

# Identification and characterisation of novel disease-causing genes and genetic modifiers in neuromuscular disorders

## ABSTRACT

Neuromuscular disorders (NMDs) represent a highly heterogeneous group of diseases affecting voluntary movements which are characterized by alterations of the peripheral nervous system and the muscle. Recent advancements in next generation sequencing technology have accelerated the identification of disease-causing variants in the genome. However, despite these advances, more than 40 % of NMD patients remain genetically undiagnosed.

The aim of this Ph.D. project was to identify novel disease-causing and -modifier genes in NMDs. Two novel disease-candidate associated genes were identified and are described in independent chapters: (I) *GPC1* as a putative genetic modifier of Friedreich's Ataxia (FRDA), and (II) *ERC2* as a candidate causative gene for an NMD with a myasthenia-like phenotype.

For *GPC1*, trio-whole genome sequencing (WGS) analysis revealed the biallelic missense variant c.424C>T, p.R142W in the *GPC1* gene in an 8-year-old girl from consanguineous Turkish parents, who presented with unsteady gait and polyneuropathy. Her symptoms started at the age of 3 years with a progressive course. The biallelic variant is located in a homology stretch and was predicted as pathogenic. *GPC1* is a heparan sulphate glycoprotein located at the extracellular side of the plasma membrane, where it regulates multiple signalling cascades, and plays roles in the nervous system development. *GPC1* has not been associated with an NMD previously, and we hypothesized that the biallelic variant may be causative for the disease. In order to identify pathomechanisms caused by the biallelic variant, we investigated the interactomes of *GPC1* WT and the missense mutation by a mass spectrometry-based proteomics approach. A total of 198 proteins were identified to be interacting with *GPC1* wild type (WT), of which 16 showed altered interaction due to *GPC1* R142W. This included CANX, which is a chaperone for N-linked glycosylated proteins; as well as several members of the vacuolar ATPase (V-ATPase) and mammalian target of rapamycin complex 1 (mTORC1) complexes, which are associated with lysosomal function. Importantly, these candidates are identified as novel *GPC1*-interacting proteins. At the age of 10.5 years, the patient developed kyphoscoliosis and dilated cardiomyopathy. Subsequent confirmation of a 104 GAA trinucleotide expansion in the *FXN* gene revealed a Friedreich's ataxia (FRDA). The patient presents with an unusually severe disease course, given the relatively low number of GAA expansion. We therefore hypothesized that the

GPC1 missense mutation might exacerbate FRDA symptoms, possibly through misregulation of GPC1 glycosylation and alteration of lysosomal homeostasis.

For *ERC2*, based on trio-whole exome sequencing (WES) analyses, the biallelic *ERC2* variants (NM\_015576.3: c.2008G>A, p.Ala670Thr and c.2712+1G>C, p.Gln856Profs\*5) were identified as potentially disease-causing in two unrelated patients with comparable myasthenia-like phenotypes. Both patients suffer from distal and proximal muscle weakness with increased fatigability, and for one patient, a motor endplate disorder was confirmed by a repetitive nerve stimulation analysis. Both patients showed disease onset in their early childhood. *ERC2* is highly conserved scaffolding protein in vertebrates, and is abundant in the nervous system functioning primarily at the presynaptic active zone of neurons. To study underlying disease pathomechanisms, we established disease-relevant cell and animal models, including *Erc2* knock-out and knock-in mouse lines, as well as an *ERC2* knock-out human iPSC (hiPSC) line, mimicking the patients' genotypes. Depletion of *ERC2* led to increased *ERC1* levels in mouse brain and spinal cord tissues, which is hypothesized to be due to a compensatory mechanism. Depletion of *ERC2* in the hiPSC-derived motor neurons resulted with increased spontaneous activity based on multielectrode array measurements. Mass-spectrometry based proteomics analyses of *ERC2* knock-out hiPSC derived motor neurons revealed regulation proteins that play roles in glutamate signalling and recycling, as well as  $Ca^{2+}$  and  $K^{+}$  homeostasis. These findings suggest that misregulation of these biological processes could underlie increased neuronal activity *in-vitro* and pathological outcomes at the organism level.

## ZUSAMMENFASSUNG

Neuromuskuläre Erkrankungen (NMDs) stellen eine hochgradig heterogene Gruppe von Krankheiten dar, deren gemeinsames Merkmal darin besteht, dass willkürliche Bewegungen von Patienten durch Veränderungen des peripheren Nervensystems und der Muskulatur beeinflusst werden. Jüngste Fortschritte in der Next Generation Sequencing (NGS) Technologie haben die Identifizierung von krankheitsverursachenden Varianten im Genom beschleunigt. Dennoch bleiben momentan mehr als 40 % der NMD-Patienten genetisch undiagnostiziert.

Ziel dieses Promotionsprojekts war es, neue krankheitsverursachende und -modifizierende Gene bei NMDs zu identifizieren. Zwei neue krankheitsverdächtige assoziierte Gene wurden identifiziert und in unabhängigen Kapiteln beschrieben: (I) *GPC1* als mögliches modifizierendes Gen der Friedreich-Ataxie (FRDA) und (II) *ERC2* als mögliches ursächliches Gen für eine NMD mit einem Myasthenie-ähnlichen Phänotyp.

Für *GPC1* identifizierte eine trio-WGS-Analyse die biallelische Missense-Variante c.424C>T, p.R142W im *GPC1*-Gen eines 8-jährigen Mädchens mit blutsverwandten türkischen Eltern, das sich mit Gangunsicherheit und Polyneuropathie vorstellte. Ihre Symptome begannen im Alter von 3 Jahren und zeigten einen progressiven Verlauf. Diese biallelische Variante, die sich in einem Homologiebereich befindet, wurde als pathogen vorhergesagt. *GPC1*, ein Heparansulfat-Glykoprotein auf der extrazellulären Seite der Plasmamembran, reguliert multiple Signalwege im Nervensystem. *GPC1* wurde bisher nicht mit einer NMD in Verbindung gebracht, und wir stellten die Hypothese auf, dass die biallelische Variante für die Krankheit ursächlich sein könnte. Um die durch die biallelische Variante verursachten Pathomechanismen zu identifizieren, untersuchten wir die Interaktome von *GPC1* WT und der *GPC1* R142W mit Hilfe eines Massenspektrometrie-basierenden Proteomik Ansatzes. Es wurden insgesamt 198 Proteine identifiziert, die mit *GPC1* WT interagieren, von denen 16 aufgrund der Missense-Mutation eine veränderte Interaktion aufwiesen. Dazu gehörten *CANX*, ein Chaperon für N-gebundene glykosylierte Proteine, sowie mehrere Mitglieder des Komplexes der vakuolären ATPase (V-ATPase) und des mammalian target of rapamycin complex 1 (mTORC1), die mit lysosomaler Funktion in Verbindung stehen. Diese Kandidaten wurden als neuartige *GPC1*-interagierende Proteine identifiziert. Die Patientin entwickelte im Alter von 10,5 Jahren eine Kyphoskoliose und eine dilatative Kardiomyopathie, woraufhin anschließend die Bestätigung einer 104 GAA-Trinukleotidexpansion im *FXN*-Gen erfolgte, und basierend darauf eine Friedreich-Ataxie (FRDA) diagnostiziert wurde. Die Patientin zeigt einen ungewöhnlich schweren Krankheitsverlauf, angesichts der relativ geringen Anzahl von GAA-Expansionen. Daher vermuten wir, dass die *GPC1*-Missense-Mutation die FRDA-Symptome

verschlimmern könnte, möglicherweise durch Fehlregulation der GPC1-Glykosylierung und Veränderung der lysosomalen Homöostase.

Für *ERC2* wurden auf der Grundlage von Trio-WES-Analysen die biallelischen *ERC2*-Varianten (NM\_015576.3: c.2008G>A, p.Ala670Thr und c.2712+1G>C, p.Gln856Profs\*5) bei zwei nicht verwandten Patienten mit vergleichbaren Myasthenie-ähnlichen Phänotypen als krankheitsverursachende Kandidaten identifiziert. Beide Patienten leiden an distaler und proximaler Muskelschwäche mit erhöhter Ermüdbarkeit, und bei einem von ihnen wurde eine Störung der motorischen Endplatte durch eine repetitive Nervenstimulation bestätigt. Bei beiden Patienten brach die Krankheit im frühen Kindesalter aus. *ERC2* ist bei Wirbeltieren hoch konserviert und wird hauptsächlich im Nervensystem exprimiert. Es kodiert für ein Gerüstprotein, das vor allem in der präsynaptischen aktiven Zone als Teil eines Komplexes fungiert, welcher eine wichtige Rolle für die Neurotransmission spielt. Aufgrund der Phänotypen der Patienten und der Funktion von *ERC2* an der Präsynapse stellen wir die Hypothese auf, dass *ERC2* ein neues krankheitsverursachendes Gen für NMDs ist. Um die Rolle von *ERC2* im neuromuskulären System und die damit verbundenen Pathomechanismen zu untersuchen, haben wir krankheitsrelevante Zell- und Tiermodelle etabliert. Mithilfe von CRISPR/Cas9-Genom-Editierung war es uns möglich *Erc2*-Knock-out- und Knock-in-Mauslinien, sowie eine humane *ERC2*-Knock-out-iPSC-Linie, zu erzeugen, die die Genotypen der Patienten widerspiegeln. Depletion von *ERC2* führte zu erhöhten *ERC1*-Spiegeln in Mausgehirn- und Rückenmarkgeweben, was auf eine Kompensation hindeutet. Menschliche aus iPSCs differenzierte *Erc2*-Knock-out-Motorneurone zeigten erhöhte spontane Aktivität basierend auf Multielektroden-Array Messungen. Durch Massenspektrometrie-basierte Proteomik Analysen von *ERC2*-Knock-out-iPSCs identifizierten wir veränderte Mengen von Proteinen, die eine Rolle bei der Glutamat-Signalübertragung und dem -Recycling sowie der Ca<sup>2+</sup>- und K<sup>+</sup>-Homöostase spielen. Diese Befunde legen nahe, dass die Fehlregulation dieser Kandidaten eine erhöhte neuronale Aktivität und pathologische Effekte auf organismischer Ebene verursachen könnten.