.44E-08	2.14E-07
19	5623256	5623257	4.84E-08	2.31E-07
3	156392411	156392435	4.92E-08	2.33E-07
1	10270514	10270515	5.80E-08	2.71E-07
2	109746169	109746691	5.83E-08	2.71E-07
1	35247052	35247055	6.15E-08	2.83E-07
1	39874873	39875069	6.32E-08	2.87E-07
1	47905066	47905067	6.34E-08	2.87E-07
3	45635692	45635702	6.58E-08	2.95E-07
13	111806220	111806221	6.70E-08	2.98E-07
6	144471642	144471643	6.77E-08	2.99E-07
2	64246938	64247030	7.40E-08	3.24E-07
16	88772657	88772686	7.72E-08	3.33E-07
20	31173014	31173147	7.74E-08	3.33E-07
2	242576817	242577535	7.89E-08	3.37E-07
11	94276897	94276903	8.10E-08	3.41E-07
21	47706089	47706161	8.16E-08	3.41E-07
4	154387493	154387603	8.18E-08	3.41E-07
14	93651415	93651423	8.76E-08	3.61E-07
11	117667836	117667862	8.80E-08	3.61E-07
15	91414592	91414844	9.00E-08	3.66E-07
2	102314137	102314465	9.36E-08	3.78E-07

1	236445063	236445184	1.03E-07	4.11E-07
4	8229158	8229159	1.05E-07	4.18E-07
14	105766977	105766987	1.06E-07	4.18E-07
9	115248955	115248990	1.11E-07	4.37E-07
3	50712232	50712233	1.18E-07	4.60E-07
10	131589140	131589189	1.24E-07	4.80E-07
10	115439305	115439306	1.26E-07	4.82E-07
10	123357694	123357754	1.27E-07	4.82E-07
18	43753898	43753899	1.28E-07	4.83E-07
7	132261764	132262058	1.34E-07	4.99E-07
14	53019544	53019545	1.34E-07	4.99E-07
11	134201951	134201993	1.39E-07	5.12E-07
1	230202494	230203043	1.40E-07	5.14E-07
10	123734240	123734694	1.45E-07	5.29E-07
17	79818536	79818537	1.50E-07	5.42E-07
1	224803836	224803858	1.51E-07	5.43E-07
15	72523563	72523569	1.53E-07	5.46E-07
17	73975159	73975160	1.59E-07	5.62E-07
14	95623698	95623699	1.61E-07	5.65E-07
2	131792520	131792528	1.75E-07	6.07E-07
11	13690104	13690122	1.76E-07	6.07E-07
14	45366563	45366564	1.76E-07	6.07E-07
1	171810467	171810490	1.77E-07	6.07E-07
5	78280881	78281103	1.84E-07	6.26E-07
11	73019336	73019501	1.89E-07	6.38E-07
12	8185343	8185526	1.97E-07	6.59E-07
8	74207182	74207199	1.98E-07	6.59E-07
5	56110880	56111630	1.99E-07	6.59E-07
14	100259351	100259804	2.10E-07	6.92E-07
8	82023928	82024104	2.24E-07	7.33E-07
15	32162818	32162823	2.36E-07	7.69E-07
17	4269554	4269695	2.44E-07	7.88E-07
12	32552278	32552279	2.64E-07	8.42E-07
17	41856237	41856238	2.65E-07	8.42E-07
4	2965194	2965203	2.65E-07	8.42E-07
17	36507991	36508002	2.67E-07	8.44E-07
1	9970256	9970257	2.74E-07	8.57E-07
10	92922570	92922571	2.75E-07	8.57E-07
11	46401701	46401702	2.79E-07	8.64E-07
14	102228215	102228221	2.94E-07	9.07E-07
5	177631660	177631661	3.03E-07	9.20E-07
19	47363944	47363949	3.03E-07	9.20E-07
10	3824386	3824516	3.04E-07	9.20E-07
10	123872500	123872501	3.08E-07	9.26E-07
1	42501598	42501599	3.12E-07	9.35E-07
4	56262268	56262445	3.15E-07	9.37E-07

9	89561174	89561826	3.20E-07	9.48E-07
19	50145353	50145354	3.35E-07	9.87E-07
2	85197807	85198553	3.43E-07	1.00E-06
1	45265702	45265852	3.48E-07	1.01E-06
9	110252211	110252512	3.68E-07	1.07E-06
12	110152299	110152300	3.86E-07	1.11E-06
3	13008957	13008958	4.06E-07	1.16E-06
13	28194656	28194706	4.10E-07	1.17E-06
13	21277705	21278287	4.29E-07	1.21E-06
3	138327706	138327728	4.36E-07	1.23E-06
12	10875368	10875369	4.38E-07	1.23E-06
17	5342413	5342543	4.51E-07	1.26E-06
11	4208722	4208723	4.69E-07	1.30E-06
7	98476370	98476371	4.70E-07	1.30E-06
1	86042673	86042934	4.74E-07	1.30E-06
7	105925538	105925539	4.76E-07	1.30E-06
1	77747874	77747875	4.94E-07	1.34E-06
10	99160095	99160130	4.95E-07	1.34E-06
12	40499810	40499811	5.28E-07	1.42E-06
9	138986378	138987085	5.31E-07	1.42E-06
1	28908370	28908371	5.39E-07	1.43E-06
13	100634629	100634630	5.46E-07	1.44E-06
11	71935848	71935881	5.60E-07	1.47E-06
7	151722620	151722623	5.61E-07	1.47E-06
17	79952527	79952786	5.74E-07	1.49E-06
7	127744388	127744389	5.76E-07	1.49E-06
19	37328979	37329090	5.94E-07	1.53E-06
22	39101983	39102445	6.00E-07	1.54E-06
1	42801022	42801087	6.01E-07	1.54E-06
19	19496566	19496667	6.19E-07	1.57E-06
10	99094218	99094219	6.55E-07	1.66E-06
4	149365875	149365887	6.59E-07	1.66E-06
22	46466182	46467123	6.79E-07	1.70E-06
17	9548830	9548898	6.96E-07	1.73E-06
20	40247208	40247231	6.97E-07	1.73E-06
3	71803097	71803308	7.22E-07	1.78E-06
20	60529383	60529758	7.30E-07	1.79E-06
3	14444066	14444133	7.44E-07	1.82E-06
10	134807901	134808159	7.82E-07	1.90E-06
17	48167095	48167096	8.14E-07	1.97E-06
11	75479824	75479866	8.25E-07	1.98E-06
1	182584177	182584188	8.25E-07	1.98E-06
2	74648753	74648754	8.35E-07	1.99E-06
5	133512409	133512410	8.40E-07	1.99E-06
9	214611	214915	8.42E-07	1.99E-06
2	201172047	201172481	8.46E-07	1.99E-06

20	18118440	18118441	8.51E-07	2.00E-06
9	124581997	124582156	8.60E-07	2.01E-06
4	54243730	54243733	8.91E-07	2.07E-06
19	17858451	17858612	9.01E-07	2.09E-06
17	8092554	8092555	9.13E-07	2.11E-06
5	176245015	176245078	9.21E-07	2.11E-06
10	91174410	91174411	9.21E-07	2.11E-06
12	44229710	44229711	9.28E-07	2.11E-06
5	72112458	72112466	9.43E-07	2.14E-06
12	112819494	112819890	9.52E-07	2.15E-06
1	111743329	111743368	9.70E-07	2.18E-06
20	13202436	13202437	1.03E-06	2.29E-06
16	21178332	21178859	1.03E-06	2.30E-06
3	66024553	66024554	1.06E-06	2.34E-06
16	128493	128494	1.06E-06	2.34E-06
17	76968749	76968750	1.08E-06	2.38E-06
3	179041356	179041357	1.14E-06	2.48E-06
2	36583116	36583536	1.14E-06	2.48E-06
19	57862553	57862554	1.18E-06	2.56E-06
10	64134154	64134160	1.19E-06	2.57E-06
18	24237059	24237060	1.21E-06	2.58E-06
10	112836687	112836785	1.21E-06	2.58E-06
1	234614966	234614967	1.21E-06	2.58E-06
21	43299788	43299789	1.22E-06	2.58E-06
17	76879171	76880071	1.26E-06	2.66E-06
8	61591938	61591939	1.27E-06	2.68E-06
4	6675789	6675881	1.30E-06	2.72E-06
7	149571110	149571111	1.30E-06	2.72E-06
2	239335446	239335466	1.30E-06	2.72E-06
17	42786044	42786676	1.35E-06	2.81E-06
16	89764254	89764255	1.39E-06	2.87E-06
1	228674906	228674966	1.41E-06	2.89E-06
4	152330106	152330107	1.43E-06	2.94E-06
6	139456247	139456526	1.46E-06	2.97E-06
2	32234976	32235240	1.48E-06	3.01E-06
12	111843938	111843939	1.53E-06	3.10E-06
4	926321	926338	1.56E-06	3.16E-06
11	4486541	4486542	1.57E-06	3.16E-06
6	35695933	35695934	1.58E-06	3.16E-06
7	26415909	26415910	1.60E-06	3.20E-06
17	17380296	17380297	1.64E-06	3.26E-06
20	39317481	39318100	1.65E-06	3.27E-06
17	1551940	1552302	1.66E-06	3.27E-06
7	26331465	26331472	1.67E-06	3.28E-06
12	122710555	122710569	1.68E-06	3.28E-06
18	6730035	6730036	1.70E-06	3.33E-06

6	42952411	42952517	1.75E-06	3.40E-06
1	7740805	7741065	1.80E-06	3.48E-06
7	157483334	157483448	1.81E-06	3.50E-06
8	103666228	103666233	1.82E-06	3.50E-06
6	159125488	159125489	1.84E-06	3.52E-06
1	19812289	19812290	1.87E-06	3.57E-06
21	46293651	46293672	1.88E-06	3.57E-06
3	123168385	123168386	1.92E-06	3.62E-06
18	59561550	59561651	1.92E-06	3.62E-06
17	80477463	80477464	1.93E-06	3.62E-06
1	109584693	109584746	1.95E-06	3.65E-06
20	3154101	3154102	2.00E-06	3.74E-06
15	83378320	83378321	2.02E-06	3.76E-06
8	101912083	101912261	2.07E-06	3.82E-06
22	45098340	45098346	2.07E-06	3.82E-06
2	148602104	148602399	2.08E-06	3.82E-06
19	42544586	42544587	2.14E-06	3.92E-06
3	107809694	107809956	2.18E-06	3.99E-06
3	125239026	125239030	2.22E-06	4.03E-06
18	30050325	30050338	2.22E-06	4.03E-06
1	93427208	93427209	2.24E-06	4.05E-06
2	242641475	242641476	2.26E-06	4.07E-06
1	58715498	58715499	2.28E-06	4.10E-06
6	24360542	24360600	2.29E-06	4.11E-06
10	11060047	11060104	2.31E-06	4.11E-06
6	166796699	166796706	2.37E-06	4.21E-06
16	9186056	9186057	2.38E-06	4.21E-06
18	9913781	9913801	2.38E-06	4.21E-06
19	47290715	47291295	2.44E-06	4.27E-06
12	32908770	32908817	2.44E-06	4.27E-06
7	99036698	99036756	2.45E-06	4.27E-06
19	8591363	8591721	2.45E-06	4.27E-06
8	67525281	67525282	2.47E-06	4.29E-06
17	4046030	4046223	2.48E-06	4.29E-06
1	179846836	179846920	2.49E-06	4.29E-06
12	121163850	121163887	2.51E-06	4.31E-06
5	94956174	94956900	2.53E-06	4.33E-06
4	184020451	184020710	2.56E-06	4.37E-06
17	27139573	27139574	2.60E-06	4.42E-06
16	89939441	89939442	2.61E-06	4.42E-06
11	76091835	76092112	2.63E-06	4.42E-06
6	30614787	30614794	2.63E-06	4.42E-06
14	73925367	73925372	2.63E-06	4.42E-06
17	40831788	40831789	2.66E-06	4.45E-06
4	25864317	25864318	2.66E-06	4.45E-06
1	145610874	145610886	2.68E-06	4.46E-06

12	31744042	31744043	2.69E-06	4.46E-06
7	94285872	94285887	2.72E-06	4.50E-06
14	68141348	68141349	2.76E-06	4.55E-06
17	74582139	74582140	2.78E-06	4.56E-06
21	38936158	38936403	2.81E-06	4.61E-06
9	99180431	99180611	2.83E-06	4.61E-06
4	119808934	119809865	2.84E-06	4.61E-06
19	10613491	10613492	2.84E-06	4.61E-06
2	241508107	241508464	2.90E-06	4.69E-06
8	61429569	61429582	2.92E-06	4.71E-06
1	860021	860185	2.95E-06	4.74E-06
6	139013756	139013757	2.96E-06	4.74E-06
22	43411100	43411101	2.98E-06	4.76E-06
14	105781311	105781366	2.99E-06	4.76E-06
3	51429233	51429234	3.06E-06	4.86E-06
1	95007225	95007331	3.07E-06	4.86E-06
15	67546978	67546979	3.09E-06	4.89E-06
1	193074553	193074554	3.15E-06	4.95E-06
7	192608	193229	3.19E-06	5.00E-06
2	75061233	75061841	3.20E-06	5.01E-06
19	40971808	40971809	3.22E-06	5.02E-06
11	67797974	67798212	3.25E-06	5.05E-06
19	40697424	40697601	3.26E-06	5.05E-06
17	73512854	73512855	3.27E-06	5.05E-06
15	80352473	80352476	3.31E-06	5.09E-06
5	134240498	134240499	3.35E-06	5.15E-06
20	8000167	8000168	3.40E-06	5.21E-06
1	68516271	68517177	3.43E-06	5.24E-06
10	21462652	21462743	3.45E-06	5.25E-06
11	67275718	67275721	3.47E-06	5.28E-06
7	16793337	16793477	3.55E-06	5.36E-06
2	43451558	43451775	3.55E-06	5.36E-06
6	108489063	108489208	3.58E-06	5.38E-06
6	101846790	101846791	3.60E-06	5.40E-06
2	230758151	230758152	3.63E-06	5.41E-06
2	205410272	205410387	3.63E-06	5.41E-06
12	62997128	62997233	3.66E-06	5.45E-06
13	29068566	29069100	3.74E-06	5.55E-06
20	34203706	34203965	3.77E-06	5.59E-06
6	42858549	42858550	3.82E-06	5.64E-06
19	10679774	10679784	3.87E-06	5.70E-06
14	88459470	88459488	3.91E-06	5.74E-06
12	113415944	113416055	3.92E-06	5.74E-06
1	110162476	110162485	3.96E-06	5.77E-06
11	10315731	10315754	3.98E-06	5.79E-06
20	54933930	54933931	4.03E-06	5.84E-06

13	25946346	25946397	4.05E-06	5.87E-06
4	78741142	78741208	4.08E-06	5.89E-06
16	50187608	50188019	4.18E-06	6.01E-06
10	13389394	13389395	4.19E-06	6.01E-06
8	145727317	145727318	4.40E-06	6.28E-06
12	68726426	68726427	4.40E-06	6.28E-06
1	231004184	231004314	4.43E-06	6.30E-06
17	32907001	32907388	4.45E-06	6.31E-06
3	23244680	23244688	4.46E-06	6.31E-06
4	171011098	171011257	4.56E-06	6.44E-06
19	46272952	46272953	4.60E-06	6.47E-06
4	186064567	186064648	4.65E-06	6.54E-06
8	66754113	66754114	4.67E-06	6.54E-06
20	33146370	33146515	4.68E-06	6.54E-06
4	123843780	123843788	4.75E-06	6.61E-06
19	10514299	10514353	4.76E-06	6.62E-06
14	48095791	48095792	4.84E-06	6.71E-06
20	60717733	60718270	4.98E-06	6.88E-06
17	8906600	8906601	4.99E-06	6.88E-06
12	26275049	26275311	5.02E-06	6.89E-06
11	67888602	67888949	5.04E-06	6.89E-06
13	41885656	41886418	5.06E-06	6.89E-06
9	34523252	34523253	5.06E-06	6.89E-06
14	39901090	39901098	5.07E-06	6.89E-06
4	153456175	153456176	5.11E-06	6.94E-06
14	77607525	77607526	5.18E-06	7.00E-06
17	31203786	31203787	5.19E-06	7.00E-06
17	46676098	46676215	5.26E-06	7.07E-06
18	5296330	5296340	5.27E-06	7.07E-06
16	84651520	84651521	5.35E-06	7.16E-06
7	138145030	138145433	5.50E-06	7.35E-06
20	44717988	44718168	5.56E-06	7.41E-06
12	45609978	45610071	5.64E-06	7.50E-06
1	169075390	169075400	5.81E-06	7.71E-06
6	144606244	144606480	5.84E-06	7.71E-06
11	67803770	67803951	5.84E-06	7.71E-06
3	143691722	143691725	5.90E-06	7.77E-06
19	47551841	47551842	5.93E-06	7.78E-06
1	51701802	51702069	5.98E-06	7.83E-06
11	65549440	65549441	6.00E-06	7.83E-06
10	116301353	116301354	6.05E-06	7.89E-06
21	44394742	44394797	6.12E-06	7.93E-06
19	57350003	57350096	6.14E-06	7.93E-06
22	46068329	46068330	6.16E-06	7.93E-06
19	13056744	13056745	6.18E-06	7.93E-06
17	34891259	34891272	6.18E-06	7.93E-06

17	65821598	65821599	6.18E-06	7.93E-06
12	56651911	56652058	6.28E-06	8.04E-06
2	73145005	73145744	6.37E-06	8.13E-06
1	90460917	90460949	6.44E-06	8.20E-06
11	128737200	128737319	6.46E-06	8.20E-06
19	18698824	18699118	6.48E-06	8.20E-06
10	181228	181232	6.48E-06	8.20E-06
16	2265124	2265388	6.51E-06	8.21E-06
18	44497501	44497604	6.59E-06	8.29E-06
15	29722599	29722600	6.61E-06	8.29E-06
10	82167763	82167774	6.62E-06	8.29E-06
16	81349060	81349483	6.68E-06	8.31E-06
17	14200589	14200790	6.68E-06	8.31E-06
20	43374631	43374860	6.68E-06	8.31E-06
15	73075934	73075975	6.77E-06	8.39E-06
6	31922462	31922556	6.78E-06	8.39E-06
2	223183931	223185049	6.90E-06	8.52E-06
13	111566801	111567049	6.92E-06	8.52E-06
1	207226455	207226467	6.94E-06	8.52E-06
22	24551899	24552379	6.95E-06	8.53E-06
17	54671145	54671343	7.02E-06	8.58E-06
5	111755480	111755514	7.10E-06	8.66E-06
19	54629093	54629249	7.14E-06	8.69E-06
3	195913800	195914017	7.18E-06	8.72E-06
18	12334510	12334548	7.29E-06	8.83E-06
6	89855993	89856484	7.35E-06	8.89E-06
3	141944515	141944519	7.53E-06	9.08E-06
1	180123795	180123798	7.69E-06	9.23E-06
20	2673201	2673604	7.69E-06	9.23E-06
5	139422524	139422625	7.91E-06	9.47E-06
13	44544576	44544993	8.11E-06	9.69E-06
19	16682353	16682861	8.32E-06	9.92E-06
11	11643413	11643485	8.41E-06	1.00E-05
14	91282280	91282597	8.46E-06	1.00E-05
13	30169481	30169727	8.47E-06	1.00E-05
3	183542913	183543146	8.70E-06	1.03E-05
11	2291263	2291347	8.71E-06	1.03E-05
7	27186992	27187102	8.78E-06	1.03E-05
12	132414753	132414800	8.79E-06	1.03E-05
5	149737302	149737331	8.84E-06	1.03E-05
2	172967619	172967665	8.91E-06	1.04E-05
19	18111388	18111623	9.05E-06	1.05E-05
2	105471964	105472193	9.45E-06	1.10E-05
8	26149054	26149300	9.46E-06	1.10E-05
8	94712809	94713206	9.83E-06	1.14E-05
3	128879835	128879943	9.86E-06	1.14E-05

20	3766960	3767625	9.93E-06	1.14E-05
16	1402055	1402058	9.95E-06	1.14E-05
6	159049625	159049761	9.95E-06	1.14E-05
13	113242513	113242518	9.99E-06	1.14E-05
14	74058485	74058654	1.01E-05	1.16E-05
14	51297391	51297401	1.03E-05	1.18E-05
11	120856800	120856918	1.03E-05	1.18E-05
7	155174253	155174508	1.05E-05	1.19E-05
1	2461412	2461610	1.05E-05	1.19E-05
2	36825272	36825350	1.06E-05	1.19E-05
14	50698468	50698663	1.06E-05	1.19E-05
3	66550917	66550922	1.08E-05	1.22E-05
8	23191080	23191110	1.08E-05	1.22E-05
17	19437156	19437181	1.11E-05	1.24E-05
4	170947394	170947779	1.11E-05	1.24E-05
19	16394498	16394574	1.11E-05	1.24E-05
13	45150989	45151437	1.12E-05	1.24E-05
10	32635412	32635814	1.12E-05	1.24E-05
15	28340111	28340213	1.12E-05	1.24E-05
20	3185357	3185377	1.13E-05	1.24E-05
11	3862740	3862931	1.13E-05	1.25E-05
6	35436164	35436189	1.14E-05	1.26E-05
13	77460297	77460322	1.16E-05	1.27E-05
5	177540737	177541051	1.16E-05	1.27E-05
11	73308990	73308997	1.18E-05	1.29E-05
11	2398772	2398776	1.19E-05	1.30E-05
19	7660968	7660977	1.20E-05	1.31E-05
13	113622707	113622738	1.20E-05	1.31E-05
16	3156380	3156573	1.21E-05	1.32E-05
14	77964711	77965284	1.22E-05	1.32E-05
1	51984743	51984752	1.22E-05	1.32E-05
20	35374155	35374581	1.25E-05	1.35E-05
20	57415954	57415978	1.28E-05	1.38E-05
11	75350497	75350940	1.28E-05	1.38E-05
10	71905956	71906310	1.29E-05	1.38E-05
19	35417878	35417911	1.29E-05	1.38E-05
7	100730577	100730929	1.31E-05	1.40E-05
8	12612370	12612811	1.34E-05	1.43E-05
5	131826475	131826479	1.34E-05	1.43E-05
13	42846304	42846648	1.37E-05	1.45E-05
17	66287912	66287915	1.41E-05	1.50E-05
2	88355057	88355401	1.42E-05	1.50E-05
2	3522681	3522905	1.43E-05	1.51E-05
11	117186895	117187002	1.44E-05	1.52E-05
14	100150997	100151561	1.46E-05	1.53E-05
14	74207855	74208165	1.47E-05	1.54E-05

1	46726949	46727745	1.49E-05	1.56E-05
1	144932310	144932369	1.51E-05	1.57E-05
11	695491	695533	1.51E-05	1.57E-05
4	183838735	183838739	1.52E-05	1.58E-05
17	18758444	18758553	1.54E-05	1.60E-05
9	94712479	94712484	1.54E-05	1.60E-05
19	663092	663107	1.54E-05	1.60E-05
9	34646549	34646610	1.55E-05	1.60E-05
1	62902661	62902681	1.56E-05	1.60E-05
11	8284311	8284746	1.56E-05	1.60E-05
3	43732324	43732571	1.71E-05	1.75E-05
10	95360813	95361208	1.71E-05	1.75E-05
3	140770012	140770306	1.93E-05	1.97E-05
7	128049862	128049885	2.17E-05	2.21E-05
11	129245567	129245691	2.82E-05	2.87E-05
19	18391328	18391883	2.91E-05	2.96E-05
2	47748041	47748839	3.13E-05	3.17E-05
3	32859437	32859587	7.25E-05	7.34E-05
11	47429712	47429939	8.49E-05	8.57E-05
11	101980806	101981050	9.10E-05	9.17E-05
1	20959825	20960202	0.000191	0.000192
19	54057207	54058085	0.000217	0.000217
19	2455898	2457077	0.000229	0.00023
4	3767072	3768078	0.000293	0.000293
Caudate Nucleus (NC)			
16	78539953	78540172	7.24E-13	2.90E-12
2	193255	193393	1.82E-08	3.65E-08
9	127562982	127563028	7.15E-08	9.54E-08
14	101158094	101158159	1.02E-07	1.02E-07
Parietal Cortex (PC)			
15	91473166	91473569	9.83E-13	5.90E-12
4	154049211	154049212	1.81E-10	5.44E-10
10	101380333	101380334	7.33E-10	1.47E-09
15	33653961	33653962	1.37E-08	2.06E-08
6	28186476	28186477	5.16E-08	6.20E-08
17	72620159	72620261	6.41E-08	6.41E-08
Cerebellum (CB)			
17	78636898	78639001	1.17E-17	3.52E-17
13	114061811	114062109	3.31E-08	4.97E-08
7	140178813	140178937	9.28E-08	9.28E-08
Correlated DMRs between blood and brain tissues			**Gene**
2	223184509	223184510	2.85E-11	3.14E-10	NA
7	154794642	154794643	3.98E-09	2.19E-08	PAXIP1-AS1; PAXIP1
10	92980549	92980550	1.27E-08	4.65E-08	PCGF5
14	101291135	101291180	3.80E-07	7.75E-07	MEG3
16	11348955	11348956	4.03E-07	7.75E-07	SOCS1

19	57350095	57350096	4.23E-07	7.75E-07	PEG3; ZIM2
6	31922462	31922463	6.03E-07	9.48E-07	NELFE
18	76740087	76740088	7.78E-07	1.07E-06	SALL3
11	120856800	120856801	1.05E-06	1.29E-06	GRIK4
18	30050337	30050338	1.67E-06	1.84E-06	GAREM
6	101846790	101846791	3.60E-06	3.60E-06	GRIK2

Supplementary table 4: RLS EWASs on cerebellum, caudate nucleus, putamen, and parietal cortex (each being a meta-analysis of two batches). Overall, 23 CpG-sites showed epigenome-wide significance after FDR < 5% correction.

CpG	Chr	Pos	Gene	beta	SE	p.meta	FDR
Caudate Nucleus (NC)							
cg03333546	19	13263065	IER2	0.99	0.18	2.32E-08	0.0177
cg26825544	5	169539492	NA	0.53	0.10	8.00E-08	0.0306
cg02748089	16	78540172	WWOX	-0.53	0.10	1.45E-07	0.0369
cg06573254	5	174909270	SFXN1	-0.57	0.11	2.14E-07	0.0409
Putamen (PU)							
cg23540272	14	80744035	DIO2-AS1	-0.72	0.13	5.25E-08	0.0400
cg17156867	15	95019541	MCTP2	0.40	0.07	5.93E-08	0.0225
cg16262357	19	36266374	SNX26	0.65	0.12	1.22E-07	0.0308
Parietal Cortex (PC)							
cg21221497	4	154049212	NA	0.64	0.10	1.81E-10	0.0001
cg15007325	10	101380334	SLC25A28	-0.89	0.14	7.33E-10	0.0003
cg08052045	15	33653962	RYR3	0.47	0.08	1.37E-08	0.0035
cg16499284	6	28186477	LOC222699	-0.45	0.08	5.16E-08	0.0098
cg07638320	13	20986392	CRYL1	-0.74	0.14	7.06E-08	0.0107
cg05276175	6	32940312	BRD2	-0.67	0.13	2.20E-07	0.0279
cg03506372	11	60897603	NA	-0.56	0.11	2.55E-07	0.0277
cg01000056	10	61927009	ANK3	0.72	0.14	3.20E-07	0.0305
cg16979479	7	39998010	CDK13	-0.66	0.13	3.55E-07	0.0300
cg25518366	5	75671504	NA	-0.80	0.16	4.12E-07	0.0313
cg25963123	11	117198671	CEP164	0.76	0.15	4.27E-07	0.0296
cg07695519	4	140374179	RAB33B	-0.37	0.08	7.39E-07	0.0469
cg27622506	3	24537160	THRB	0.85	0.17	7.41E-07	0.0434
cg25083270	16	69220946	SNTB2	0.57	0.11	7.48E-07	0.0407
cg18087672	17	46824915	NA	0.75	0.15	7.73E-07	0.0392
cg25834692	3	29535414	RBMS3	-0.50	0.10	8.49E-07	0.0404

Supplementary table 5: Functions of genes located at differentially methylated CpG sites (FDR < 5%) in the analyzed brain regions. Genes linked according to Illumina annotation.

Gene	Region	Function of gene product
Neurodevelopment and Neurotransmission		
MCTP2	PU	Presynaptic calcium sensor, leading to homeostatic stabilization of synaptic transmission in drosophila[3].
SNX26	PU	Brain-enriched RHO GTPase-activating protein essential for dendritic spine arborization during early neocortex development. In mature neurons a role in the activity-dependent structural change of dendritic spines is assumed[4].
WWOX	NC	In the brain WWOX is mainly expressed in the medial entorhinal cortex, frontal cortex layer 5 and GABAergic basket and granule cells in the cerebellum. Variants are associated with multiple neurodevelopmental and neurodegenerative disorders[5].
RYR3	PC	Ryanodine receptor RYR3 functions in intracellular calcium release, promoting. among others, the activation of RAC1, a molecular factor for axonal growth. Ryanodine receptors are required in early stages of neuronal polarization and axonal development[6].
BRD2	PC	Heterozygous loss of function of BRD2 has been linked to a deficit of GABAergic neurons in the basal ganglia of mice, increasing seizure susceptibility in myoclonic epilepsy[7].
ANK3	PC	ANK3 encodes multiple isoforms of ankyrin-G, which contribute to neuronal development. It mediates the localization of proteins to the axon's initial segment or to the dendritic shaft and spines. Ankyrin-G and its binding partners are associated with multiple neuropsychiatric diseases[8].
CDK13	PC	CDK13 and CDK12, which share 92% identity in the kinase domain, regulate axonal growth during neural development. Their depletion impairs the ability to extend axons over larger distances[9].
RAB33B	PC	Rab small GTPases are crucial for intracellular membrane trafficking. In the zebrafish model, RAB33A and RAB33B mediate axonal outgrowth and forebrain commissure formation[10].
RBMS3	PC	RBMS3 belongs to the c-Myc binding proteins which have been implicated in cell growth control. Myc genes are complexly involved in neurodegenerative cell death but also neuronal repair[11].
Metabolism		

IER2	NC	IER2 might act as a transcription factor for the human myo-inositol phosphate synthase[12]. Inositol-containing molecules act as membrane components and second messengers affecting proliferation, neurostimulation, secretion, and contraction.
SFXN1	NC	The mitochondrial serine transporter SFXN1 functions in one-carbon metabolism which is required for various anabolic processes. As the pathway has cytosolic and mitochondrial branches the transportation of serine into mitochondria is crucial[13].
SLC25A28	PC	SLC25A28 (Mitoferrin 2) is one of two homologous mitochondrial iron transporters that mediate iron ion uptake. Mitochondrial iron supply is essential for neuronal energy metabolism[14].
CRYL1	PC	CRYL1 functions in the uronate pathway which accounts for about 5% of daily glucose catabolism.
SNTB2	PC	Syntrophin Beta 2 is a dystrophin/utrophin-associated protein that appears to function in control of lipid levels[15, 16].
Others		
DIO2-AS1	PU	-
LOC222699	PC	-
CEP164	PC	Involvement in ciliogenesis, DNA repair, and chromosome segregation. Variants have been associated with Bardet–Biedl syndrome[17].
THRB	PC	Nuclear hormone receptor for triiodothyronine. It is one of the several receptors to mediate the biological activities of thyroid hormones.

Supplementary table 6: CpG sites included in the optimized methylation risk score with chromosomal location, effect size, significance level p, and false discovery rate q.

CpG	Chr	Pos	logFC	p	q
cg07786668	chr16	73092391	-0.69	1.4e-27	7.5e-22
cg04192862	chr10	28966472	-0.32	7e-25	1.3e-19
cg08668411	chr15	85525019	-0.67	6.6e-25	1.3e-19
cg17218495	chr19	11071743	-0.74	1.5e-24	1.7e-19
cg04794887	chr4	1858231	-0.32	1.6e-24	1.7e-19
cg18766912	chr15	25683909	-0.38	2.1e-24	1.8e-19
cg22512322	chr3	64009096	-0.51	2.3e-24	1.8e-19
cg07182872	chr3	136471214	-0.51	3.7e-24	2.5e-19
cg10849092	chr15	40574697	-0.34	6.8e-24	4.1e-19
cg16041611	chr6	43139680	-0.74	6.1e-23	3e-18
cg14119263	chr10	5708327	-0.51	1.3e-22	5.9e-18
cg08655589	chr3	14444175	-0.63	1.9e-22	7.9e-18
cg03183872	chr20	3140552	-0.61	3.7e-22	1.3e-17
cg21160472	chr1	212782112	-0.65	7.1e-22	2.4e-17
cg09238598	chr5	14871908	-0.38	1.1e-21	3.3e-17
cg08987887	chr6	76311629	-0.7	1.1e-21	3.4e-17
cg06819235	chr10	181229	-0.65	1.7e-21	4.7e-17
cg09643587	chr3	107809710	-0.38	4.7e-21	1.1e-16

cg22167789	chr19	1754780	-0.45	5.1e-21	1.1e-16
cg17626178	chr2	205410273	-0.41	5.6e-21	1.2e-16
cg14864852	chr22	39102110	-0.56	6.4e-21	1.3e-16
cg14810343	chr5	139028149	-0.68	8.9e-21	1.8e-16
cg06642177	chr6	134496341	-0.76	1e-20	2e-16
cg17662034	chr8	74207518	-0.88	2.5e-20	4.3e-16
cg26606224	chr9	123476699	-0.51	3.5e-20	5.8e-16
cg12100751	chr1	109203672	-0.61	4.1e-20	6.4e-16
cg02547025	chr2	30454275	-0.59	4.1e-20	6.4e-16
cg25869317	chr15	101792241	-0.35	5.3e-20	7.7e-16
cg01896926	chr17	685509	-0.83	8.1e-20	1.1e-15
cg25139493	chr1	39957400	-0.3	9.2e-20	1.2e-15

Extended online methods

Quality control, EWAS, and meta-analyses: Exclusion of potential confounders

Phillips et al[18] reported associations of RLS with alcohol consumption, BMI, diabetes, income, physical activity, and smoking. Although we focused on primary RLS and used controls matched for sex and age, we assessed these traits as possible confounders. To do so, we referred to DNA methylation risk scores because our case samples did not include detailed information on these traits. Searching the MRC-IEU catalog (http://ewascatalog.org) we extracted CpG sites and weights of the risk scores from respective studies on individuals of European descent[19, 20, 21, 22, 23, 24]. In case of meta-analyses that included KORA[21, 24], we used weights derived from other European subsamples in these studies. For the quantitative traits the risk scores for each individual were calculated as the weighted sums of the CpG site methylation levels (M values as in the respective publications), while in case of smoking we calculated the probability for being a current smoker using the definition given in the respective study[19]. We then validated the risk scores in the KORA control samples (that came with information on the traits in question) by Pearson or biserial correlation analyses which were highly significant for alcohol consumption, BMI, diabetes, and smoking, but non-significant for physical activity and income. The 4 risk scores which significantly correlated with the respective traits in the control samples were then compared between cases and controls by 2-sided t-test (or Wilcoxon test if bimodality of binary trait scores was expected), followed by random-effect meta-analysis between studies. All comparisons were non-significant. Therefore, in order not to overfit our EWAS model, we did not include these risk scores. (Of note, these findings cannot

refute the RLS-associations reported by Phillips et al[18] because our sample of primary RLS cases is not suited to uncover associations that relate to secondary RLS).

Epigenetic age

We applied Horvath's epigenetic clock[25], a multi-tissue estimator of DNA methylation age (DNAmAge), to blood and brain DNA samples from RLS cases and controls in order to examine whether biological age is advanced in RLS. Horvath's epigenetic clock had been constructed on Illumina Infinium HumanMethylation450 (450K array), based on 353 CpGs. 19 of these sites are missing from the EPIC array. However, McEwen et al.[26] have shown that age prediction is largely unaffected by the platform differences. We calculated the difference between DNAmAge and chronological age and conducted linear regression analysis of that difference on RLS status, correcting for chronological age, sex, and - in case of the brain samples - post-mortem delay until autopsy. Results of individual batches were combined by random effect meta-analysis. For 41 individuals, the difference between DNAmAge and chronological age could be assessed in all 4 brain regions; then, in order to take advantage of the multiple measurements in these individuals, the average difference between DNAmAge and chronological age in the 4 regions was regressed on the RLS status and meta-analyzed across batches as described above. In addition, we used Shireby's Cortical clock[27] to estimate the DNAmAge in the brain DNA samples.

Supplementary References

1.   Wang B, Shi W, Miao Z. Confidence analysis of standard deviational ellipse and its extension into higher dimensional euclidean space. PLoS ONE 2015;10(3):e0118537.
2.   Xu Z, Xie C, Taylor JA, Niu L. ipDMR: identification of differentially methylated regions with interval P-values. Bioinformatics 2021;37(5):711-713.
3.   Genç Ö, Dickman DK, Ma W, Tong A, Fetter RD, Davis GW. MCTP is an ER-resident calcium sensor that stabilizes synaptic transmission and homeostatic plasticity. Elife 2017;6.
4.   Kim Y, Ha CM, Chang S. SNX26, a GTPase-activating protein for Cdc42, interacts with PSD-95 protein and is involved in activity-dependent dendritic spine formation in mature neurons. J Biol Chem 2013;288(41):29453-29466.
5.   Aldaz CM, Hussain T. WWOX Loss of Function in Neurodevelopmental and Neurodegenerative Disorders. Int J Mol Sci 2020;21(23).
6.   Wilson C, Muñoz-Palma E, Henríquez DR, et al. A Feed-Forward Mechanism Involving the NOX Complex and RyR-Mediated Ca2+ Release During Axonal Specification. J Neurosci 2016;36(43):11107-11119.
7.   Velíšek L, Shang E, Velíšková J, et al. GABAergic neuron deficit as an idiopathic generalized epilepsy mechanism: the role of BRD2 haploinsufficiency in juvenile myoclonic epilepsy. PLoS ONE 2011;6(8):e23656.
8.   Yoon S, Piguel NH, Penzes P. Roles and mechanisms of ankyrin-G in neuropsychiatric disorders. Exp Mol Med 2022;54(7):867-877.
9.   Chen HR, Lin GT, Huang CK, Fann MJ. Cdk12 and Cdk13 regulate axonal elongation through a common signaling pathway that modulates Cdk5 expression. Exp Neurol 2014;261:10-21.
10.   Huang L, Urasaki A, Inagaki N. Rab33a and Rab33ba mediate the outgrowth of forebrain commissural axons in the zebrafish brain. Sci Rep 2019;9(1):1799.
11.   Marinkovic T, Marinkovic D. Obscure Involvement of MYC in Neurodegenerative Diseases and Neuronal Repair. Mol Neurobiol 2021;58(8):4169-4177.
12.   Takaya T, Kasatani K, Noguchi S, Nikawa J. Functional analyses of immediate early gene ETR101 expressed in yeast. Biosci Biotechnol Biochem 2009;73(7):1653-1660.
13.   Kory N, Wyant GA, Prakash G, et al. SFXN1 is a mitochondrial serine transporter required for one-carbon metabolism. Science 2018;362(6416).
14.   Baldauf L, Endres T, Scholz J, et al. Mitoferrin-1 is required for brain energy metabolism and hippocampus-dependent memory. Neurosci Lett 2019;713:134521.
15.   Buechler C, Boettcher A, Bared SM, Probst MC, Schmitz G. The carboxyterminus of the ATP-binding cassette transporter A1 interacts with a beta2-syntrophin/utrophin complex. Biochem Biophys Res Commun 2002;293(2):759-765.
16.   Krautbauer S, Neumeier M, Haberl EM, et al. The utrophin-beta 2 syntrophin complex regulates adipocyte lipid droplet size independent of adipogenesis. Mol Cell Biochem 2019;452(1-2):29-39.
17.   Shamseldin HE, Shaheen R, Ewida N, et al. The morbid genome of ciliopathies: an update. Genet Med 2022;24(4):966.
18.   Phillips B, Young T, Finn L, Asher K, Hening WA, Purvis C. Epidemiology of restless legs symptoms in adults. Arch Intern Med 2000;160(14):2137-2141.
19.   Bollepalli S, Korhonen T, Kaprio J, Anders S, Ollikainen M. EpiSmokEr: a robust classifier to determine smoking status from DNA methylation data. Epigenomics 2019;11(13):1469-1486.
20.   Florath I, Butterbach K, Heiss J, et al. Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults. Diabetologia 2016;59(1):130-138.
21.   Liu C, Marioni RE, Hedman Å K, et al. A DNA methylation biomarker of alcohol consumption. Mol Psychiatry 2018;23(2):422-433.
22.   EH VANR, Dugué PA, Jung CH, et al. Physical Activity, Television Viewing Time, and DNA Methylation in Peripheral Blood. Med Sci Sports Exerc 2019;51(3):490-498.

23. Song N, Sim JA, Dong Q, et al. Blood DNA methylation signatures are associated with social determinants of health among survivors of childhood cancer. Epigenetics 2022;17(11):1389-1403.

24. Wahl S, Drong A, Lehne B, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature 2017;541(7635):81-86.

25. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14(10):R115.

26. McEwen LM, Jones MJ, Lin DTS, et al. Systematic evaluation of DNA methylation age estimation with common preprocessing methods and the Infinium MethylationEPIC BeadChip array. Clin Epigenetics 2018;10(1):123.

27. Shireby GL, Davies JP, Francis PT, et al. Recalibrating the epigenetic clock: implications for assessing biological age in the human cortex. Brain 2020;143(12):3763-3775.

Full summary statistics are available online as Excel spreadsheet (Data S1):

https://movementdisorders.onlinelibrary.wiley.com/doi/10.1002/mds.29440

References

1. Palamara, Pier F., et al., *Leveraging Distant Relatedness to Quantify Human Mutation and Gene-Conversion Rates.* The American Journal of Human Genetics, 2015. **97**(6): p. 775-789.

2. Kong, A., et al., *Rate of de novo mutations and the importance of father's age to disease risk.* Nature, 2012. **488**(7412): p. 471-475.

3. Campbell, C.D., et al., *Estimating the human mutation rate using autozygosity in a founder population.* Nature Genetics, 2012. **44**(11): p. 1277-1281.

4. Francioli, L.C., et al., *Genome-wide patterns and properties of de novo mutations in humans.* Nature Genetics, 2015. **47**(7): p. 822-826.

5. Riordan, J.D. and J.H. Nadeau, *From Peas to Disease: Modifier Genes, Network Resilience, and the Genetics of Health.* The American Journal of Human Genetics, 2017. **101**(2): p. 177-191.

6. Schwartz, P.J., L. Crotti, and A.L. George, Jr., *Modifier genes for sudden cardiac death.* Eur Heart J, 2018. **39**(44): p. 3925-3931.

7. Agarwala, V., et al., *Evaluating empirical bounds on complex disease genetic architecture.* Nat Genet, 2013. **45**(12): p. 1418-27.

8. Marian, A.J., *Clinical Interpretation and Management of Genetic Variants.* JACC: Basic to Translational Science, 2020. **5**(10): p. 1029-1042.

9. Manolio, T.A., et al., *Finding the missing heritability of complex diseases.* Nature, 2009. **461**(7265): p. 747-53.

10. Lander, E.S., *The new genomics: global views of biology.* Science, 1996. **274**(5287): p. 536-9.

11. Reich, D.E. and E.S. Lander, *On the allelic spectrum of human disease.* Trends Genet, 2001. **17**(9): p. 502-10.

12. Pritchard, J.K. and N.J. Cox, *The allelic architecture of human disease genes: common disease-common variant...or not?* Hum Mol Genet, 2002. **11**(20): p. 2417-23.

13. Li, P., et al., *An overview of SNP interactions in genome-wide association studies.* Brief Funct Genomics, 2015. **14**(2): p. 143-55.

14. Buniello, A., et al., *The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019.* Nucleic Acids Res, 2019. **47**(D1): p. D1005-d1012.

15. Nica, A.C. and E.T. Dermitzakis, *Expression quantitative trait loci: present and future.* Philos Trans R Soc Lond B Biol Sci, 2013. **368**(1620): p. 20120362.

16. Gibson, G., *Rare and common variants: twenty arguments.* Nat Rev Genet, 2012. **13**(2): p. 135-45.

17. Chen, W., B.J. Coombes, and N.B. Larson, *Recent advances and challenges of rare variant association analysis in the biobank sequencing era.* Front Genet, 2022. **13**: p. 1014947.

18. Boycott, K.M., et al., *International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases.* The American Journal of Human Genetics, 2017. **100**(5): p. 695-705.

19. Wright, C.F., et al., *Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data.* Lancet, 2015. **385**(9975): p. 1305-14.

20. Gruber, C. and D. Bogunovic, *Incomplete penetrance in primary immunodeficiency: a skeleton in the closet.* Human Genetics, 2020. **139**(6): p. 745-757.

21. Oetjens, M.T., et al., *Quantifying the polygenic contribution to variable expressivity in eleven rare genetic disorders.* Nature Communications, 2019. **10**(1): p. 4897.

22. Bateman, M.S., et al., *Incomplete penetrance, variable expressivity, or dosage insensitivity in four families with directly transmitted unbalanced chromosome abnormalities.* Am J Med Genet A, 2018. **176**(2): p. 319-329.

23. Niemi, M.E.K., et al., *Common genetic variants contribute to risk of rare severe neurodevelopmental disorders.* Nature, 2018. **562**(7726): p. 268-271.

24. Wainschtein, P., et al., *Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data.* Nature Genetics, 2022. **54**(3): p. 263-273.

25. Momozawa, Y. and K. Mizukami, *Unique roles of rare variants in the genetics of complex diseases in humans.* Journal of Human Genetics, 2021. **66**(1): p. 11-23.

26. Li, N., et al., *Investigation of monogenic causes of familial breast cancer: data from the BEACCON case-control study.* NPJ Breast Cancer, 2021. **7**(1): p. 76.

27. Fahn, S., *Classification of movement disorders.* Mov Disord, 2011. **26**(6): p. 947-57.

28. Compston, A., *Progressive lenticular degeneration: a familial nervous disease associated with cirrhosis of the liver, by S. A. Kinnier Wilson, (From the National Hospital, and the Laboratory of the National Hospital, Queen Square, London) Brain 1912: 34; 295–509.* Brain, 2009. **132**(8): p. 1997-2001.

29. Baumann, J., *THE CLASSIFICATION OF THE DISEASES OF THE EXTRAPYRAMIDAL SYSTEM.* Acta Neurol Scand Suppl, 1963. **39**(4): p. Suppl4:102-7.

30. Klein, C., *Movement disorders: classifications.* J Inherit Metab Dis, 2005. **28**(3): p. 425-39.

31. Olgiati, S., M. Quadri, and V. Bonifati, *Genetics of movement disorders in the next-generation sequencing era.* Mov Disord, 2016. **31**(4): p. 458-70.

32. Di Fonzo, A., E. Monfrini, and R. Erro, *Genetics of Movement Disorders and the Practicing Clinician; Who and What to Test for?* Curr Neurol Neurosci Rep, 2018. **18**(7): p. 37.

33. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.

34. Committee on the Review of Omics-Based Tests for Predicting Patient Outcomes in Clinical, T., et al., in *Evolution of Translational Omics: Lessons Learned and the Path Forward*, C.M. Micheel, S.J. Nass, and G.S. Omenn, Editors. 2012, National Academies Press (US)

35. Petersen, B.S., et al., *Opportunities and challenges of whole-genome and -exome sequencing.* BMC Genet, 2017. **18**(1): p. 14.

36. van Dijk, E.L., et al., *Ten years of next-generation sequencing technology.* Trends Genet, 2014. **30**(9): p. 418-26.

37. Chandler, N., et al., *Rapid prenatal diagnosis using targeted exome sequencing: a cohort study to assess feasibility and potential impact on prenatal counseling and pregnancy management.* Genet Med, 2018. **20**(11): p. 1430-1437.

38. Valencia, C.A., et al., *Clinical Impact and Cost-Effectiveness of Whole Exome Sequencing as a Diagnostic Tool: A Pediatric Center's Experience.* Front Pediatr, 2015. **3**: p. 67.

39. Petrovski, S., et al., *Whole-exome sequencing in the evaluation of fetal structural anomalies: a prospective cohort study.* Lancet, 2019. **393**(10173): p. 758-767.

40. Meng, L., et al., *Use of Exome Sequencing for Infants in Intensive Care Units: Ascertainment of Severe Single-Gene Disorders and Effect on Medical Management.* JAMA Pediatr, 2017. **171**(12): p. e173438.

41. Posey, J.E., et al., *Resolution of Disease Phenotypes Resulting from Multilocus Genomic Variation.* N Engl J Med, 2017. **376**(1): p. 21-31.

42. Yang, Y., et al., *Molecular findings among patients referred for clinical whole-exome sequencing.* Jama, 2014. **312**(18): p. 1870-9.

43.   Zech, M., et al., *Monogenic variants in dystonia: an exome-wide sequencing study.* Lancet Neurol, 2020. **19**(11): p. 908-918.

44.   Hamosh, A., et al., *Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders.* Nucleic Acids Res, 2005. **33**(Database issue): p. D514-7.

45.   Harrer, P., et al., *Recessive NUP54 Variants Underlie Early-Onset Dystonia with Striatal Lesions.* Ann Neurol, 2023. **93**(2): p. 330-335.

46.   Karczewski, K.J., et al., *The mutational constraint spectrum quantified from variation in 141,456 humans.* Nature, 2020. **581**(7809): p. 434-443.

47.   Landrum, M.J., et al., *ClinVar: public archive of interpretations of clinically relevant variants.* Nucleic Acids Res, 2016. **44**(D1): p. D862-8.

48.   Stenson, P.D., et al., *The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies.* Hum Genet, 2017. **136**(6): p. 665-677.

49.   Richards, S., et al., *Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.* Genet Med, 2015. **17**(5): p. 405-24.

50.   Consortium, T.G., et al., *The GTEx Consortium atlas of genetic regulatory effects across human tissues.* Science, 2020. **369**(6509): p. 1318-1330.

51.   Uhlén, M., et al., *Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.

52.   Rentzsch, P., et al., *CADD: predicting the deleteriousness of variants throughout the human genome.* Nucleic Acids Research, 2018. **47**(D1): p. D886-D894.

53.   Sobreira, N., et al., *GeneMatcher: a matching tool for connecting investigators with an interest in the same gene.* Hum Mutat, 2015. **36**(10): p. 928-30.

54.   Wang, B., et al., *Reviving the Transcriptome Studies: An Insight Into the Emergence of Single-Molecule Transcriptome Sequencing.* Front Genet, 2019. **10**: p. 384.

55.   Zubair, M., et al., *Proteomics approaches: A review regarding an importance of proteome analyses in understanding the pathogens and diseases.* Front Vet Sci, 2022. **9**: p. 1079359.

56.   consortium, w., *Protein Data Bank: the single global archive for 3D macromolecular structure data.* Nucleic Acids Research, 2018. **47**(D1): p. D520-D528.

57.   Steinegger, M., M. Mirdita, and J. Söding, *Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold.* Nat Methods, 2019. **16**(7): p. 603-606.

58.   Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.

59.   Wilhelm, M., et al., *Mass-spectrometry-based draft of the human proteome.* Nature, 2014. **509**(7502): p. 582-7.

60.   Aslam, B., et al., *Proteomics: Technologies and Their Applications.* J Chromatogr Sci, 2017. **55**(2): p. 182-196.

61.   Pillai-Kastoori, L., A.R. Schutz-Geschwender, and J.A. Harford, *A systematic approach to quantitative Western blot analysis.* Analytical Biochemistry, 2020. **593**: p. 113608.

62.   Laurila, K. and M. Vihinen, *Prediction of disease-related mutations affecting protein localization.* BMC Genomics, 2009. **10**(1): p. 122.

63.   Park, S., et al., *Protein localization as a principal feature of the etiology and comorbidity of genetic diseases.* Mol Syst Biol, 2011. **7**: p. 494.

64.   Ferrari, S., F. Pellati, and M. Costi, *Disruption of Protein-Protein Interfaces.* 2013, Springer, Berlin.

65. Lu, H., et al., *Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials.* Signal Transduction and Targeted Therapy, 2020. **5**(1): p. 213.

66. Ni, D., P. Xu, and S. Gallagher, *Immunoblotting and Immunodetection.* Curr Protoc Mol Biol, 2016. **114**: p. 10.8.1-10.8.37.

67. Edidin, M., *Fluorescence resonance energy transfer: techniques for measuring molecular conformation and molecular proximity.* Curr Protoc Immunol, 2003. **Chapter 18**: p. Unit 18.10.

68. Alam, M.S., *Proximity Ligation Assay (PLA).* Curr Protoc Immunol, 2018. **123**(1): p. e58.

69. Peschansky, V.J. and C. Wahlestedt, *Non-coding RNAs as direct and indirect modulators of epigenetic regulation.* Epigenetics, 2014. **9**(1): p. 3-12.

70. van Otterdijk, S.D. and K.B. Michels, *Transgenerational epigenetic inheritance in mammals: how good is the evidence?* Faseb j, 2016. **30**(7): p. 2457-65.

71. Eckersley-Maslin, M.A., C. Alda-Catalinas, and W. Reik, *Dynamics of the epigenetic landscape during the maternal-to-zygotic transition.* Nat Rev Mol Cell Biol, 2018. **19**(7): p. 436-450.

72. Kobayashi, T. and M.A. Surani, *On the origin of the human germline.* Development, 2018. **145**(16).

73. Janssen, S.M. and M.C. Lorincz, *Interplay between chromatin marks in development and disease.* Nature Reviews Genetics, 2022. **23**(3): p. 137-153.

74. Carter, B. and K. Zhao, *The epigenetic basis of cellular heterogeneity.* Nature Reviews Genetics, 2021. **22**(4): p. 235-250.

75. Kim, M. and J. Costello, *DNA methylation: an epigenetic mark of cellular memory.* Experimental & Molecular Medicine, 2017. **49**(4): p. e322-e322.

76. Yousefi, P.D., et al., *DNA methylation-based predictors of health: applications and statistical considerations.* Nature Reviews Genetics, 2022. **23**(6): p. 369-383.

77. Sánchez-Pernaute, O. and A. Gabucio, *Epigenetics and complex diseases, a new era in the assessment of exposure and risk?* Rheumatology, 2017. **57**(10): p. 1693-1694.

78. Nabais, M.F., et al., *An overview of DNA methylation-derived trait score methods and applications.* Genome Biology, 2023. **24**(1): p. 28.

79. Seiler Vellame, D., et al., *Characterizing the properties of bisulfite sequencing data: maximizing power and sensitivity to identify between-group differences in DNA methylation.* BMC Genomics, 2021. **22**(1): p. 446.

80. Houseman, E.A., et al., *DNA methylation arrays as surrogate measures of cell mixture distribution.* BMC Bioinformatics, 2012. **13**: p. 86.

81. Caggiano, C., et al., *Comprehensive cell type decomposition of circulating cell-free DNA with CelFiE.* Nat Commun, 2021. **12**(1): p. 2717.

82. Hannon, E., et al., *Leveraging DNA-Methylation Quantitative-Trait Loci to Characterize the Relationship between Methylomic Variation, Gene Expression, and Complex Traits.* Am J Hum Genet, 2018. **103**(5): p. 654-665.

83. Min, J.L., et al., *Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation.* Nat Genet, 2021. **53**(9): p. 1311-1321.

84. Luo, C., P. Hajkova, and J.R. Ecker, *Dynamic DNA methylation: In the right place at the right time.* Science, 2018. **361**(6409): p. 1336-1340.

85. Zeilinger, S., et al., *Tobacco smoking leads to extensive genome-wide changes in DNA methylation.* PLoS One, 2013. **8**(5): p. e63812.

86. Horvath, S., *DNA methylation age of human tissues and cell types.* Genome Biol, 2013. **14**(10): p. R115.

87. Kabacik, S., et al., *The relationship between epigenetic age and the hallmarks of aging in human cells.* Nature Aging, 2022. **2**(6): p. 484-493.

88. Horvath, S. and K. Raj, *DNA methylation-based biomarkers and the epigenetic clock theory of ageing.* Nature Reviews Genetics, 2018. **19**(6): p. 371-384.

89. Shireby, G.L., et al., *Recalibrating the epigenetic clock: implications for assessing biological age in the human cortex.* Brain, 2020. **143**(12): p. 3763-3775.

90. Maleki, F., et al., *Gene Set Analysis: Challenges, Opportunities, and Future Research.* Front Genet, 2020. **11**: p. 654.

91. Kanehisa, M., et al., *KEGG as a reference resource for gene and protein annotation.* Nucleic Acids Res, 2016. **44**(D1): p. D457-62.

92. Watanabe, K., et al., *Functional mapping and annotation of genetic associations with FUMA.* Nature Communications, 2017. **8**(1): p. 1826.

93. Docherty, L.E., et al., *Genome-wide DNA methylation analysis of patients with imprinting disorders identifies differentially methylated regions associated with novel candidate imprinted genes.* J Med Genet, 2014. **51**(4): p. 229-38.

94. Ziller, M.J., et al., *Charting a dynamic DNA methylation landscape of the human genome.* Nature, 2013. **500**(7463): p. 477-481.

95. Albanese, A., et al., *Phenomenology and classification of dystonia: a consensus update.* Mov Disord, 2013. **28**(7): p. 863-73.

96. Balint, B., et al., *Dystonia.* Nat Rev Dis Primers, 2018. **4**(1): p. 25.

97. Jankovic, J., *Treatment of hyperkinetic movement disorders.* Lancet Neurol, 2009. **8**(9): p. 844-56.

98. Steeves, T.D., et al., *The prevalence of primary dystonia: a systematic review and meta-analysis.* Mov Disord, 2012. **27**(14): p. 1789-96.

99. Williams, L., et al., *Epidemiological, clinical and genetic aspects of adult onset isolated focal dystonia in Ireland.* Eur J Neurol, 2017. **24**(1): p. 73-81.

100. Asgeirsson, H., et al., *Prevalence study of primary dystonia in Iceland.* Mov Disord, 2006. **21**(3): p. 293-8.

101. Sitburana, O., et al., *Motor overflow and mirror dystonia.* Parkinsonism Relat Disord, 2009. **15**(10): p. 758-61.

102. Stamelou, M., et al., *The non-motor syndrome of primary dystonia: clinical and pathophysiological implications.* Brain, 2012. **135**(Pt 6): p. 1668-81.

103. Fasano, A., et al., *Non-DYT1 early-onset primary torsion dystonia: comparison with DYT1 phenotype and review of the literature.* Mov Disord, 2006. **21**(9): p. 1411-8.

104. Keller Sarmiento, I.J. and N.E. Mencacci, *Genetic Dystonias: Update on Classification and New Genetic Discoveries.* Curr Neurol Neurosci Rep, 2021. **21**(3): p. 8.

105. Lin, J.P. and N. Nardocci, *Recognizing the Common Origins of Dystonia and the Development of Human Movement: A Manifesto of Unmet Needs in Isolated Childhood Dystonias.* Front Neurol, 2016. **7**: p. 226.

106. Bhatia, K.P. and C.D. Marsden, *The behavioural and motor consequences of focal lesions of the basal ganglia in man.* Brain, 1994. **117 (Pt 4)**: p. 859-76.

107. Jinnah, H.A., V. Neychev, and E.J. Hess, *The Anatomical Basis for Dystonia: The Motor Network Model.* Tremor Other Hyperkinet Mov (N Y), 2017. **7**: p. 506.

108. Pantano, P., et al., *A transverse and longitudinal MR imaging voxel-based morphometry study in patients with primary cervical dystonia.* AJNR Am J Neuroradiol, 2011. **32**(1): p. 81-4.

109. Poston, K.L. and D. Eidelberg, *Functional brain networks and abnormal connectivity in the movement disorders.* Neuroimage, 2012. **62**(4): p. 2261-70.

110. Berardelli, A., et al., *The pathophysiology of primary dystonia.* Brain, 1998. **121 (Pt 7)**: p. 1195-212.

111. Calabresi, P., et al., *Hyperkinetic disorders and loss of synaptic downscaling.* Nat Neurosci, 2016. **19**(7): p. 868-75.

112. Martella, G., et al., *Regional specificity of synaptic plasticity deficits in a knock-in mouse model of DYT1 dystonia.* Neurobiol Dis, 2014. **65**: p. 124-32.

113. Gonzalez-Latapi, P., N. Marotta, and N.E. Mencacci, *Emerging and converging molecular mechanisms in dystonia.* Journal of Neural Transmission, 2021. **128**(4): p. 483-498.

114. Jinnah, H.A. and Y.V. Sun, *Dystonia genes and their biological pathways.* Neurobiol Dis, 2019. **129**: p. 159-168.

115. Franco, R., I. Reyes-Resina, and G. Navarro, *Dopamine in Health and Disease: Much More Than a Neurotransmitter.* Biomedicines, 2021. **9**(2).

116. Ichinose, H., et al., *Hereditary progressive dystonia with marked diurnal fluctuation caused by mutations in the GTP cyclohydrolase I gene.* Nat Genet, 1994. **8**(3): p. 236-42.

117. Ribot, B., et al., *Dystonia and dopamine: From phenomenology to pathophysiology.* Prog Neurobiol, 2019. **182**: p. 101678.

118. Ijomone, O.M., et al., *The aging brain: impact of heavy metal neurotoxicity.* Crit Rev Toxicol, 2020. **50**(9): p. 801-814.

119. Chen, P., M.R. Miah, and M. Aschner, *Metals and Neurodegeneration.* F1000Res, 2016. **5**.

120. Ticci, C., et al., *Movement Disorders in Children with a Mitochondrial Disease: A Cross-Sectional Survey from the Nationwide Italian Collaborative Network of Mitochondrial Diseases.* J Clin Med, 2021. **10**(10).

121. Distelmaier, F., et al., *Treatable mitochondrial diseases: cofactor metabolism and beyond.* Brain, 2017. **140**(2): p. e11.

122. Stenton, S.L., et al., *Leigh Syndrome: A Study of 209 Patients at the Beijing Children's Hospital.* Ann Neurol, 2022. **91**(4): p. 466-482.

123. Stenton, S.L. and H. Prokisch, *Genetics of mitochondrial diseases: Identifying mutations to help diagnosis.* EBioMedicine, 2020. **56**: p. 102784.

124. Brini, M., et al., *Calcium in health and disease.* Met Ions Life Sci, 2013. **13**: p. 81-137.

125. Brini, M., et al., *Neuronal calcium signaling: function and dysfunction.* Cell Mol Life Sci, 2014. **71**(15): p. 2787-814.

126. Labrum, R.W., et al., *Large scale calcium channel gene rearrangements in episodic ataxia and hemiplegic migraine: implications for diagnostic testing.* J Med Genet, 2009. **46**(11): p. 786-91.

127. Atasu, B., et al., *HPCA confirmed as a genetic cause of DYT2-like dystonia phenotype.* Mov Disord, 2018. **33**(8): p. 1354-1358.

128. Feyma, T., et al., *Dystonia in ATP2B3-associated X-linked spinocerebellar ataxia.* Mov Disord, 2016. **31**(11): p. 1752-1753.

129. Zech, M., et al., *A Neurodevelopmental Disorder With Dystonia and Chorea Resulting From Clustering CAMK4 Variants.* Mov Disord, 2021. **36**(2): p. 520-521.

130. Dzinovic, I., et al., *Myoclonic dystonia phenotype related to a novel calmodulin-binding transcription activator 1 sequence variant.* Neurogenetics, 2021. **22**(2): p. 137-141.

131. Dzinovic, I., J. Winkelmann, and M. Zech, *Genetic intersection between dystonia and neurodevelopmental disorders: Insights from genomic sequencing.* Parkinsonism & Related Disorders, 2022. **102**: p. 131-140.

132. Erro, R., N.E. Mencacci, and K.P. Bhatia, *The Emerging Role of Phosphodiesterases in Movement Disorders.* Mov Disord, 2021. **36**(10): p. 2225-2243.

133.   Zech, M., et al., *Haploinsufficiency of KMT2B, Encoding the Lysine-Specific Histone Methyltransferase 2B, Results in Early-Onset Generalized Dystonia.* Am J Hum Genet, 2016. **99**(6): p. 1377-1387.

134.   Fuchs, T., et al., *Mutations in the THAP1 gene are responsible for DYT6 primary torsion dystonia.* Nat Genet, 2009. **41**(3): p. 286-8.

135.   Barbagiovanni, G., et al., *KMT2B Is Selectively Required for Neuronal Transdifferentiation, and Its Loss Exposes Dystonia Candidate Genes.* Cell Rep, 2018. **25**(4): p. 988-1001.

136.   Faundes, V., et al., *Histone Lysine Methylases and Demethylases in the Landscape of Human Developmental Disorders.* Am J Hum Genet, 2018. **102**(1): p. 175-187.

137.   Mirza-Schreiber, N., et al., *Blood DNA methylation provides an accurate biomarker of KMT2B-related dystonia and predicts onset.* Brain, 2022. **145**(2): p. 644-654.

138.   Cheng, F., et al., *Unraveling Molecular Mechanisms of THAP1 Missense Mutations in DYT6 Dystonia.* J Mol Neurosci, 2020. **70**(7): p. 999-1008.

139.   Goodchild, R.E. and W.T. Dauer, *Mislocalization to the nuclear envelope: an effect of the dystonia-causing torsinA mutation.* Proc Natl Acad Sci U S A, 2004. **101**(3): p. 847-52.

140.   Chen, P., et al., *The early-onset torsion dystonia-associated protein, torsinA, is a homeostatic regulator of endoplasmic reticulum stress response.* Hum Mol Genet, 2010. **19**(18): p. 3502-15.

141.   Hewett, J.W., et al., *Mutant torsinA interferes with protein processing through the secretory pathway in DYT1 dystonia cells.* Proc Natl Acad Sci U S A, 2007. **104**(17): p. 7271-6.

142.   Rittiner, J.E., et al., *Functional Genomic Analyses of Mendelian and Sporadic Disease Identify Impaired eIF2α Signaling as a Generalizable Mechanism for Dystonia.* Neuron, 2016. **92**(6): p. 1238-1251.

143.   Willis, T. and E. Philiatros, *The London practice of physick.* (No Title), 1977.

144.   Ekbom, K.-A., *Restless Legs: A Clinical Study of a Hitherto Overlooked Disease in the Legs.* 1945.

145.   Para, K.S., et al., *Suicidal thought and behavior in individuals with restless legs syndrome.* Sleep Med, 2019. **54**: p. 1-7.

146.   Trenkwalder, C., et al., *Socioeconomic impact of restless legs syndrome and inadequate restless legs syndrome management across European settings.* Eur J Neurol, 2021. **28**(2): p. 691-706.

147.   Picchietti, D.L., et al., *Achievements, challenges, and future perspectives of epidemiologic research in restless legs syndrome (RLS).* Sleep Med, 2017. **31**: p. 3-9.

148.   Fulda, S., et al., *We need to do better: A systematic review and meta-analysis of diagnostic test accuracy of restless legs syndrome screening instruments.* Sleep Med Rev, 2021. **58**: p. 101461.

149.   Ma, J.F., et al., *Restless legs syndrome in Chinese elderly people of an urban suburb in Shanghai: a community-based survey.* Parkinsonism Relat Disord, 2012. **18**(3): p. 294-8.

150.   Kim, T.J., et al., *Prevalence and Characteristics of Restless Legs Syndrome in Korean Adults: A Study in Two Independent Samples of the General Population.* Neuroepidemiology, 2019. **52**(3-4): p. 193-204.

151.   Manconi, M., et al., *When gender matters: restless legs syndrome. Report of the "RLS and woman" workshop endorsed by the European RLS Study Group.* Sleep Med Rev, 2012. **16**(4): p. 297-307.

152.   Prosperetti, C. and M. Manconi, *Restless Legs Syndrome/Willis-Ekbom Disease and Pregnancy.* Sleep Med Clin, 2015. **10**(3): p. 323-9, xiv.

153. Trenkwalder, C., et al., *Restless legs syndrome associated with major diseases: A systematic review and new concept.* Neurology, 2016. **86**(14): p. 1336-1343.

154. Szentkirályi, A., et al., *Multimorbidity and the risk of restless legs syndrome in 2 prospective cohort studies.* Neurology, 2014. **82**(22): p. 2026-33.

155. Whittom, S., et al., *Age-at-onset in restless legs syndrome: a clinical and polysomnographic study.* Sleep Med, 2007. **9**(1): p. 54-9.

156. Allen, R.P. and C.J. Earley, *Defining the phenotype of the restless legs syndrome (RLS) using age-of-symptom-onset.* Sleep Med, 2000. **1**(1): p. 11-19.

157. Manconi, M., et al., *Restless legs syndrome.* Nature Reviews Disease Primers, 2021. **7**(1): p. 80.

158. Montplaisir, J., et al., *Clinical, polysomnographic, and genetic characteristics of restless legs syndrome: a study of 133 patients diagnosed with new standard criteria.* Mov Disord, 1997. **12**(1): p. 61-5.

159. Gamaldo, C., et al., *Evaluating daytime alertness in individuals with Restless Legs Syndrome (RLS) compared to sleep restricted controls.* Sleep Med, 2009. **10**(1): p. 134-8.

160. Walters, A.S., et al., *Validation of the International Restless Legs Syndrome Study Group rating scale for restless legs syndrome.* Sleep Med, 2003. **4**(2): p. 121-32.

161. Ohayon, M.M., R. O'Hara, and M.V. Vitiello, *Epidemiology of restless legs syndrome: a synthesis of the literature.* Sleep Med Rev, 2012. **16**(4): p. 283-95.

162. Allen, R.P., et al., *Family history study of the restless legs syndrome.* Sleep Med, 2002. **3 Suppl**: p. S3-7.

163. Ondo, W.G., K.D. Vuong, and Q. Wang, *Restless legs syndrome in monozygotic twins: clinical correlates.* Neurology, 2000. **55**(9): p. 1404-6.

164. Schormair, B., et al., *Identification of novel risk loci for restless legs syndrome in genome-wide association studies in individuals of European ancestry: a meta-analysis.* Lancet Neurol, 2017. **16**(11): p. 898-907.

165. Akçimen, F., et al., *Transcriptome-wide association study for restless legs syndrome identifies new susceptibility genes.* Commun Biol, 2020. **3**(1): p. 373.

166. Tilch, E., et al., *Identification of Restless Legs Syndrome Genes by Mutational Load Analysis.* Ann Neurol, 2020. **87**(2): p. 184-193.

167. Roy, A., et al., *Developing a biomarker for restless leg syndrome using genome wide DNA methylation data.* Sleep Medicine, 2021. **78**: p. 120-127.

168. de Paiva, J.P.Q., et al., *Sensorimotor white matter projections and disease severity in primary Restless Legs Syndrome/Willis-Ekbom disease: a multimodal DTI analysis.* Sleep Med, 2020. **73**: p. 106-116.

169. Kocar, T.D., H.P. Müller, and J. Kassubek, *Differential functional connectivity in thalamic and dopaminergic pathways in restless legs syndrome: a meta-analysis.* Ther Adv Neurol Disord, 2020. **13**: p. 1756286420941670.

170. Allen, R.P. and C.J. Earley, *The role of iron in restless legs syndrome.* Mov Disord, 2007. **22 Suppl 18**: p. S440-8.

171. Godau, J., et al., *Substantia nigra hypoechogenicity: definition and findings in restless legs syndrome.* Mov Disord, 2007. **22**(2): p. 187-92.

172. Schmidauer, C., et al., *Transcranial ultrasound shows nigral hypoechogenicity in restless legs syndrome.* Ann Neurol, 2005. **58**(4): p. 630-4.

173. Earley, C.J., et al., *Abnormalities in CSF concentrations of ferritin and transferrin in restless legs syndrome.* Neurology, 2000. **54**(8): p. 1698-700.

174. Connor, J.R., et al., *Profile of altered brain iron acquisition in restless legs syndrome.* Brain, 2011. **134**(Pt 4): p. 959-68.

175. Allen, R., *Dopamine and iron in the pathophysiology of restless legs syndrome (RLS).* Sleep Med, 2004. **5**(4): p. 385-91.

176. Connor, J.R., et al., *Altered dopaminergic profile in the putamen and substantia nigra in restless leg syndrome.* Brain, 2009. **132**(Pt 9): p. 2403-12.

177. Earley, C.J., et al., *Connectome and molecular pharmacological differences in the dopaminergic system in restless legs syndrome (RLS): plastic changes and neuroadaptations that may contribute to augmentation.* Sleep Med, 2017. **31**: p. 71-77.

178. Manconi, M., et al., *Restless legs syndrome.* Nat Rev Dis Primers, 2021. **7**(1): p. 80.

179. Allen, R.P., *Restless Leg Syndrome/Willis-Ekbom Disease Pathophysiology.* Sleep Med Clin, 2015. **10**(3): p. 207-14, xi.

180. Khan, F.H., et al., *Iron, dopamine, genetics, and hormones in the pathophysiology of restless legs syndrome.* J Neurol, 2017. **264**(8): p. 1634-1641.

181. Lanza, G. and R. Ferri, *The neurophysiology of hyperarousal in restless legs syndrome: Hints for a role of glutamate/GABA.* Adv Pharmacol, 2019. **84**: p. 101-119.

182. Ferré, S., et al., *New Insights into the Neurobiology of Restless Legs Syndrome.* Neuroscientist, 2019. **25**(2): p. 113-125.

183. Kapur, N. and R. Friedman, *Oral ketamine: a promising treatment for restless legs syndrome.* Anesth Analg, 2002. **94**(6): p. 1558-9, table of contents.

184. Silver, N., et al., *A 10-year, longitudinal assessment of dopamine agonists and methadone in the treatment of restless legs syndrome.* Sleep Med, 2011. **12**(5): p. 440-4.

185. Winkelmann, J., et al., *Treatment of restless legs syndrome: Evidence-based review and implications for clinical practice (Revised 2017)(§).* Mov Disord, 2018. **33**(7): p. 1077-1091.

186. Garcia-Borreguero, D., et al., *Pregabalin versus pramipexole: effects on sleep disturbance in restless legs syndrome.* Sleep, 2014. **37**(4): p. 635-43.

187. McGahan, M.C., et al., *Iron alters glutamate secretion by regulating cytosolic aconitase activity.* Am J Physiol Cell Physiol, 2005. **288**(5): p. C1117-24.

188. Yepes, G., et al., *Targeting hypersensitive corticostriatal terminals in restless legs syndrome.* Ann Neurol, 2017. **82**(6): p. 951-960.

189. Ferré, S., et al., *Pivotal Role of Adenosine Neurotransmission in Restless Legs Syndrome.* Front Neurosci, 2017. **11**: p. 722.

190. Quiroz, C., et al., *Adenosine receptors as markers of brain iron deficiency: Implications for Restless Legs Syndrome.* Neuropharmacology, 2016. **111**: p. 160-168.

191. Ferré, S., et al., *Essential Control of the Function of the Striatopallidal Neuron by Pre-coupled Complexes of Adenosine A(2A)-Dopamine D(2) Receptor Heterotetramers and Adenylyl Cyclase.* Front Pharmacol, 2018. **9**: p. 243.

192. Garcia-Borreguero, D., et al., *A Randomized, Placebo-Controlled Crossover Study with Dipyridamole for Restless Legs Syndrome.* Mov Disord, 2021. **36**(10): p. 2387-2392.

193. Harrer, P., et al., *Dystonia Linked to EIF4A2 Haploinsufficiency: A Disorder of Protein Translation Dysfunction.* Movement Disorders, 2023. **n/a**(n/a).

194. Harrer, P., et al., *Epigenetic Association Analyses and Risk Prediction of RLS.* Mov Disord, 2023.

195. Zuk, O., et al., *Searching for missing heritability: designing rare variant association studies.* Proc Natl Acad Sci U S A, 2014. **111**(4): p. E455-64.

196. Rodenburg, R.J., *The functional genomics laboratory: functional validation of genetic variants.* J Inherit Metab Dis, 2018. **41**(3): p. 297-307.

197. Strambio-De-Castillia, C., M. Niepel, and M.P. Rout, *The nuclear pore complex: bridging nuclear transport and gene regulation.* Nat Rev Mol Cell Biol, 2010. **11**(7): p. 490-501.

198. Nofrini, V., D. Di Giacomo, and C. Mecucci, *Nucleoporin genes in human diseases.* European Journal of Human Genetics, 2016. **24**(10): p. 1388-1395.

199. Guglielmi, V., S. Sakuma, and M.A. D'Angelo, *Nuclear pore complexes in development and tissue homeostasis.* Development, 2020. **147**(23).

200. Basel-Vanagaite, L., et al., *Mutated nup62 causes autosomal recessive infantile bilateral striatal necrosis.* Ann Neurol, 2006. **60**(2): p. 214-22.

201. Worman, H.J. and W.T. Dauer, *The Nuclear Envelope: An Intriguing Focal Point for Neurogenetic Disease.* Neurotherapeutics, 2014. **11**(4): p. 764-772.

202. Di Fonzo, A., A. Albanese, and H.A. Jinnah, *The apparent paradox of phenotypic diversity and shared mechanisms across dystonia syndromes.* Curr Opin Neurol, 2022. **35**(4): p. 502-509.

203. Mannino, P.J. and C.P. Lusk, *Quality control mechanisms that protect nuclear envelope identity and function.* J Cell Biol, 2022. **221**(9).

204. Jokhi, V., et al., *Torsin mediates primary envelopment of large ribonucleoprotein granules at the nuclear envelope.* Cell Rep, 2013. **3**(4): p. 988-95.

205. VanGompel, M.J., et al., *A novel function for the Caenorhabditis elegans torsin OOC-5 in nucleoporin localization and nuclear import.* Mol Biol Cell, 2015. **26**(9): p. 1752-63.

206. Laudermilch, E., et al., *Dissecting Torsin/cofactor function at the nuclear envelope: a genetic study.* Mol Biol Cell, 2016. **27**(25): p. 3964-3971.

207. Doye, V. and E. Hurt, *From nucleoporins to nuclear pore complexes.* Curr Opin Cell Biol, 1997. **9**(3): p. 401-11.

208. Schwartz, T.U., *The Structure Inventory of the Nuclear Pore Complex.* J Mol Biol, 2016. **428**(10 Pt A): p. 1986-2000.

209. Milles, S., et al., *Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors.* Cell, 2015. **163**(3): p. 734-45.

210. Hershey, J.W.B., N. Sonenberg, and M.B. Mathews, *Principles of Translational Control.* Cold Spring Harb Perspect Biol, 2019. **11**(9).

211. Kuipers, D.J.S., et al., *EIF2AK2 Missense Variants Associated with Early Onset Generalized Dystonia.* Ann Neurol, 2021. **89**(3): p. 485-497.

212. Musacchio, T., et al., *A Recurrent EIF2AK2 Missense Variant Causes Autosomal-Dominant Isolated Dystonia.* Ann Neurol, 2021. **89**(6): p. 1257-1258.

213. Camargos, S., et al., *DYT16, a novel young-onset dystonia-parkinsonism disorder: identification of a segregating mutation in the stress-response protein PRKRA.* Lancet Neurol, 2008. **7**(3): p. 207-15.

214. Ge, L., et al., *EIF2AK2 selectively regulates the gene transcription in immune response and histones associated with systemic lupus erythematosus.* Molecular Immunology, 2021. **132**: p. 132-141.

215. Collart, M.A., *The Ccr4-Not complex is a key regulator of eukaryotic gene expression.* Wiley Interdiscip Rev RNA, 2016. **7**(4): p. 438-54.

216. Clark, M.M., et al., *Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases.* NPJ Genom Med, 2018. **3**: p. 16.

217. Smedley, D., et al., *100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report.* N Engl J Med, 2021. **385**(20): p. 1868-1880.

218. Lunke, S., et al., *Integrated multi-omics for rapid rare disease diagnosis on a national scale.* Nature Medicine, 2023. **29**(7): p. 1681-1691.

219. Hon, T., et al., *Highly accurate long-read HiFi sequencing data for five complex genomes.* Scientific Data, 2020. **7**(1): p. 399.

220. Mahmoud, M., et al., *Structural variant calling: the long and the short of it.* Genome Biol, 2019. **20**(1): p. 246.

221. Kingsmore, S.F., et al., *A Randomized, Controlled Trial of the Analytic and Diagnostic Performance of Singleton and Trio, Rapid Genome and Exome Sequencing in Ill Infants.* Am J Hum Genet, 2019. **105**(4): p. 719-733.

222. Brockman, D.G., et al., *Randomized prospective evaluation of genome sequencing versus standard-of-care as a first molecular diagnostic test.* Genet Med, 2021. **23**(9): p. 1689-1696.

223. Schaid, D.J., W. Chen, and N.B. Larson, *From genome-wide associations to candidate causal variants by statistical fine-mapping.* Nat Rev Genet, 2018. **19**(8): p. 491-504.

224. Cano-Gamez, E. and G. Trynka, *From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases.* Front Genet, 2020. **11**: p. 424.

225. McLaren, W., et al., *The Ensembl Variant Effect Predictor.* Genome Biol, 2016. **17**(1): p. 122.

226. Yang, H. and K. Wang, *Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR.* Nat Protoc, 2015. **10**(10): p. 1556-66.

227. Rizzo, G., et al., *Brain imaging and networks in restless legs syndrome.* Sleep Med, 2017. **31**: p. 39-48.

228. Stefani, A., et al., *Multimodal Magnetic Resonance Imaging reveals alterations of sensorimotor circuits in restless legs syndrome.* Sleep, 2019. **42**(12).

229. Iranzo, A., et al., *Restless legs syndrome in Parkinson's disease and other neurodegenerative diseases of the central nervous system.* Mov Disord, 2007. **22 Suppl 18**: p. S424-30.

230. Gan-Or, Z., et al., *Sleep disorders and Parkinson disease; lessons from genetics.* Sleep Med Rev, 2018. **41**: p. 101-112.

231. Postuma, R.B., et al., *Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study.* Brain, 2019. **142**(3): p. 744-759.

232. Schulte, D. and D. Geerts, *MEIS transcription factors in development and disease.* Development, 2019. **146**(16).

233. Hannon, E., et al., *Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes.* Epigenetics, 2015. **10**(11): p. 1024-32.

234. Gunasekara, C.J., et al., *A machine learning case-control classifier for schizophrenia based on DNA methylation in blood.* Transl Psychiatry, 2021. **11**(1): p. 412.

235. Chen, J., et al., *Association of a Reproducible Epigenetic Risk Profile for Schizophrenia With Brain Methylation and Function.* JAMA Psychiatry, 2020. **77**(6): p. 628-636.

236. Xie, Y., et al., *Integrated Analysis of Methylomic and Transcriptomic Data to Identify Potential Diagnostic Biomarkers for Major Depressive Disorder.* Genes (Basel), 2021. **12**(2).

237. Levy, M.A., et al., *Functional correlation of genome-wide DNA methylation profiles in genetic neurodevelopmental disorders.* Hum Mutat, 2022. **43**(11): p. 1609-1628.

238. Hüls, A. and D. Czamara, *Methodological challenges in constructing DNA methylation risk scores.* Epigenetics, 2020. **15**(1-2): p. 1-11.

239. Claussnitzer, M., et al., *A brief history of human disease genetics.* Nature, 2020. **577**(7789): p. 179-189.

240. Lappalainen, T. and D.G. MacArthur, *From variant to function in human disease genetics.* Science, 2021. **373**(6562): p. 1464-1468.

241. Smirnov, D., N. Konstantinovskiy, and H. Prokisch, *Integrative omics approaches to advance rare disease diagnostics.* J Inherit Metab Dis, 2023.

242.   Wortmann, S.B., et al., *How to proceed after "negative" exome: A review on genetic diagnostics, limitations, challenges, and emerging new multiomics techniques.* J Inherit Metab Dis, 2022. **45**(4): p. 663-681.

243.   Falkenberg, K.D., et al., *Allelic Expression Imbalance Promoting a Mutant PEX6 Allele Causes Zellweger Spectrum Disorder.* Am J Hum Genet, 2017. **101**(6): p. 965-976.

244.   Alaimo, J.T., et al., *Integrated analysis of metabolomic profiling and exome data supplements sequence variant interpretation, classification, and diagnosis.* Genet Med, 2020. **22**(9): p. 1560-1566.

245.   Fowler, D.M., et al., *An Atlas of Variant Effects to understand the genome at nucleotide resolution.* Genome Biology, 2023. **24**(1): p. 147.

246.   Morris, J.A., et al., *Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens.* Science, 2023. **380**(6646): p. eadh7699.

247.   Gloyn, A.L. and D.J. Drucker, *Precision medicine in the management of type 2 diabetes.* Lancet Diabetes Endocrinol, 2018. **6**(11): p. 891-900.

248.   Ahlqvist, E., et al., *Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables.* Lancet Diabetes Endocrinol, 2018. **6**(5): p. 361-369.

249.   Schrader, S., et al., *Novel Subgroups of Type 2 Diabetes Display Different Epigenetic Patterns That Associate With Future Diabetic Complications.* Diabetes Care, 2022. **45**(7): p. 1621-1630.

Figures and Tables

My parents, mentors, and colleagues.

Full list of Publications

Harrer P*, Škorvánek M*, Kittke V*, Dzinovic I, Borngräber F, Thomsen M, Mandel V, Svorenova T, Ostrozovicova M, Kulcsarova K, Berutti R, Busch H, Ott F, Kopajtich R, Prokisch H, Kumar KR, Mencacci NE, Kurian MA, Di Fonzo A, Boesch S, Kühn AA, Blümlein U, Lohmann K, Haslinger B, Weise D, Jech R, Winkelmann J†, Zech M†. Dystonia Linked to EIF4A2 Haploinsufficiency: A Disorder of Protein Translation Dysfunction. Mov Disord. 2023 Jul 23. doi: 10.1002/mds.29562. Epub ahead of print. PMID: 37485550.

Harrer P*, Mirza-Schreiber N*, Mandel V*, Roeber S, Stefani A, Naher S, Wagner M, Gieger C, Waldenberger M, Peters A, Högl B, Herms J, Schormair B, Zhao C, Winkelmann J†, Oexle K†. Epigenetic Association Analyses and Risk Prediction of RLS. Mov Disord. 2023 May 22. doi: 10.1002/mds.29440. Epub ahead of print. PMID: 37212434.

Harrer P*, Schalk A*, Shimura M*, Baer S, Calmels N, Spitz MA, Warde MA, Schaefer E, Kittke VMS, Dincer Y, Wagner M, Dzinovic I, Berutti R, Sato T, Shirakawa T, Okazaki Y, Murayama K, Oexle K, Prokisch H, Mall V, Melčák I, Winkelmann J†, Zech M†. Recessive NUP54 Variants Underlie Early-Onset Dystonia with Striatal Lesions. Ann Neurol. 2023 Feb;93(2):330-335. doi: 10.1002/ana.26544. Epub 2022 Nov 18. PMID: 36333996.

Harrer P, Leppmeier V, Berger A, Demund S, Winkelmann J, Berweck S, Zech M. A de novo BCL11B variant case manifesting with dystonic movement disorder regarding the article "BCL11B-related disorder in two canadian children: Expanding the clinical phenotype (Prasad et al., 2020).". Eur J Med Genet. 2022 Nov;65(11):104635. doi: 10.1016/j.ejmg.2022.104635. Epub 2022 Oct 3. PMID: 36202297.

Bartesaghi L*, Wang Y*, Fontanet P, Wanderoy S, Berger F, Wu H, Akkuratova N, Bouçanova F, Médard JJ, Petitpré C, Landy MA, Zhang MD, **Harrer P**, Stendel C, Stucka R, Dusl M, Kastriti ME, Croci L, Lai HC, Consalez GG, Pattyn A, Ernfors P, Senderek J, Adameyko I, Lallemend F†, Hadjab S†, Chrast R†. PRDM12 Is Required for Initiation of the Nociceptive Neuron Lineage during Neurogenesis. Cell Rep. 2019 Mar 26;26(13):3484-3492.e4. doi: 10.1016/j.celrep.2019.02.098. PMID: 30917305; PMCID: PMC7676307.

Reynhout S*, Jansen S*, Haesen D, van Belle S, de Munnik SA, Bongers EMHF, Schieving JH, Marcelis C, Amiel J, Rio M, Mclaughlin H, Ladda R, Sell S, Kriek M, Peeters-Scholte CMPCD, Terhal PA, van Gassen KL, Verbeek N, Henry S, Scott Schwoerer J, Malik S, Revencu N, Ferreira CR, Macnamara E, Braakman HMH, Brimble E, Ruzhnikov MRZ, Wagner M, **Harrer P**, Wieczorek D, Kuechler A, Tziperman B, Barel O, de Vries BBA, Gordon CT, Janssens V†, Vissers LELM†. De Novo Mutations Affecting the Catalytic Cα Subunit of PP2A, PPP2CA, Cause Syndromic Intellectual Disability Resembling Other PP2A-Related Neurodevelopmental Disorders. Am J Hum Genet. 2019 Jan 3;104(1):139-156. doi: 10.1016/j.ajhg.2018.12.002. Epub 2018 Dec 27. Erratum in: Am J Hum Genet. 2019 Feb 7;104(2):357. PMID: 30595372; PMCID: PMC6323609.