

Search for novel genes involved in primary microcephaly and functional characterization of KIAA0408

Inaugural-Dissertation

zur

Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln



vorgelegt von

Syeda Seema Waseem

aus

Gujranwala, Pakistan

Köln, 2021

Vorsitzender: Prof. Dr. Siegfried Roth

1. Berichterstatter: Prof. Dr. Peter Nürnberg

2. Berichterstatter: Prof. Dr. Michael Nothnagel

Beisitzer: Dr. Birgit Budde

Mündliche Prüfung: 05.11.2021

The present research work was carried out from June 2015 to August 2021 at the Cologne Center for Genomics (CCG) and Centre for Biochemistry, Institute for Biochemistry I, Medical Faculty, University of Cologne, Cologne, Germany, under the supervision of Prof. Dr. Peter Nürnberg, Prof. Dr. Angelika A. Noegel and Dr. Muhammad Sajid Hussain

Die vorliegende Arbeit wurde in der Zeit von Juni 2015 bis August 2021 unter der Anleitung von Prof. Dr. Peter Nürnberg, Prof. em. Dr. Angelika A. Noegel und Dr. Sajid Hussain am Cologne Center for Genomics (CCG) und Biochemischen Institut I der Medizinischen Fakultät der Universität zu Köln, Köln, Deutschland, angefertigt.

Dedicated to

My family, my beloved and affectionate parents Syeda Faseeha Riffat and Syed Waseem Uddin, my husband Abubakar Moawia and **especially my daughter Hanna Abubakar**. Without their tremendous understanding and encouragement, it would be impossible for me to complete my study.

“I am among those who think that science has great beauty.” – Marie Curie

Acknowledgements

First and foremost, I am extremely grateful to my supervisor, **Prof. Dr. Peter Nürnberg**, director at the Cologne Center for Genomics (CCG), University of Cologne, Germany for his invaluable advice, continuous support, and patience during my academic tenure of PhD. His immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life. I feel privileged to have had him as my supervisor.

I would also like to express my gratitude to my PhD thesis committee members **Prof. Dr. Michael Nothnagel**, **Prof. Dr. Siegfried Roth** and my supervisor **Prof. Dr. Peter Nürnberg** for their valuable advices, contributions and great support through the project and to my education. I would like to express my sincere thanks to **Prof. Dr. Angelika A. Noegel** and **Dr. Muhammad Sajid Hussain** for giving me the opportunity to carry out my PhD-related biochemical experiments at institute for Biochemistry I, University of Cologne.

My special thanks go to **Prof. Dr. Shahid Mahmood Baig**, S.I., chairman at Pakistan Science Foundation for his intensive support during this work. His friendly and professional assistance ensured a successful and nice cooperation with his team in Pakistan.

My great appreciation goes to the Graduate School for Biological Sciences at University of Cologne for providing me equal opportunities to contribute to various skill training programs. Also, I gratefully acknowledge all my colleagues especially **Nina Dalibor** and **Gerti Meyer** at the Cologne center for Genomics (CCG) for their valuable advices that have made my study and life in the Germany a wonderful time. I am also thankful to the faculty at the Cologne Center for Genomics (CCG) and Institute for Biochemistry I, CMMC and CECAD at University of Cologne for their project related support. My gratitude extends to **Dr. Birgit Budde** at CCG whom fruitful discussions and valuable suggestions have contributed greatly to the quality of this work. My cordial thanks go to **Gabriele Thorn** for her tremendous support in official work and for her generous cooperation in every matter.

I would like to extend my sincere and special thanks to **Dr. Muhammad Tariq** and his team at National institute for Biotechnology and Genetic Engineering (NIBGE),

Faisalabad, Pakistan for their kind cooperation and a precious help in providing clinical investigations necessary to accomplish the research work of my PhD study and publications successfully.

This research work sounds incomplete without acknowledging my collaborators **Univ.- Prof. Jay Gopalakrishnan and Dr. Elke Gabriel** for their insightful suggestions that contributed greatly to the productivity of my PhD research work. My great appreciation also goes to the patients and the families who participated in this study.

Above all, I would like to express my profound gratitude to my beloved parents, my husband, my daughter and my brother **Syed Abdul Azeem** for their never-ending encouragement and confidence and for helping me any time I needed and for standing all this time beside me, for their patience and affection. This work is dedicated to you

Last but not least, I express my deepest thanks to my friends especially **Dr. Shivani Rana** and **Maria Iqbal** who supported and encouraged me at any time.

Syeda Seema Waseem

Summary

Microcephaly is the consequence of a small but architecturally normal brain and has two types; congenital microcephaly (CM) and postnatal microcephaly. While CM is characterized by a reduced head circumference at birth and non-progressive mental retardation, a microcephaly that developed only after birth is usually associated with further anomalies and also referred to as syndromic microcephaly. CM is usually caused by single-gene defects and mostly referred to as autosomal recessive primary microcephaly (MCPH). MCPH is a rare and heterogeneous neurodevelopmental disease that arises from a deficient production of cortical neurons within the neurogenic epithelium during human brain development. As of July 2021, the Online Mendelian Inheritance in Man (OMIM) database lists 27 genes to be implicated in MCPH. Most of these MCPH genes encode centrosomal proteins or have a link to the cell cycle machinery. The discovery of most of these 27 genes became only possible due to tremendous advances in exome and genome sequencing that have been expediting the discovery of disease-causing variants in novel genes over recent years. However, many families are still waiting for a genetic diagnosis and we are definitely far from completely understanding the molecular mechanisms underlying microcephaly.

In this study, I aimed at dissecting the genetic cause of microcephaly in 8 families from Pakistan afflicted with this disorder. In 7 families I was able to identify the disease-causing variants using a combined approach of linkage analysis and exome sequencing. Interestingly, a total of four novel genes (*AKNA*, *KBTBD6*, *TBC1D8B*, *KIAA0408*) could be associated with MCPH in the study cohort. While sufficient information was available in the literature for *AKNA*, *KBTBD6* and *TBC1D8B* to interpret the identified variants and genes in terms of their roles in disease aetiology, hardly anything was known about *KIAA0408* prior to my investigations. Therefore, I focused my efforts on the functional characterization of this novel and unknown MCPH-associated gene.

A homozygous *KIAA0408* missense variant (NM_014702.4: c.1640T>C, p.Leu547Ser) was identified as the most likely cause of MCPH in a family with two affected siblings recruited from Northern Punjab. Segregation analysis and homozygosity mapping corroborated this assumption. *In silico* analyses predicted the novel *KIAA0408* variant to be disease-causing, damaging or deleterious, depending on the tool used for prediction.

Though KIAA0408 was a largely uncharacterized protein at that time, I found it listed in centrosomeDB (<http://centrosome.dacya.ucm.es>), a database containing all centrosome associated proteins. Using immunofluorescence microscopy, I could verify its centrosomal localization during interphase and mitosis in different cell lines (HEK293, RPE1 and human fibroblasts). I saw KIAA0408 also on the mid body during cytokinesis. Hence, I could define KIAA0408 as a novel centrosome and mid-body protein. I also documented the cellular localization of KIAA0408 in mutant fibroblasts during interphase and at different stages of mitosis. Furthermore, I also determined KIAA0408 as a basal body protein in RPE1 cells (ciliary cells). Several cellular phenotypes like an abnormal number of centrosomes, misshapen nuclei and multipolar spindles were also noticed in mutant cells. Additionally, quantitative PCR and western blotting revealed overexpression and an increased amount of KIAA0408 in mutant cells. Interestingly, the amount of other centrosomal proteins was also increased in mutant cells as compared to wild-type cells.

The most insightful step of this project was to reprogram patient fibroblasts with the *KIAA0408* variant into induced pluripotent stem cells (iPSCs) and then differentiate them into neural progenitor cells (NPCs). This approach aimed at investigating the effect of the identified *KIAA0408* variant on the differentiation of the NPCs during brain development. Further experiments validated the premature differentiation of NPCs into neurons for mutant cells as compared to the wild-type ones, which would explain the reduced cerebral cortex observed in the patients with the *KIAA0408* variation. Notably, the observed premature switching of NPCs to neurons in mutant cells would result in thinning of the cerebral cortex, thus suggesting premature cortical neurogenesis as the pathomechanism in patients of the family affected by the *KIAA0408* variant.

Taken together, validated genetic and functional data of this study provide valuable insights into the complex nature of human brain development. Moreover, this study also expands the genotypic and phenotypic spectra of MCPH. In particular, it also highlights the role of genetic modifiers in determining the severity of the expressed microcephaly phenotype.

Zusammenfassung

Mikrozephalie ist die Folge eines zu kleinen, aber architektonisch normal gestalteten Gehirns. Es werden zwei Formen unterschieden: kongenitale Mikrozephalie (CM) und postnatale Mikrozephalie. Während CM durch einen bereits bei der Geburt verringerten Kopfumfang und eine nicht-progressive geistige Retardierung gekennzeichnet ist, entwickelt sich die postnatale Mikrozephalie erst nach der Geburt und ist in der Regel mit weiteren Anomalien assoziiert und wird daher auch als syndromale Mikrozephalie bezeichnet. CM wird in der Regel durch einen einzelnen Gendefekt verursacht und meist als autosomal rezessive primäre Mikrozephalie (MCPH) bezeichnet. MCPH ist eine seltene und heterogene neuronale Entwicklungsstörung, die durch eine verminderte Produktion von kortikalen Neuronen innerhalb des neurogenen Epithels entsteht. Stand Juli 2021 weist die *Online Mendelian Inheritance in Man* (OMIM)-Datenbank 27 Gene aus, die mit MCPH in Zusammenhang stehen. Die meisten dieser MCPH-Gene kodieren für zentrosomale Proteine oder haben eine Verbindung zur Zellzyklusmaschinerie. Die Entdeckung der meisten dieser 27 Gene wurde erst durch die enormen Fortschritte bei der Exom- und Genomsequenzierung möglich, die in den letzten Jahren die Entdeckung von krankheitsverursachenden Varianten in neuen Genen beschleunigt haben. Dennoch warten noch immer viele Familien auf eine genetische Diagnose, und wir sind weit davon entfernt, die molekularen Mechanismen, die der Mikrozephalie zugrunde liegen, vollständig zu verstehen.

In dieser Studie habe ich versucht, die genetische Ursache der Mikrozephalie in 8 Familien aus Pakistan, die von dieser Krankheit betroffen sind, zu entschlüsseln. In 7 Familien konnte ich die krankheitsverursachenden Varianten mit einem kombinierten Ansatz aus Kopplungsanalyse und Exom-Sequenzierung identifizieren. Interessanterweise konnten in der Studienkohorte insgesamt vier neue Gene (*AKNA*, *KBTBD6*, *TBC1D8B*, *KIAA0408*) mit MCPH in Verbindung gebracht werden. Während für *AKNA*, *KBTBD6* und *TBC1D8B* in der Literatur ausreichend Informationen verfügbar waren, um die identifizierten Varianten und Gene im Hinblick auf ihre Rolle in der Ätiologie der Krankheit zu interpretieren, war über *KIAA0408* vor meinen Untersuchungen kaum etwas bekannt. Daher konzentrierte ich meine Bemühungen auf die funktionelle Charakterisierung dieses neuen und unbekanntes MCPH-assoziierten Gens.

In einer Familie mit zwei betroffenen Geschwistern aus dem nördlichen Punjab wurde die homozygote KIAA0408-Missense-Variante NM_014702.4:c.1640T>C,p.Leu547Ser als wahrscheinlichste Ursache für die MCPH identifiziert. Segregationsanalysen und Homozygotiekartierung bestätigten diese Vermutung. Verschiedene In-silico-Analysen sagten die neuartige Variante als krankheitsverursachend, schädlich oder nachteilig voraus, je nachdem, welches Tool für die Vorhersage verwendet wurde.

Obwohl KIAA0408 zu diesem Zeitpunkt ein weitgehend uncharakterisiertes Protein war, fand ich es in der centrosomeDB (<http://centrosome.dacya.ucm.es>), einer Datenbank, die alle mit dem Zentrosom in Verbindung stehenden Proteine enthält. Mittels Immunfluoreszenzmikroskopie konnte ich seine zentrosomale Lokalisierung in verschiedenen Zelllinien (HEK293, RPE1 und menschliche Fibroblasten) während der Interphase und Mitose nachweisen. Während der Zytokinese war KIAA0408 auch am „mid body“ zu sehen. Damit konnte ich KIAA0408 als ein neuartiges Zentrosom- und „mid body“-Protein definieren. Ich dokumentierte auch die zelluläre Lokalisierung von KIAA0408 in mutierten Fibroblasten während der Interphase und in verschiedenen Stadien der Mitose. Darüber hinaus habe ich KIAA0408 auch als Basalkörperprotein in RPE1-Zellen (Ziliarzellen) nachgewiesen. Mehrere zelluläre Phänotypen wie eine abnorme Zahl von Zentrosomen, missgestaltete Kerne und multipolare Spindeln wurden ebenfalls in den mutierten Zellen festgestellt. Darüber hinaus zeigten quantitative PCR und Western Blotting eine Überexpression bzw. eine erhöhte Menge von KIAA0408 in den mutierten Zellen. Interessanterweise war auch die Menge anderer zentrosomaler Proteine in den mutierten Zellen im Vergleich zu Wildtyp-Zellen erhöht.

Der aufschlussreichste Schritt dieses Projekts bestand darin, Patientenfibroblasten mit der KIAA0408-Variante in induzierte pluripotente Stammzellen (iPSCs) umzuprogrammieren und sie dann in neurale Vorläuferzellen (NVZ) zu differenzieren. Mit diesem Ansatz sollte die Wirkung der identifizierten KIAA0408-Variante auf die Differenzierung der NVZ während der Gehirnentwicklung untersucht werden. Weitere Experimente bestätigten eine vorzeitige Differenzierung von NVZ in Neuronen bei den mutierten Zellen im Vergleich zu den Wildtyp-Zellen, was die bei Patienten mit der KIAA0408-Variante beobachtete reduzierte Großhirnrinde erklären würde. Insbesondere würde die beobachtete vorzeitige Umwandlung von NVZ in Neuronen bei mutierten Zellen zu einer Ausdünnung der Großhirnrinde führen, was auf eine

vorzeitige kortikale Neurogenese als Pathomechanismus bei den Patienten der von der KIAA0408-Variante betroffenen Familie hindeutet.

Zusammengenommen bieten die validierten genetischen und funktionellen Daten dieser Studie wertvolle Einblicke in die komplexe Natur der menschlichen Gehirnentwicklung. Darüber hinaus erweitert diese Studie auch das genotypische und phänotypische Spektrum von MCPH. Insbesondere unterstreicht sie die Rolle genetischer Modifizier bei der Ausprägung der Schwere des Mikrozephalie-Phänotyps.