Molecular Genetics of Primary Microcephaly and Microcephalic Primordial Dwarfism in Consanguineous Families from Pakistan

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Ilyas Ahmad

aus Dir Lower, Khyber Pukhtoonkhwa (Pakistan)

2016

This thesis is dedicated to my beloved mother, father, brothers and sisters

Acknowledgements

First and foremost, I would like to express my gratitude and appreciation to Prof. Dr. Peter Nürnberg, who has been an excellent supervisor and has supported me in every way throughout these four years. His active and engaged supervision was preeminent to finalize my PhD thesis. Thank you for your guidance, immense patience and valuable suggestions to complete this journey. I have had a lovely time in the lab and feel lucky to have had him as my mentor. My sincere indebtedness also goes to my co-supervisor Prof. Dr. Angelika Anna Noegel who has been helpful and patient to me; without her involvement, this work will not have been possible. Her encouragement and kindness will always be remembered. She was a tireless proofreader at various stages of this thesis. Thank you again for your quick feedback during the writing and revision process. I am also obliged to Prof. Dr. Hartmut Arndt and Dr. Holger Thiele for agreeing to be part of my exam committee on a very short notice.

I also extends my sincere regards to Prof. Shahid Mahmmod Baig, National Institute for Biotechnology and Genetic Engineering (NIBGE), Pakistan for providing me the opportunity to work on this collaborative project. I am also thankful for providing their assistance and logistical support during my sampling visits in Pakistan. Also thanks given to Dr. Muhmmad Sajid Hussain for supervising all the experiments at Biochemistry 1 and always helpful with ideas over the last four years. Thanks for teaching me how to use adobe photoshop and image analysis, and to edit this thesis.

Next, I would like to thank all the participating patients and their family members in this study. I would also extend my acknowledgments to the members of Human Molecular Genetics lab, NIBGE, Pakistan for their help in DNA extraction and gathering clinical information. I am also obliged to all the informants who gave me valuable information about patients during my field work. I warmly thank to all the present and past members of the CCG for providing such a warm working atmosphere. In particular I would like to thank: Dr. Mohammad Reza Toliat for whole-genome genotyping; Dr. Bigirt Budde and Gudrun Nürnberg for homozygosity mapping and linkage analysis; Dr. Holger Thiele and other members of bioinformatics team for whole exome sequencing; Hamid Shabbir for help with qPCR; Nina Dalibor, Elisabeth Kirst and Gerti Meyer for all technical assistance. I am sincerely thankful to Dr. Wilfred Gunia for their quick response to provide IT support and Holger Trucks and Dennis Lal for helpful discussion regarding various genetic tools and concepts.

Table of contents

1. Abstract

The centrosome is known as the coordinator of key organizational processes of a cell. Mutations in genes of centrosome associated proteins have been found to cause primary microcephaly (MCPH) or related growth pathologies such as microcephalic primordial dwarfism syndromes in humans. Defects in mitotic progression, chromosome segregation, microtubule assembly and DNA damage-induced checkpoint arrests are the proposed links for the underlying mechanisms. Genes that are part of networks involved in the regulation of the cell cycle can be potential candidates for both diseases. To identify new genes involved in the disease process, twenty nine consanguineous MCPH families and a family with microcephalic primordial dwarfism were recruited from various areas of Pakistan. Homozygosity mapping complemented with exome sequencing or targeted gene sequencing identified several mutations in known MCPH associated genes and in a new gene, *PLK4*. In case of MCPH, mutations were identified in 28 families linked to known loci. In one family sequencing of the whole exome failed to detect a mutation, suggesting the involvement of regulatory regions. Nine novel mutations were found in *ASPM*, two in *MCPH1* and two in *CDK5RAP2*. Three known mutations were found in *ASPM* and one in *WDR62*. Furthermore, the exact breakpoints were mapped in two families carrying overlapping homozygous microdeletions of 164 kb and 577 kb at the MCPH1 locus.

A biallelic loss-of-function mutation, p.Arg936Serfs*1, was identified near the C-terminus of PLK4 in a family segregating inherited microcephalic primordial dwarfism. PLK4 kinase is an essential regulator of centriole biogenesis. Cellular experiments carried out with patient-derived fibroblasts demonstrated a significant reduction at the protein level. Centriole biogenesis was significantly impaired and resulted in monopolar configurations, centriole loss and defective spindle formation. Further, overexpression studies revealed that mutant PLK4 still localized to the centrosome but centriole duplication was significantly impaired. Moreover, patient cells showed a reduced cellular proliferation and delayed mitotic progression. Combining genetic data with the results of cellular experiments, *PLK4* was established as a new disease-associated gene in microcephalic primordial dwarfism.

This study adds to the mutational spectra of MCPH-associated genes and quantifies the genetic heterogeneity in the Pakistani population. It further confirms the importance of the centrosome for normal brain and body size and extends the genetic understanding and phenotypic spectrum associated with genes relevant for centriole biogenesis.

2. Zusammenfassung

Zentrale organisatorische Prozesse der Zelle werden vom Zentrosom koordiniert. Tragen die Gene, die für zentrosomassoziierte Proteine kodieren, Mutationen, so kann das zu primärer Mikrozephalie (MCPH) oder ähnlichen Wachstumsstörungen, wie z. B. den mikrozephalen primordialen Kleinwuchssyndromen (MPD), führen. Die zugrundeliegenden Defekte führen zu Störungen der mitotischen Progression, der Chromosomensegregation und der Mikrotubuli-Anordnung sowie zu Arretierungen an Zellzykluskontrollpunkten, die mit DNA-Schäden im Zusammenhang stehen. Grundsätzlich sind alle Gene für Komponenten von Netzwerken, die den Zellzyklus regulieren, potentielle Kandidaten für die Verwicklung in die Ausprägung derartiger Erkrankungen. Um neue involvierte Gene zu identifizieren, wurden 29 konsanguine MCPH-Familien und eine MPD-Familie aus verschiedenen Regionen Pakistans rekrutiert. Mit Hilfe von Homozygotie-Kartierungen ergänzt mit Exom- bzw. Kandidatengen-Sequenzierungen konnten verschiedene Mutationen in bekannten MCPH-assoziierten Genen sowie in einem neuen Gen (*PLK4*) identifiziert werden. Bezogen auf MCPH wurden in 28 Familien, die mit bekannten Loci koppelten, Mutationen gefunden. In einer Familie konnte die kausale Mutation trotz der Sequenzierung des gesamten Exoms nicht detektiert werden. Hier könnte die Mutation in einer der nicht erfassten regulatorischen Regionen liegen. Es wurden zahlreiche neue Mutationen gefunden, neun in *ASPM*, zwei in *MCPH1* und weitere zwei in *CDK5RAP2*. Weiterhin wurden drei bekannte Mutationen in dem Gen *ASPM* und eine bekannte Mutation in *WDR62* gefunden. Darüber hinaus konnten in zwei Familien die genauen Bruchpunkte einer 164-kb- und einer 577-kb-Deletion am MCPH1-Locus kartiert werden, die einander überlappten und in den jeweiligen Familien homozygot vorlagen.

In der MPD-Familie wurde in der Nähe des C-Terminus von PLK4 die biallelische Funktionsverlustmutation p.Arg936SerFs*1 identifiziert. Die PLK4-Kinase ist ein essentieller Regulator der Centriol-Biogenese. Fibroblastenzellen von betroffenen Probanden wiesen eine signifikant reduzierte Proteinmenge auf. Außerdem war die Centriol-Biogenese stark beeinträchtigt. Dies resultierte in monopolaren Strukturen, einem Verlust von Centriolen und einer fehlerhaften Spindelbildung. Des Weiteren haben Überexpressionsstudien gezeigt, dass das mutierte PLK4-Protein zwar noch zum Zentrosom lokalisiert, jedoch die Fähigkeit der Kinase, die Centriol-Duplikation zu befördern, signifikant eingeschränkt ist. Darüber hinaus zeigten die Patientenzellen eine reduzierte Proliferation und eine verzögerte mitotische Progression. Die genetischen Daten zusammen mit den Resultaten der zellulären Experimente haben *PLK4* überzeugend als ein neues MPD-assoziiertes Gen etabliert.
Die vorliegende Studie erweitert die Mutationsspektren einiger MCPH-assoziierter Gene und quantifiziert die genetische Heterogenität in der Pakistanischen Bevölkerung hinsichtlich der

untersuchten seltenen Erkrankungen. Diese Arbeit bestätigt außerdem die große Bedeutung der Zentrosomenfunktion für eine normale Gehirnentwicklung und Körpergröße. Sie erweitert ferner unser Wissen über die für die Centriol-Biogenese relevanten Gene sowie das phänotypische Spektrum, das mit diesen assoziiert sein kann.

3.Introduction

The word 'genetics' was coined by William Bateson after the rediscovery of Gregory Mendel's work at the beginning of the 20th century (Harper 2005). Archibald Garrod established the first link between a medical condition and Mendel's research. He recognized the pattern of Mendelian inheritance in human diseases and published his observation on alkaptonuria in 1902 (Kennedy 2001). It sparked a new scientific quest to understand the biology of genetic information and content. Scientific progress revolutionized human genetics after the strenuous efforts in a time frame of 100 years. It started from the understanding of the basic rules of hereditary to the ultimate mapping of the complete sequence of the human genome (Kennedy 2001).

Human genetic diseases have traditionally been divided into 'monogenic, simple and rare' and 'multigenic, complex and common' diseases (Bamshad et al. 2011). The elucidation of the underlying molecular cause of disease phenotypes is the major goal of medical genetics. Positional cloning was employed successfully to identify the mutated genes underlying human monogenic diseases (Botstein and Risch 2003). The development of DNA sequencing technologies and target enrichment strategies resulted in whole exome sequencing projects, which further expedited the identification of mutations in rare conditions (Ng et al. 2009). There are about 7,000 rare monogenic diseases and the molecular etiology is known only for >3,500 (~ 50%) (Boycott et al. 2013). Still, a large chunk of monogenic conditions with missing links is present in isolated and inbred populations across the world (Antonarakis and Beckmann 2006). At present, the realistic aim is to find the candidate gene for every recessive disorder through global coordination and cooperation. This will not only provide enormous service to the families but also provide many novel entry points to understand the pathomechanisms of the diseases (Bamshad et al. 2011).

3.1 Pakistani Population

Pakistan lies on the postulated costal route with the evidence of settlers from 60,000 years ago (Mehdi et al. 1999). The country borders India in the east, Afghanistan in the north-west, Iran in the west, China in the north and the Arabian Sea in the south. Pakistan has world's sixth largest population of 182 million people (WHO, 2013) who belong to at least 18 ethnic groups and speak >60 different languages (Grimes and Grimes 1993). In the different periods of history, Aryans, Macedonians, Arabs and Mongols settled after invasion and add to considerable ethnic diversity. The population is divided into Punjabi (44.15%), Pashtuns (15.42%), Serakis (10.53%), Kashmiri, Balti, Burushes and Hazaras in the north, and Balochis (3.57%), Brahuis, Makranis and Sindis (14.1%) in the south (Mehdi et al. 1999; Syed and

Shehla 2011). After the 1947 partition, many ethnic groups migrated from India who are grouped as Muhajir (7.57%) (Syed and Shehla 2011).

Ethno-cultural diversity had influenced the settlement and development of communities in the sub-continent (Ibbetson 1982). Consanguineous marriages are highly favoured due to different sociocultural, ethnic and geographical background. In Pakistan, 60% of marriage unions are consanguineous (Akram et al. 2008; Hussain 1999). Cousin marriages can lead to the multigenerational pedigrees with increase prevalence of recessive disorders, spontaneous abortion and other congenital malformations (Devi et al. 1987; Hussain 1999). Two independent studies found various congenital malformations in 14.79% and 40% cases where the parents were first or second cousins (Akram et al. 2008; Hafeez et al. 2007). Beta-thalassemia is the most common inherited disorders with carrier frequency >5.6% and 9 million estimated carriers in Pakistan (Baig et al. 2012). The other most common inherited conditions are microcephaly and deafness. It is estimated that the primary microcephaly and profound bilateral hearing loss occur for 1 and 1.6 per 1000 births respectively (Ali 2010; Elahi et al. 1998; Woods 2004).

The high degree of consanguinity and different ethnicities constitute a unique resource to reveal the genetic basis of many traits (Hussain 1999; Mehdi et al. 1999). At present, many Pakistani families with multiple affected births having congenital inherited traits are left undiagnosed. Moreover, information and counselling regarding recurrence rates is never provided to the parents. In addition to counselling and symptomatic diagnosis, prenatal diagnosis is also a legal option in an Islamic society like Pakistan (Baig et al. 2012). In the current study, I focused on the molecular genetics of two neurodevelopmental disorders, primary microcephaly and microcephalic primordial dwarfism in the Pakistani population. The use of powerful tools in the field of modern human genetics provided us the best opportunity to identify the disease-causing variants in genes essential for embryonic development.

3.2 Microcephaly

Humans have large, convoluted and highly specialized brain structure which makes them a distinguished species on earth. The human brain associates with cognition, emotions and the power to orchestrate different thoughts always puzzled scientists and philosophers, and remained the most intriguing question for them (Genin et al. 2012; Hu et al. 2014). On the contrary, microcephaly is known to humans for a long time but it was only at the end of 19th century that attracted researchers to gain insights about the underlying pathomechanism of the disease (Opitz and Holt 1990). The term microcephaly is originated from Greek language and it is the combination of two words "Mikros" means small and "Kephalos" for head or skull. So, people born with smaller brain and head size than usual are microcephalic.

Microcephaly can be broadly divided on the basis of aetiology in to primary and secondary microcephaly. Primary microcephaly occurs at 32 weeks of gestation and has genetic background. Microcephaly can occur isolated, with no obvious abnormities, or it may be part of a syndrome of congenital anomalies, and therefore associated with other abnormities (Abuelo 2007; Passemard et al. 2013; Woods 2004). In secondary microcephaly, the child is born with normal head size but postnatal growth fails. Secondary microcephaly occurs through non genetic causes including environmental insults such as exposure to alcohol (foetal alcohol syndrome), toxins, foetal irradiation, placental insufficiency or congenital infections, (e.g. CMV, toxoplasmosis, HIV, rubella) and, rarely, maternal metabolic disorders e.g. maternal Phenylketonuria (PKU).

3.2.1 Primary microcephaly

Autosomal recessive primary microcephaly (MCPH, MIM 251200) or "microcephaly vera" is a genetically heterogeneous human brain disorder and a rare subgroup of congenital microcephaly (Tan et al. 2014a; Woods et al. 2005). It is defined as a reduced occipital frontal circumference (OFC) at birth that is less than four standard deviations (-4SD) below the age, sex and ethnicity (Woods et al. 2005).The affected individuals have varying degree of intellectual disability without other neurological findings. The presence of fits does not exclude the diagnosis. MCPH appears to be due to a failure of normal developmental processes rather than regression or degeneration in the neuronal tissue. In general, the architecture of the brain is not grossly affected (Thornton and Woods 2009; Woods et al. 2005) although in some cases structural defects in neural migration have been reported (Bond et al. 2002; Hussain et al. 2013).

The prevalence of microcephaly varies in different populations. The incidence of MCPH is 1 in a million in the Yorkshire region of Britain, 1/30,000 in Japan, 1/250,000 in Holland and 1/ 2000,000 in Scotland (Komai et al. 1955; Tolmie et al. 1987; Van Den Bosch 1959).The prevalence of MCPH is higher in populations where consanguinity is widely practiced. The highest incidence of 1/10,000 is present in families originating from Northern Pakistan. The overall incidence is higher in South Asian and Arabic populations than Caucasian population (Woods et al. 2005). Genetic heterogeneity was expected in MCPH due to the broad spectrum of phenotypes (Cowie 1960). Majority of MCPH cases show normal height, chromosomal analysis, weight and MRI scan without gross structural abnormalities. The growing number of MCPH genes has increased the phenotypic spectrum like height reduction, chromosome condensation, lisencephaly, schizencephaly, polymicrogyria and hypoplasia of corpus callosum. Therefore, the diagnosis of MCPH no longer excludes the presence of these phenotypes (Woods et al. 2005).

Figure 3.1: MCPH patient from Pakistan **(a)** front view **(b)** side view showing the reduced head size and sloping forehead.

3.2.1.1 Molecular genetics of primary microcephaly

Primary microcephaly is a genetically heterogeneous disorder. So far, thirteen genes have been associated with primary microcephaly, eight of which have been published in the last six years as a result of rapid advancement of next generation sequencing and high–throughput deep sequencing (Hussain et al. 2013; Khan et al. 2014). All known genes for MCPH are summarized in Table 3.1. *ASPM* is the most common implicated gene in MCPH (Nicholas et al. 2009). Mutations in *ASPM* had been reported in many populations, including Pakistan, Iran, India, Middle Eastern, Caucasian, Turkey, African, African American, Hispanic (Bond et al. 2003; Darvish et al. 2010; Kumar et al. 2004; Nicholas et al. 2009; Tan et al. 2014a). The majority of the MCPH proteins are associated with centrosome or spindle pole and are the part of many biological processes from centrosome maturation, centriole biogenesis, spindle position, to DNA repair (Thornton and Woods 2009). In the following lines, I will briefly discuss the genes where I found mutations in the current study.

3.2.1.1.1 Microcephalin (*MCPH1*)

The first molecular defect was identified in *MCPH1* in families of Pakistani origin (Jackson et al. 2002). Microcephalin, the *MCPH1* product, consists of 835 amino acids with three BRCA1-carboxy terminal domains (BRCT domains) (Jackson et al. 2002). Microcephalin has been shown to play a variety of roles in cell cycle regulation of neuronal progenitors and DNA damage response pathways. It has been demonstrated that microcephalin co-localises with DNA damage checkpoint 1 (MDC1) protein and γH2AX in foci induced by ionizing radiation which are required for the recruitment of many downstream DNA damage response proteins (Lin et al. 2005; Xu et al. 2004). The phenotypic spectrum of *MCPH1* mutations is not only limited to primary microcephaly but several studies have been conducted to address the genetic, epigenetic and functional role of *MCPH1* in various forms of cancer (Venkatesh and Suresh 2014).

3.2.1.1.2 WD repeat-containing protein 62 (*WDR62*)

WDR62 (MCPH2; OMIM 604317) is the second most common genetic cause of primary microcephaly. The *WDR62* consists of 32 exons and encodes a 1523 amino acids long protein. Mutations in *WDR62* account for 10% of the MCPH cases (Nicholas et al. 2010; Thornton and Woods 2009). Various mutations in *WDR62* can be associated with additional cortical abnormalities including lissencephaly, polymicrogyria and heterotopia. WDR62 is associated with the mitotic spindle poles and is essential for the proliferation of neural progenitor cells (NPCs) (Bilguvar et al. 2010; Nicholas et al. 2010; Yu et al. 2010). Depletion experiments in mice showed spindle instability, mitotic arrest and cell death for NPCs and displayed the phenotypes of primordial dwarfism and microcephaly (Chen et al. 2014).

3.2.1.1.3 Cyclin-dependent kinase 5 regulatory subunit-associated protein 2 (*CDK5RAP2*)

CDK5RAP2 encodes a 1893 amino acid, 215 kDa, pericentriolar material (PCM) protein that is important for the recruitment of other PCM components. It recruits the γ-tubulin ring complex (the microtubule nucleator) to the centrosome through its N-terminus (Bond et al. 2005; Fong et al. 2008) and interacts with Pericentrin through C-terminus (Wang et al. 2010). It also plays a pivotal role in centrosome maturation (Dobbelaere et al. 2008) and centrosome cohesion (Graser et al. 2007) and regulates the assembly of microtubules (Fong et al. 2008). Mutations in human *CDK5RAP2* caused MCPH phenotypes (MCPH3; OMIM 604804). The patients showed reduced brain size without other cortical abnormalities (Bond et al. 2005). It has been shown that Cdk5rap2 is highly expressed in the mouse embryonic brain and is essential for the survival of cortical progenitors and their proliferation. Its depletion leads to a reduction in the proportion of apical progenitor cells. There was also an increased level of apoptosis observed in Cdk5rap2 deficient mouse embryos (Buchman et al. 2010; Lizarraga et al. 2010a).

3.2.1.1.4 Abnormal spindle-like microcephaly- associated protein (*ASPM*)

ASPM, the human orthologue of Drosophila melanogaster abnormal spindle gene (asp), is the most frequently mutated gene in >50% cases with MCPH phenotypes (Thornton and Woods 2009). It is involved in the organization and binding of microtubules at spindle poles and localises to the central spindle during meiosis and mitosis (Riparbelli et al. 2002; Ripoll et al. 1985). It is highly expressed during the cerebral cortical neurogenesis especially in the ventricular zone in the mouse brain (Bond et al. 2002). The cortical size is then determined by the innate property of neural progenitor cells to switch from symmetric proliferation to asymmetric. In the absence of ASPM, cleavage plane is changed from the normal ventricular

11

orientation and result in a premature switch to asymmetric division of NE cells (Fish et al. 2006). The aspm mutant mice exhibit microcephaly and loss of germ cells which were rescued with the human ASPM transgene (Pulvers et al. 2010).

Table 3.1: List of known MCPH genes.

Gene	Disease locus	Cytogenetic location	OMIM #	Reported in Population	Protein localization
MCPH1	MCPH1	8p23	251200	Pakistani	centrosome and nucleas
WDR62	MCPH2	19q13.12-q13.2	604317	Pakistani	mitotic spindle pole
CDK5RAP2	MCPH3	9q33.2	604804	Pakistani	centrosome
CASC5	MCPH4	15q14-15.1	604321	Moroccan	kinetochore
ASPM	MCPH5	1q31.3	608716	Pakistani	mitotic spindle and mid body
CENPJ	MCPH6	13q12.12	608393	Brazilin	centrosome
STIL	MCPH7	1p33	612703	Indian	centrosome
CEP135	MCPH8	4p14-q12	614673	Pakistani	centrosome
CEP152	MCPH9	15q21.1	614852	Pakistani	centrosome
ZNF335	MCPH10	20q13.12	615095	Israeli	nucleus
PHC1	MCPH11	12p13.31	615414	Saudi Arabian	nucleus
CDK6	MCPH12	7q21.2	616080	Pakistani	centrosome
SASS6	MCPH14	1p21.2	616402	Pakistani	centrosome

CENPE (MCPH13, MIM 117143) was omitted from the above list because the gene is reported with different phenotypes.

3.2.2 Microcephalic primordial dwarfism (MPD)

Primordial dwarfism is a group of monogenic disorders in which growth of the individuals is restricted from the prenatal stage onward and is continuously reduced after birth. These individuals were described as "miniature construction with approximately normal anthropometric proportion" or were the centre of attraction in sideshows and circuses in the 19th and 20th century (Hall 2013; McKusick 1955). When the individuals have a reduced head size in proportion to or smaller than the body size, the condition constitutes a separate distinguished group, microcephalic primordial dwarfism (Fig 3.2-A) (Klingseisen and Jackson 2011). This group encompasses several subtypes of human genetic disorders including Seckel syndrome (Majewski et al. 1982; Seckel 1960), Microcephalic Osteodysplastic Primordial Dwarfism (MOPD), type I and II (Hall et al. 2004; Majewski et al. 1982) and Meier-Gorlin Syndrome (MGS) (Gorlin et al. 1975). There is a substantial number of microcephalic primordial dwarfism cases which do not fit into the current diagnostic scheme. However, all these disorders have autosomal recessive segregation and the cardinal features include intrauterine growth restriction, proportionate dwarfism, severe microcephaly and varying degrees of intellectual impairment (Hall et al. 2004; Klingseisen and Jackson 2011).

Genes that are described for microcephalic primordial dwarfism are listed in Table 2. *PLK4* is the new gene identified in a consanguineous Pakistani family in this study. This list includes some of the genes which are also involved in isolated primary microcephaly cases. To date, the underlying cause for these genetic overlaps is unknown. This is a very intriguing question that will need to be addressed in the future.
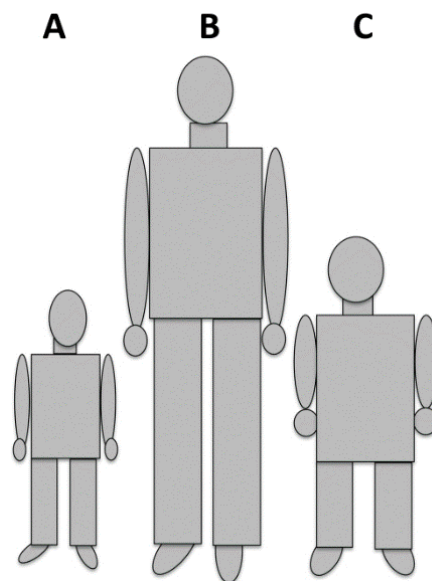
Figure 3.2: Scaling of height and head size to normal average adult: **(A)** depicts microcephalic primordial dwarfism. **(B)** Normal adult individual. **(C)** Primordial dwarfism. Adopted from (Klingseisen and Jackson 2011).

Table 3. 2: Summary of the identified genes in MPD

Type of MPD	Genes involved	Disease locus	OMIM*	Cytogenetic location	Protein localization
Seckel syndrome	*ATR*	SCKL1	210600	3q23	centrosome and nucleus
	RBBP8	SCKL2	606744	18q11.2	nucleus
	ATRIP	-		3p21.31	nucleus
	CENPJ	SCKL4	613676	13q12.2	centrosome
	CEP152	SCKL5	613823	15q21.1	centrosome
	CEP63	SCKL6	614728	3q22.2	centrosome
	NIN	SCKL7	614851	14q22.1	centrosome
	DNA2	SCKL8	615807	10q21.3	nucleus
MOPD I	*RNU4ATAC*	-	210710	2q14.2	component of the spliceosome
MOPD II	*PCNT*	MOPD2	210720	21q22.3	centrosome
	CENPE	-	616051	4q24	kinetochore
Meier-Gorlin Syndrome	*ORC1*	MGORS1	224690	1p32.3	centrosome and nucleus
	ORC4	MGORS2	613800	2q22.3-q23.1	nucleus
	ORC6	MGORS3	613803	16q11.2	kinetochore
	CDT1	MGORS4	613804	16q24.3	kinetochore
	CDC6	MGORS5	613805	17q21.2	nuclear and cytoplasmic
Uncategorized types	*IGF1*	-	608747	12q23.2	nucleus
	LIG4	-	606593	13q33.3	nucleus
	CRIPT	-	615789	2p21	nucleus and cytoplasm
	PLK4	-	616171	4q28.2	centrosome
	XRCC4	-	-	5q14.2	nucleus and cell cytoplasm

3.2.2.1 Plk4 (Polo- like kinase 4)

Polo-like kinases (Plks) are a family of serine-threonine protein kinases. The founding member (Plk1) was first identified in Drosophila where a mutant exhibited defective spindle pole behaviour (Sunkel and Glover 1988). The PLKs family consists of five members: Plk1, Plk2/Snk, Plk3/Fnk/Prk, Plk4/Sak and Plk5 (de Carcer et al. 2011; Sillibourne and Bornens 2010). PLKs has miscellaneous roles and is critical for many cellular processes including regulation of mitosis, chromosome segregation, cytokinesis, spindle pole formation, and cell cycle checkpoints in response to DNA damage (de Carcer et al. 2011). Plk4 has no direct homologue in *C. elegans* but zyg-1 has been proposed as the functional complement due to its essential role in centriole duplication (O'Connell et al. 2001). All members of the Plks family have a special modular architecture that connects a C-terminal structure (Polo box domain) to the N-terminal kinase domain. Plk1-3 contains two Polo box domains at their carboxyl tail domain, while Plk4 has one (Sillibourne and Bornens 2010). However, Plk4 possesses a unique central region called cryptic polo box (CPB) domain (cassette of PB1-PB2) which make Plk4 a distinguished memeber in the Plks family (Fig 3.3). Plks have 3 PEST sequences which are known to have role in kinase stability (Sillibourne and Bornens 2010). The C-terminal Polo box domain, the PB3 domain, exists as homodimer, composed of twisted antiparallel ß sheets and a C-terminal helix (Jana et al. 2012) (Fig 3.3-D). The precise role of PB3 is largely unknown. It was thought that it confers the ability to homodimerize but recent studies demonstrated that PB3 is not necessary for Plk4 dimerization (Klebba et al. 2015a; Leung et al. 2002). Loss of the PB3 domain causes a reduced centriole targetting. In a recent study a novel role was assigned to the PB3 domain where it was shown that autoinhibition of Plk4 is relieved through the PB3 domain. Further, the kinase activity was impaired in a Plk4-ΔPB3 protein (Klebba et al. 2015a). Plk4 dimerizes through the CPB. It brings together a degron motif located near the kinase domain, which then recruits the SCF [Slimb/β-TrCP] ubiquitin ligase, resulting in its ubiquitination and proteasomal degradation (Klebba et al. 2013). Morever, the CPB was shown to interact with the N-terminus of Asl/CEP152 and drive centriole biogenesis but recent studies revealed that the C-terminus of Asl binds to the CBP of Plk4 to initiate centriole duplication while the N-teminus of Asl promotes Plk4 homodimerization (Klebba et al. 2015b).

Loss of Plk4 in mice (Plk4[-/-]) leads to an arrest of the embryo at E7.5 and is embryonically lethal with marked increase in mitotic and apoptotic cells, indicating that Plk4 is essential for embryonic development (Hudson et al. 2001). Plk4[+/-] mice grew normally but gradually developed tumors in adulthood, especially in liver and lung tissues. In addition, Plk4[+/-] MEFs displayed centrosome amplification, reduced cell growth and defective spindle formation (Ko et al. 2005). In support of these findings, another study found multinucleation, centrosome amplification and tetraploidy in early passages of Plk4[-/+] derived mouse embryonic fibroblasts

(MEF) and proposed cytokinesis failure as underlying mechanism (Rosario et al. 2010). On contrary, a recent study showed no centrosome amplification and gross chromosomal abnormalities in Plk4[+/-] MEF cells and further indicated that Plk4 has no direct role in cytokinesis but rather leads to centriole loss and spindle defects (Holland et al. 2012a). Plk4 is a conserved master regulator of centriole biogenesis, and its overexpression leads to centriole amplification and denovo centriole assembly (Avidor-Reiss and Gopalakrishnan 2013; Bettencourt-Dias et al. 2005). Centrosome amplification effects were investigated in the developing mouse brain and serendipitously a microcephaly phenotype was found (Marthiens et al. 2013). However, Plk4 is autoregulated and promotes its own destruction to limit centriole duplication once per cell cycle (Holland et al. 2012b). Depletion of SAK/Plk4 in *Drosophila* causes centriole loss, mitotic abnormalities, and defective spermatides without axonemes structure (Bettencourt-Dias et al. 2005). In our study we found centriole loss, monoploar spindles and delayed mitotic progression in PLK4 deficient cells from patients with microcephalic primordial dwarfism, the first human phenotype associated with PLK4 (Martin et al. 2014).
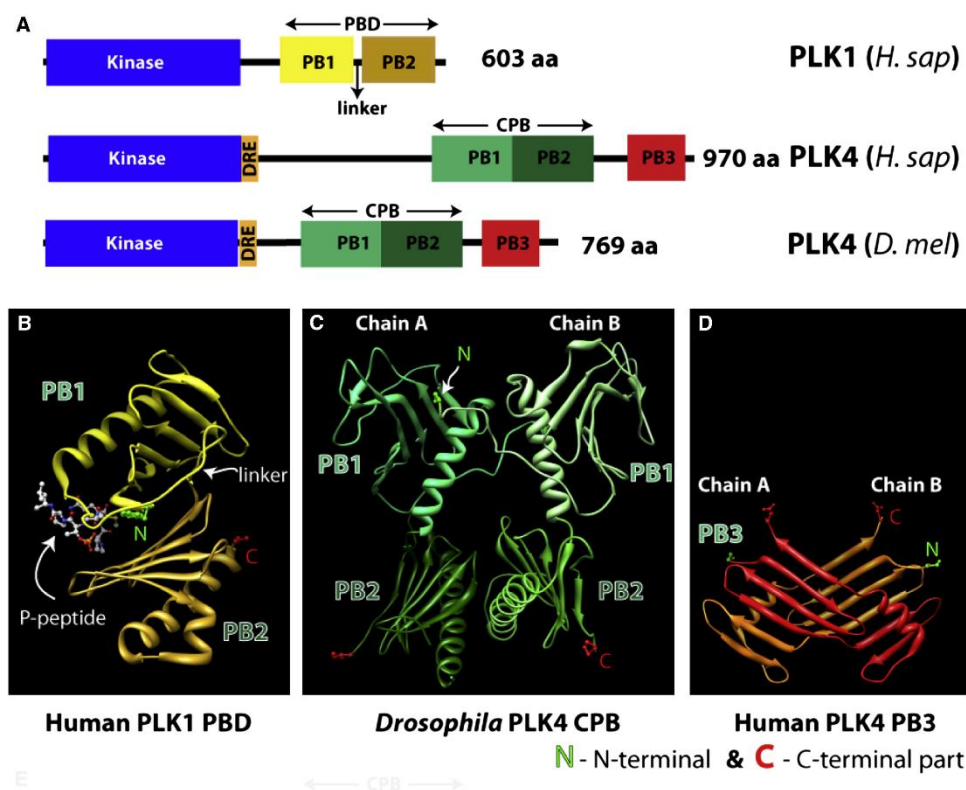
Figure 3.3: Organization and structure of PLKs: **(A)** Domain organization of human and *Drosophila* PLK4 and human PLK1**. (B**) Quaternary structures of the polo box domain in human PLK1. **(C)** Quaternary structures of the cryptic polo box in *Drosophila* PLK4 **(D)** PLK4 domain-swapped PB3 dimer. Adopted from (Jana et al. 2012).

3.3 The centrosome

The vertebrate centrosome is composed of two orthogonally arranged centrioles that are embedded in a pericentriolar material (PCM), which is a matrix that surrounds the centrioles and is enriched in different proteins required for centrosome associated functions. Centrioles are barrel-shaped structures composed of nine triplets of microtubules with dimensions of 0.5 µm length and 0.2 µm in diameter (Fig 3.4). The centrioles also contain highly stable distal and sub distal appendages (Bettencourt-Dias 2013; Doxsey 2001). The PCM is highly dynamic and contains a high proportion of coiled-coiled proteins which provides docking sites for various proteins, particularly members of the γ-tubulin ring complex (γTuRC) (Doxsey 2001). Centrioles are the major players which determine most properties of the centrosome. The centrosome has several functions that act primarily around cell cycle regulation and microtubule organization. Besides microtubule nucleation, it facilitates mitotic spindle assembly, chromosome segregation fidelity, and spindle orientation to the cell cortex in the developing embryo. It lays a patch for the assembly of primary cilia and is involved in signal transduction. In additional, it is also engaged in cell migration, vesicle trafficking, and cell polarity (Bettencourt-Dias 2013; Bornens 2012; Chavali et al. 2014). Centrosomes are not essential in all organisms and there are many exceptions existing in the eukaryotic tree. Centrosomes are absent in many species such as fungi, seed plants, and in many classes of protists that are thought to have lost the ability to encode proteins essential for centrosome assembly (Bettencourt-Dias 2013; Carvalho-Santos et al. 2011).
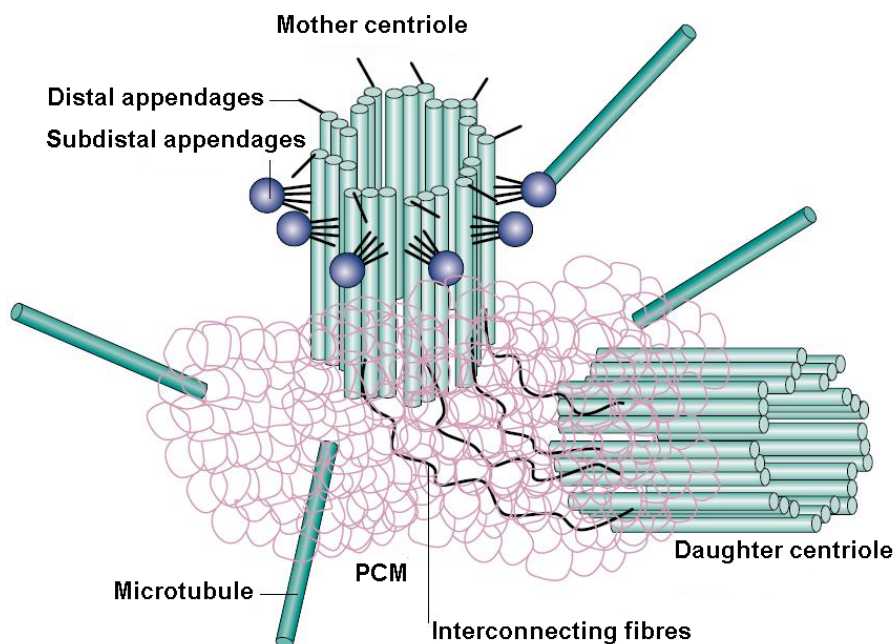
Figure 3.4: Structure of the centrosome. Schematic of the centrosome showing a centrosome consisting of the centrioles and the surrounding PCM. Adopted from (Doxsey 2001).

Despite the lack of centrosomes these organisms have a well-organized microtubule cytoskeleton system. There are some classes of animals like *Planaria* which assemble cilia through centrioles but have no centrosome. In mammals, the female oocyte and the mouse embryo lacks centrioles until the 64 cells stage. Along these limitations of the centrosome, it is quite essential in other organisms for tissue development (Bettencourt-Dias 2013).

3.3.1 Centrosome duplication cycle

The centrosome plays an important role in the cell cycle, having a dominant role in guiding spindle formation. The centrosome number is regulated in a controlled manner during the cell cycle, and division occurs once per cell cycle in a semi-conservative manner (Holland et al. 2010). The centrosome cycle is completed after a sequence of many key events (Fig 3.5). GI phase cells possess one centrosome that contains one mature centrioles called the mother centriole, generated at least two cycles before, and the other is the daughter or immature centriole, generated in the previous cell cycle (Azimzadeh and Bornens 2007).
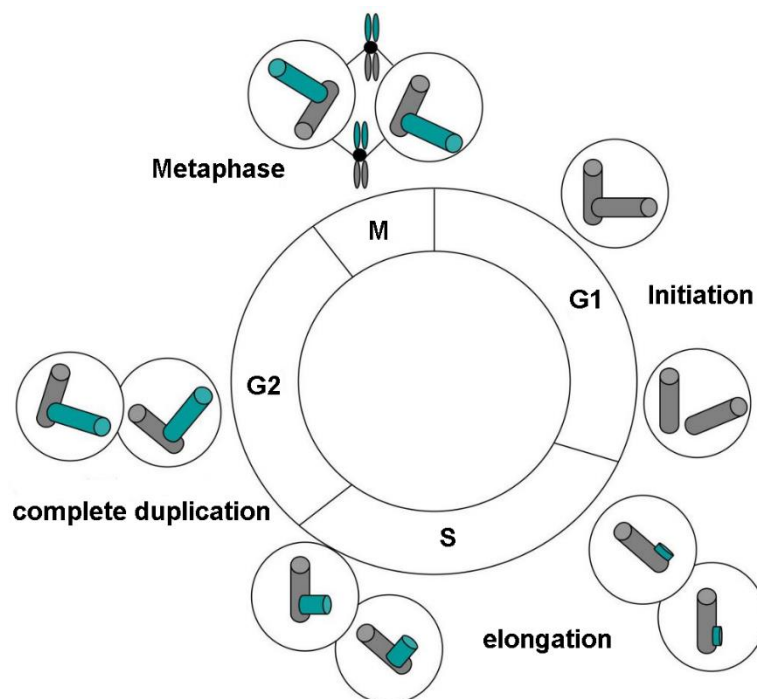
Figure 3.5: Centrosome duplication cycle: Schematic diagram showing the different stages of the centrosome duplication cycle. Mother centrioles are shown in dark gray and daughter centriole in blue-green. This image is taken from (Afonso et al. 2007).

The mother centriole is distinguished by the presence of appendages at the distal end. At the early G1 phase before completion of cytokinesis, the two centrioles disengage and the new procentriole assembles at a perpendicular angle next to each parental centriole in S phase. As the cell progresses through S and G2 phases, procentrioles elongate and attain their final length at G2. At the G2-M phase transition, the two centrosomes separate, start to instruct the formation of spindle poles and gradually move to the opposite poles. At late cell division, the connection between mother and procentriole are removed and disengage both centrioles. The daughter centrioles mature and secure distal and sub distal appendages in the following cycles (Azimzadeh and Bornens 2007; Holland et al. 2010).

3.4 Centrosome and primary microcephaly proteins

Centrosome dysfunction has been linked with many human diseases such as defects in brain development, cancer and ciliopathies. MCPH has been found the bona fide model genetic disorder for centrosome and associated proteins. In general, microcephaly is the prominent feature of many developmental brain disorders. The genes afflicted in MCPH can be classified into three categories according to their function; (i) centriole duplication; (ii) centrosome maturation and spindle pole formation; (iii) other cell cycle associated functions (Table 3.1; Fig 3.6). Five MCPH proteins (CENPJ, STIL, CEP135, CEP152 and HsSAS-6) play important roles in centriole duplication. The formation of new centrioles requires core centrosomal components, including PLK4, STIL/Ana2/SAS-5, CPAP/CENPJ/SAS-4, CEP135/Bld10 and CEP152/Asl (Bettencourt-Dias et al. 2011; Chavali et al. 2014). Interestingly, four of these carry MCPH-linked mutations, and PLK4 is now established as the cause of MPD through the current study (Martin et al. 2014), and by an independent study (Shaheen et al. 2014a). CEP152 acts as scaffold and interacts with CEP192 to recruit polo-like kinase 4 (PLK4) to trigger centriole biogenesis. This is followed by the recruitment of HsSAS-6 (MCPH12), STIL (MCPH7) and then finally CENPJ (MCPH6) protein. CENPJ (MCPH6) is the key button of the MCPH protein network, because it binds STIL, CEP152 and CEP135, albeit most likely not concurrently (Bettencourt-Dias et al. 2011; Chavali et al. 2014).

CDK5RAP2 is essential for centrosome maturation, centriole cohesion, and shackles the centrosome to the mitotic spindle poles. In the absence of CDK5RAP2, centrioles detached from the spindle poles (Barr et al. 2010). The orthologue of ASPM has been shown to be crucial for mouse and Drosophila. On depletion, centrosomes were detached from the spindles (Fish et al. 2006; Gonzalez et al. 1990). WDR62 and CDK6 proteins are also associated with spindle poles during mitosis, and mimic the localization of ASPM, CENPJ and CEP152 have also role in centrosome maturation (Bettencourt-Dias et al. 2011; Hussain et al. 2013).

The remaining four MCPH proteins show no clear functional overlap. MCPH1 is a nuclear and centrosomal protein which is involved in the control of the cell cycle, DNA damage and repair (Cox et al. 2006). CASC5 is a kinetochore scaffold protein and required for the correct attachment of centromeres to the mitotic apparatus, and is necessary for spindle-assembly checkpoint (SAC) signaling (Genin et al. 2012). The ZNF335 protein is localized in the nucleus and plays a role in chromatin remodeling. It binds to the promoter region of transcription factor REST/NRSF, which has a crucial role in neurogenesis and neural differentiation (Yang et al. 2012). PHC1, a member of the polycomb group, localizes to the nucleus and promotes the ubiquitination of histone H2A. It regulates the levels of GEMININ, a protein which is involved in cell cycle regulation and partly localizes to the centrosome (Barbelanne and Tsang 2014; Khan et al. 2014).

3.5 Short commentary on cellular pathways in microcephalic primordial dwarfism (MPD)

The molecular genetics of primary microcephaly and primordial dwarfism has highlighted the phenotypic and genetic overlap. At least ≥ 21 genes have been implicated in microcephalic primordial dwarfism. The genetic aberrations in these genes are not only associated with reduced brain size but also cause substantial growth failure. These genes encode proteins that are involved in cell cycle regulation and cell divisions. Most of the genes mutated in primordial dwarfism activate proteins which are part of the ataxia telangiectasia mutated (ATR) signaling pathway (Chavali et al. 2014; Klingseisen and Jackson 2011). ATR and its main substrate CHK1 coordinate the response to DNA damage and replicative stress. ATR is essential from S to M phase progression or for stabilizing of stalled replication forks (Alderton et al. 2004; Chavali et al. 2014). Proteins (ORC1, ORC4, ORC6, CDT1 and CDC6) associated with the replicative complex formation, and origin activation showed alterations in cell-cycle kinetics through replicative stress in Meier-Gorlin syndrome (Bicknell et al. 2011; Bleichert et al. 2013). The accumulation of replicative stress disrupts development and tissue homoeostasis. Mutations in *ATR, ATRRIP, CtIP* genes associated with Seckel syndrome lead to deficient ATR signaling pathways (Ogi et al. 2012; Qvist et al. 2011). Similarly, core centrosomal protein PCNT mutant cells exhibit impaired ATR signaling with defective G2/M checkpoint (Tibelius et al. 2009). Interestingly, the MCPH gene *CEP152*, also appears in the maintenance of genomic integrity and caused Seckel syndrome. In CEP152 deficient fibroblasts CHK1 function is normal and activation of the Ataxia telengiectasia-mutated (ATM) singaling pathway occurs (Kalay et al. 2011). The accumulation of replicative stress and the premature checkpoint arrest in ATM or ATR signaling pathways create antiproliferative effects that impede cellular growth during development (Chavali et al. 2014; Klingseisen and Jackson 2011). Centrosome amplification and aneuploidy is another cause that can affect the cellular growth as

demonstrated in the patient cells that were deficient in PCNT, CEP152, ORC1 and CPAP proteins (Alderton et al. 2004; Chavali et al. 2014; Hossain and Stillman 2012; Kalay et al. 2011). Cells with excess centrosomes should cluster them in two poles (Quintyne et al. 2005). Otherwise cells with centrosome amplification undergo multipolar cytokinesis and cell death. In spite of this crucial role, centrosome clustering leads to a delay in mitosis and genomic instability as observed in CPAP-Seckel mouse cells and CEP152-Seckel lymphocytes (Kalay et al. 2011; McIntyre et al. 2012). It seems that genes associated with microecphalic primordial dwarfism prevent centrosomal duplication. This notion is further supported by the finding that ORCI suppresses CDK2-cyclin E dependent centriole duplication (Hossain and Stillman 2012).

Lastly, recent studies indicated that cilia dysfunction also contributes to the pathomechanism of MPD. The depletion of ORC1 in wild type cells, and ORC1-deficient patient primary fibroblasts leads to defective cilia formation and impaired sonic hedgehog signaling (Shh). The Shh pathway appears to have an important role in embryogenesis (Stiff et al. 2013).
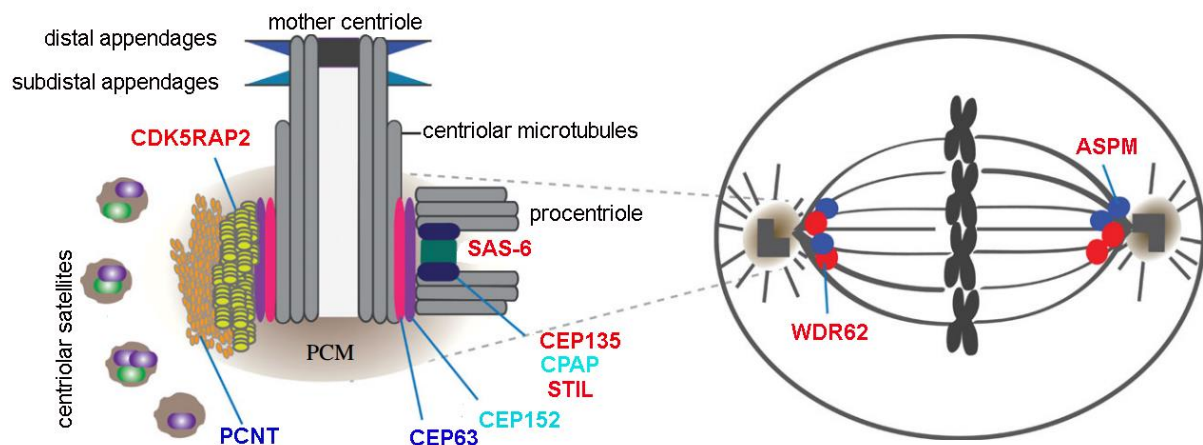
Figure 3.6: Schematic of the vertebrate centrosome and the mitotic spindle. The location of MCPH and MPD associated proteins are shown. MCPH protein are listed in red, MPD proteins in blue and proteins involved in both disorders are shown in bluish color. This image is modified from (Chavali et al. 2014).

3.6 Corticogenesis and primary microcephaly

Humans have more than 100 billion nerve cells in their brain (Stiles and Jernigan 2010). Scientists have now gained major insights to understand the critical components of normal brain development and functions. Primary microcephaly has been identified as the disease of neurogenic mitosis which results in a reduction of the cell number in the developing human brain. A unifying theme stemming from the molecular genetic studies is the role of the centrosome and its association with the neural progenitors (Cox et al. 2006; Hu et al. 2014). The association of MCPH proteins with the centrosome and cell cycle was already discussed

under the heading 3.4. In the next section, current concepts concerning primary microcephaly are reviewed.

A. Symmetric and asymmetric cell division

In human, the neural tube is formed during the third week of gestation. The neural progenitor cells line the center of the neural tube. When the brain becomes larger, the lumen and neural progenitor cells form the ventricular zone (VZ). At the end of gastrulation, neural progenitor cells undergo symmetrical division and two identical neural progenitor cells at each division are generated and hence they exponentially enlarge the progenitor populations (Stiles and Jernigan 2010). At E42, the mode of cell division switches from symmetrical to asymmetrical, generating one progenitor and a neuron that leaves the ventricular zone and divides no further. This results in increasing numbers of neurons and cortical thickness. The new progenitor cells remain in the proliferative zone and continue to divide. In human, cortical neurogenesis is completed by approximately at E108 (Clancy et al. 2001; Stiles and Jernigan 2010; Wodarz and Huttner 2003).

The differentiated neurons migrate radially from the proliferative zones to the developing neocortex. In earliest migration neurons adopt a somal translocation mode but as development proceeds, the neurons use radial glial cells for migration. The migration ultimately leads to the formation of the 6-layered cortex. This occurs with the earlier neurons constituting the deepest layer and the following neurons forming the outer layers in an inside-out fashion (Stiles and Jernigan 2010).The balance between progenitor cell expansion and differentiation is essential for the precise numbers of neurons and brain size (Chenn and Walsh 2002; Noctor et al. 2004). Spindle-pole positioning relative to the apical–basal axis of neuroepithelial cells determines the overall orientation of cleavage. Vertical cleavage leads to two identical progenitor daughter cells while the outcome of horizontal cleavage are an apical and a basal daughter cell (Gotz and Huttner 2005).

Disturbance of these divisions affect the neural progenitor pool. Genes that are mutated in MCPH encode proteins that regulate several key functions and are expressed in embryonic neurogenesis. Most of these proteins are localized at the centrosome and the spindle poles (Fig 3.6). So it is speculated that these MCPH proteins influence the orientation of the mitotic spindle and the mode of cell division in stem cell division (Fig 3.7). As a consequence of the premature shift from symmetric to asymmetric cell divisions, the progenitor cell pool is depleted and thus a reduced number of neurons results in the congenital condition of primary microcephaly (Cox et al. 2006).
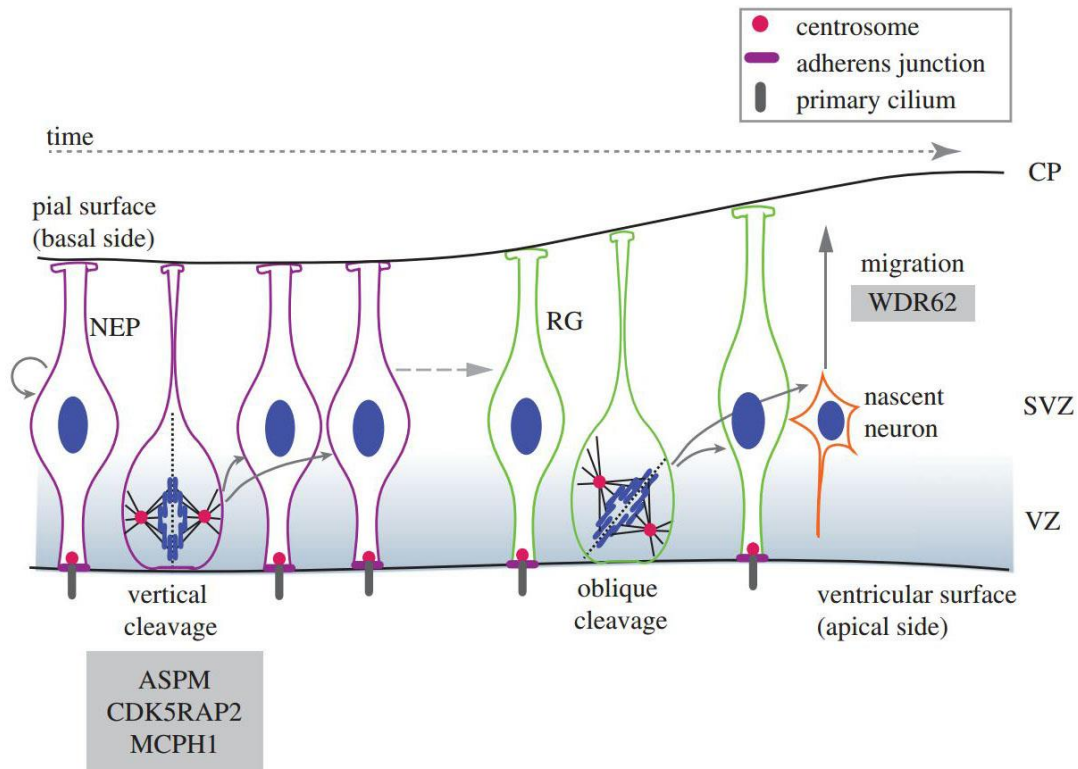
Figure 3.7: Schematic representation for the neurogenesis in the rodent brain. The neural progenitor cells (NEP) expand through symmetric division to exponentially enlarge the progenitor population but after the onset of neurogenesis, NEPs give rise to the radial glial (RG) cells which divide by symmetric division and asymmetric division. They are lying in the ventricular zone. The progenitor cells display a vertical cleavage plane to the apical surface in case of symmetric division while in case of asymmetric division, they display oblique or horizontal cleavage planes. This image is taken from (Chavali et al. 2014).

B. Apoptosis

In addition to the cell fate choice, cell death and premature cell cycle exit during neurogenesis has been proposed to contribute to the pathomechanism of MCPH. Microcephalin mediated DNA-damage response through ATM or ATR signaling pathways caused reduced neuron progenitor proliferation by inducing apoptosis. In recent studies, centrosome defects and errors in chromosome segregation are observed in microcephaly mouse models which lead to aneuploidy, mitotic arrest, and ultimately apoptosis (Gruber et al. 2011; Lizarraga et al. 2010b). The overexpression of PLK4 in the mouse brain did not significantly disturb spindle orientation but rather promoted p53 dependent apoptosis, seemingly as the result of aneuploidy (Marthiens et al. 2013).

4. Aims and objectives of the thesis

The main focus of my PhD thesis is to identify genetic variants that cause two types of inherited neurodevelopmental disorders: primary microcephaly, a typically lethal condition characterized by small brains, and microcephalic primordial dwarfism (MPD), i.e. when primary microcephaly is combined with a global growth failure.

Individuals suffering from MPD are usually described as the smallest people in the world. Mostly, the genetic components of both disorders are involved in the regulation of the cell cycle or are part of the core cellular machinery. The highly consanguineous Pakistani population is a goldmine for genetic studies to identify the defects underlying the diseases. The identification of new genes involved will improve diagnostics and management of these rare genetic disorders. It can increase our understanding of how the body regulates brain and cellular growth during embryonic development. The specific aims of this thesis are:

1. To analyze Pakistani consanguineous families segregating primary microcephaly and microcephalic primordial dwarfism.

2. To screen for mutations in previously reported genes to broaden our knowledge on the mutational spectra

3. To identify causative variants in novel genes associated with MCPH and MPD

4. To understand the basic cellular mechanisms and their role in the development of pathological phenotypes

4. To uncover the functional consequences of the identified causative variants and the role of the newly associated genes in the etiology of the corresponding disease.

5. Materials and Methods

5.1 Families studied

For the study presented here, twenty nine consanguineous families of autosomal recessive primary microcephaly (MCPH) were recruited from the Khyber Pakhtunkhwa, Sindh and Punjab provinces of Pakistan. Among these families, there were 113 MCPH affected members with a male to female ratio of 63:50 respectively. This cohort had families from different ethnicities and geographical origin of Pakistan. The number of affected members in each family varied from 2 to 13. In addition, a family segregating inherited microcephalic primordial dwarfism with seven affected individuals was ascertained from Khyber Pakhtunkhwa, province of Pakistan. Informed consent was obtained from families who agreed to participate in this study. All families demonstrated the autosomal recessive mode of inheritance. Clinical examination was performed meticulously with careful measurement of head circumference, a simple method used to asses gross brain size (Woods et al. 2005). All affected individual displayed reduced head circumference (-5 SD to -13 SD) whereas their parents were normal. There was no indication of growth retardation or other neurological deficits except seizure was reported in some patients which are no longer excluded from the MCPH diagnostic criteria (Woods et al. 2005). The degree of intellectual disability varied from mild to severe or even profound and it differed among patients within the families. In most MCPH patients communication barrier was obvious while some patients had the ability to speak short sentences and basic knowledge of native language. Cognitive deficits were evident in most cases while some patients had delayed motor skills. The phenotypes of all twenty nine congenital MCPH families are documented by photos (Fig 5.1).

The patients in MCP68 were diagnosed as having microcephalic primordial dwarfism in accordance with the classification of primary microcephaly-seckel spectrum disorders (Verloes et al. 1993). The general criteria set for such syndromes are the prenatal onset of proportionate dwarfism, marked microcephaly, intellectual disability and autosomal recessive segregation.

5.2 Pedigree construction and analysis

To make genetic conjecture, extensive pedigrees were drawn for each family using standard methods. The inheritance pattern in the families was inferred by observing the mode of segregation. The genealogical relationships for all affected individuals were established through personal interviews of the elderly of the family. The males were symbolized by squares and females by circles. The phenotypically normal individuals were indicated with unfilled symbols while the affected individuals were depicated by filled symbols. A number enclosed within a symbol indicates the number of sibs, male or females, as the case may be. Deceased

individuals were presented as line across the symbols. A horizontal line connecting a circle and square is a marriage line while a double horizontal line shows consanguinity. An arrow pointing at a particular affected individual is the starting point for construction of the pedigree (general proband, propositus for male and proposita for female). Pedigrees were constructed using HaploPainter v.1.043 (Thiele and Nürnberg 2005). All the family data were arranged in the form of pedigrees, the mode of inheritance of all studied families was ascertained as autosomal recessive.

Figure 5.1: Clinical features of affected individuals of primary microcephaly families.

5.3 Blood sampling

Peripheral blood samples were taken from all available affected individuals, parents and normal members of the family with informed consent. Blood was collected in 10 ml BD vacutainer® (BD-Plymouth, UK) containing EDTA as anticoagulant. These blood samples were stored at 4 °C for short term or at -80 °C for longer periods until DNA extraction.

5.4 DNA extraction

Genomic DNA was extracted with phenol- chloroform isoamyl alcohol, a common protocol for isolation of DNA from blood. 0.75 ml of blood was taken in a 1.5 ml sterile microcentrifuge tube with equal volume of 0.75 ml of solution A and kept at room temperature for 10-15 minutes after mixing the contents. The tubes were then centrifuged for 1 minute at 13,000 rpm and discarded the supernatant. The pellet was resuspended in 400 µl of solution A and repeated centrifugation and discarding the supernatant. Nuclear pellet was resuspended in 400 µl of solution B, 25 µl of 10% SDS and 8 µl of proteinase K and incubated at 65 °C for three hours in water bath or at 37 °C for overnight. After this step, 0.5 ml of freshly prepared solution C and D was added into samples, mixed and centrifuged for 10 minutes at 13,000 rpm. The top (aqueous) phase was collected carefully in a new 1.5 ml microcentrifuge tube and added equal volume of solution D. Centrifugation was then performed at 13,000 rpm for 10 minutes. The aqueous phase was transferred to a new sterile tube and added 55 µl of 3 M sodium acetate (pH 6) and equal volume of isopropanol. Tubes were then inverted several times to precipitate DNA. DNA was incubated with 350 µl of 70% ethanol (stored at -20 °C) and centrifuged at 13,000 rpm for 10 minutes. Ethanol was drawing off from side without disturbing DNA pellet. DNA was dried and dissolved in appropriate amount of buffer FG3 (Qiagen) or ultrapure water.

Composition of solutions

Solution A (pH 7.5)

0.32 M Sucrose	54.72 g
10 mM Tris-HCl (pH 7.5)	5 ml (1 M)
5 mM MgCl$_2$	2.5 ml (1 M)

First, the aforementioned chemicals were dissolved in 400 ml of distilled water and adjusted to pH 7.5. The volume was made up to 495 ml with distilled water and autoclaved. After autoclaving, the bottle was allowed to cool at room temperature. The volume of solution was raised to 500 ml (final volume) by adding 5 ml of Triton X-100 (1% v/v) and was autoclaved before use.

Solution B

10 mM Tris-HCl (pH 7.5)	5 ml
400 mM NaCl	40 ml
2 mM EDTA (pH 8.0)	2 ml

The solution was brought to 500 ml (final volume) and autoclaved.

Solution C (saturated phenol pH 8.0)

100 ml Phenol (Sigma-Aldrich) was heated to 68 °C in a water bath. Then it was mixed with 100 ml of 0.5 M Tris-HCl (pH 8.0) and 100 mg of hydroxyquiniline (0.1%). The mixture was stirred for 15-20 minutes on a magnetic stirrer. The aqueous upper phase was discarded and added 0.1 M Tris-HCl (pH 8.0) of equal volume to the lower phase. The solution was stirred for 15 minutes. The upper phase was removed again and repeated it several times until the pH of phenol reached to 7.8. The pH of the phenol was measured within the upper phase. Next, 10 ml of Tris-HCl was added that had 200 µl of β-mercaptoethanol. The solution was stored for 1 month at 4 °C.

Solution D

Chloroform   24 volume

Isoamy alcohol   01 volume

5.4.1 DNA quantification

DNA integrity was assessed through 1% agarose gel and quantified through NanoDrop TM 1000 Spectrophotometer (Wilmington, USA) for further experiments. 1.5 µl of well mixed DNA dilutions was directly placed onto the ND-1000 Spectrophotometer. For each sample, DNA light absorbance at 260 nm and 280 nm was measured three times. The DNA purity and concentration were calculated directly by the provided software.

5.4.2 Polymerase chain reaction (PCR) and gel electrophoresis

Polymerase chain reactions were run in different amounts and on different PCR profiles according to the downstream application. The primers are listed in appendix 1-5 and appendex 7. Mastermixes was run on DNA Engine® Tetrad 2 Thermal Cycler (Bio-Rad) and GeneAmp® PCR System 9700 Gene (AB Applied Biosysytem). Different Taq polymerases were used to amplify targets of various lengths. Innu Taq DNA polymerase (Analtika Jena) was used to amplify products for sequencing and genotyping with microsatellite markers. The Q5 High-Fidelity DNA Polymerase (New England Biolabs GmbH) was used for amplicons in break point determination. Low melting point agarose (Invitrogen ™) was used for preparative gels (0.7 to 2%). The agarose was melted in IXTAE buffer (40 mM Tris (pH 7.6), 20 mM acetic acid and 1 mM EDTA), stained with ethidium bromide, and run the gel on horizontal gel electrophoresis chambers (Analytik Jena) to resovle the DNA fragments. The DNA molecular weight markers, 1 kb Plus DNA Ladder (Invitrogen) and 100 bp DNA Ladder (Fermentas) were used for size determination of the amplicons. Ethidium bromide (0.5 µg/ml final concentration) was used to

visulized nucleic acids under UV light. Images of the gels were taken by Molecular Imager® Gel Doc™ XR System and Alpha Innotech Multi Image Light Cabinet (Biozym).

5.5 Exclusion mapping using short tandem repeats (STRs)

In order to locate the disease gene in six (MCP101, MCP104, MCP105, MCP106, MCP117 and MCP120) out of twenty nine MCPH families, a set of fluorescently labelled STR markers of the most common MCPH loci were used to type the DNA of all available individuals in a pedigree. Four to six microsatellite markers were selected for each of the loci MCPH1 to MCPH7. Segregation of haplotypes was analyzed in healthy and affected individuals to see if the families are linked to or excluded from the known loci. Appendix 1 summarizes microsatellite markers located within the region of known loci. These microsatellite markers were selected for each locus from the Marshfield genetic map (Broman et al. 1998), and were labeled with fluorescent dyes e.g. FAM, TET, HEX and NED at 5′ end of forward primers. Haplotypes were constructed (not shown) and analyzed for linkage or exclusion to known loci. Polymerase chain reaction (PCR) amplification was carried out in a total volume of 10 µl containing 6 ng of genomic DNA, 10x PCR buffer II with 15 mM ($MgCl_2$), 10 mM of deoxyribonucleotide triphosphates (dNTPs), 2.5 U of Innu polymerase, forward and reverse primers each of 10 mM. Two different amplification conditions (MS and TDM) were used as given in the table 5.1. After amplification, the PCR products were diluted and added 5 µl to the 7 µl of internal size standard (1:100 diluted) (GeneScan-500 ROX; Applied Biosystems). The amplified markers were then separated by capillary electrophoresis on ABI 3730 genetic analyzer (Applied Biosystem, Foster City, CA). Allele sizes of microsatellite markers (bp) were determined using GeneMarker version 1.51 software (SoftGenetics) by comparison with those of the internal size standard. Haplotypes were constructed by EasyLinkage Plus v5.02 and HaploPainter v.1.043 (Thiele and Nürnberg 2005). If all affected individuals in a pedigree demonstrated homozygosity and unaffected parents heterozygosity for a marker set spanning a particular region, the family was considered as linked to that particular locus. Heterozygosity of the markers in affected individuals excluded the locus under investigation.

Table 5.1: PCR profiles to amplify microsatellite markers for known MCPH loci

MS			TDM		
1	95 °C	5 min.	**1**	95 °C	5 min.
2	95 °C	30 Sec.	**2**	95 °C	30 Sec.
3	61 °C	45 Sec.	**3**	65 °C	30 Sec.
4	72 °C	1 min.	Decrease by 0.5°C every cycle		
Step 2-4 repeat 2 more times			**4**	72 °C	30 Sec.
5	95 °C	30 Sec.	Cycle to step 2 for 19 more times		
6	59 °C	45 Sec.	**5**	95 °C	30 Sec.
7	72 °C	1 min.	**6**	55 °C	30 Sec.
Step 5-7 repeat 2 more times			**7**	72 °C	30 Sec.
8	94 °C	30 Sec.	Cycl to step 7 for 19 more times		
9	57 °C	45 Sec.	**8**	72 °C	10 min.
10	72 °C	1 min.	**9**	4 °C	Forever
Step 8-10 repeat 2 more times					
11	95 °C	30 Sec.			
12	55 °C	45 Sec.			
13	72 °C	1 min.			
Step 14-16 repeat 2 more times					
14	72 °C	30 min.			
15	4 °C	Forever			

5.6 SNP genotyping methods

The initial draft of human genome in 2001 and follow-on studies spawned a new era of functional characterization and cataloguing of genetic variations. Several million SNPs were discovered and characterized in terms of linkage disequilibrium (LD) in different populations (Naidoo et al. 2011). Genetic variants that are in LD tend to be inherited together. The first linkage map as constructed with 2,988 SNPs and used as a linkage analysis tool to investigate the underlying genetic causes of human diseases (Matise et al. 2003). In addition to linkage mapping, the genome-wide SNP analysis offers the possibility of direct dectection of structural genetic variations (Gibbs and Singleton 2006). New genotyping technologies were introduced over time, starting from 10K arrays (Affymetrix;Santa Clara, CA, in 2003) to the latest 5 million marker tool (Illumina; San Diego, CA). In the present studies, the following arrays were used for homozygosity mapping and structural variation identification.

5.6.1 Axiom™ Genome-Wide CEU 1 Array

The Axiom™ Genome-Wide CEU 1 Array was used for the geonotyping of nine MCPH families including MCP105, MCP106, MCP118, MCP120, MCP121, MCP125, MCP129, MCP130 and MCP131. Genotyping was performed by our in-house core facility at the Cologne Center for Genomics (CCG), Cologne, Germany. Axiom Genome-Wide CEU 1 Array interrogates more than 587,000 SNPs and 11,000 single base insertions/deletions. The SNP content and the insertions/deletions on this array were selected from the dbSNP database and the Axiom Genomic Database. Arrays were run according to standard Affymetrix protocols (part number 901608, www.affymetrix.com). Genotyping Console Software version 4.2 (Affymetrix) was used for allele calling and quality control of called genotypes. Genotyping data were exported in PLINK format (Anderson et al. 2010; Purcell et al. 2007) and primary data were analysed with the PLINK software package (version 1.07; http://pngu.mgh.harvard.edu/purcell/plink/). Data files were recoded in binary file format. Sex of sample was determined from the genotypes on X chromosome and checked against recorded gender information in the ped file generated from the pedigree. Samples with call rates above 96% were considered for further analysis.

5.6.2 Illumina HumanCoreExome-12 v1.1 BeadChip

Four MCPH families MCP104, MCP123, MCP165, MCP171 and a family with microcephalic primordial dwarfism were genotyped at the CCG using the HumanCoreExome 12v1-1 bead chip (Illumina Inc., San Diego, CA) according to the manufacturer's protocol. The Illumina HumanExome-12 V1-1 BeadChip (microarray) targets >240,000 variants, the majority of which are protein-altering variants (i.e. SNPs in splice site, stop variants, non-synonymous SNPs) selected from ~12,000 individual exome and whole-genome sequenced genomes representing multiple ethnicities and range of common conditions (complex traits). Genotype calling was carried out with the GenomeStudio software version 2011.1 (Illumina, Inc. San Diego, USA). A standard protocol implemented in PLINK was used to perform quality control (QC) and data clean-up of genotype data (Anderson et al. 2010; Purcell et al. 2007). SNPs that had call rates of less than 97% were excluded from further analysis.

5.6.3 CytoScan HD Array

The Genome-wide Human CytoScan HD Array (Affymetrix, CA, USA) was used for DNA samples from MCP118 and MCP125 to analyse genomic alterations. Experiments were performed and analysed at ATLAS Biolabs GmbH, Berlin according to the manufacturer's protocol. The array contains 2,690 K markers that include 750K polymorphic SNPs and 1,900 K non-polymorphic copy number variation (CNV) markers. These probes are spaced at 880

bp on average and covered 96% of all genes. Samples with QC call above 80% were considered for further analysis. For CNV detection, an algorithm of the Chromosome Analysis Suite (ChAS) v2.0 was used (Affymetrix Inc., USA).

5.7 SNP array mapping and linkage analysis

Thirteen consanguineous MCPH families (MCP104, MCP105, MCP106, MCP118, MCP120, MCP121, MCP123, MCP125, MCP129, MCP130, MCP131, MCP165 and MCP171) and a family (MCP68) with MPD were subjected to genome mapping and linkage analysis with the arrays mentioned above. Sample genders were checked by counting the heterozygous SNPs on the X chromosome and comparing it to the given information in pedigree file (PED file). Misspecification or errors of the assumed relationships in the pedigree (e.g sib-sib, parent-offspring etc.) were evaluated with the help of the program Graphical Relationship Representation (GRR). Linkage analysis was performed assuming autosomal recessive inheritance, full penetrance, consanguinity and a disease gene frequency of 0.0001. Mendelian inconsistencies were detected by the program PedCheck. Multipoint LOD scores were calculated using the program ALLEGRO (Gudbjartsson et al. 2000). Haplotypes were reconstructed with ALLEGRO and presented graphically with HaploPainter (Thiele and Nürnberg 2005).

5.8 Sequencing reaction

5.8.1 Sanger sequencing

Sanger sequencing is a linear amplification processes in which the amplified DNA act as template molecule. In the sequencing reaction one primer either forward or reverse is used which anneals to specific region of double-stranded DNA and start the enzymatic extension. In addition to unlabeled dNTPs, the reaction also contains fluorescence labelled dideoxynulceotides (ddNTPs) that terminate the elongation of the DNA reaction due to absence of 3'-hydroxy group. In result, DNA fragments are produced of different sizes which can be subsequently separated by capillary electrophoresis and detected after excitation by a laser. For this reason, Sanger sequencing is also called dideoxynucleotide chain termination. Sequences were viewed with the Seqman program of DNASTAR Software (DNASTAR, Inc. Madison, USA).

PCR was performed for all coding exons of the respective MCPH candidate gene. The 15 µl master mix prepared for the PCR reaction included 3 µl DNA (2 ng/µl), 9.48 µl ddH2O, 1.50 µl 10xPCR-Buffer II (10x), 0.30 µl of each primer (10 mM), 0.30 µl dNTP´s (10 mM) and 0.12 µl of Taq (5 U/µl). DNA was amplified using standard conditions: (1) 95 °C for 5 min; (2) 95 °C

30 sec ; (3) 60 °C for 30 sec ; (4) 72 °C for 30 sec (5) repeat 2-4 step 44 more times; (6) and 72 °C for 5 min. In addition, TDM PCR profile (Table 5.1) was also used.

Cleaning of PCR products with Exo-SAP: PCR products were diluted and cleaned up through Exo-SAP reaction. Two enzymes, Exonuclease I (20 U/µl, New England Biolabs) and the Shrimp Alkaline Phosphatase (SAP) (1U/µl, New England Biolabs) were used to remove the residual nucleotides from PCR products. The exonuclease I removes unused primers, while the Shrimp Alkaline Phosphatase removes of the residual dNTPs. For a total reaction mixture of 10 µl, 8 µl of the PCR product was added to the 0.075 µl Exonuclease I (20 U/µl), 0.3 µl SAP (1U/ µl) and 1.6 µl water. The PCR was performed at 37 °C for 30 minutes, and then 85 °C for 15 minutes.

BigDye® terminator cycle sequencing: Sequencing reactions (final volume of 10 µl) contained 2 µl clean up amplified product, 0.25 µl ABI Big Dye version1.1 (Applied Biosystems GmbH, Germany), 1.75 µl 5 X sequencing buffer (Applied Biosystems GmbH, Germany), 0.25 µl primer (10 µM), and 5.75 µl water. The sequencing reaction was performed on a DNA Engine® Tetrad 2 Thermal Cycler (Bio-Rad Laboratories, Inc., Germany). The thermal cycler conditions were: (i) initial step of 96 °C for 10 seconds; (ii) 55 °C for 5 seconds; (iii) 60 °C for 4 minutes; (iv) repeat i-iii steps 32 times. Sequencing products were purified with CleanSEQ buffer (Agencourt) and subsequently run on ABI 3730 genetic analyzer (Applied Biosystem), sequencing facility at CCG. The sequencing data were analysed with DNA Star (Lasergene) and Mutation Surveyor (SoftGenetics).
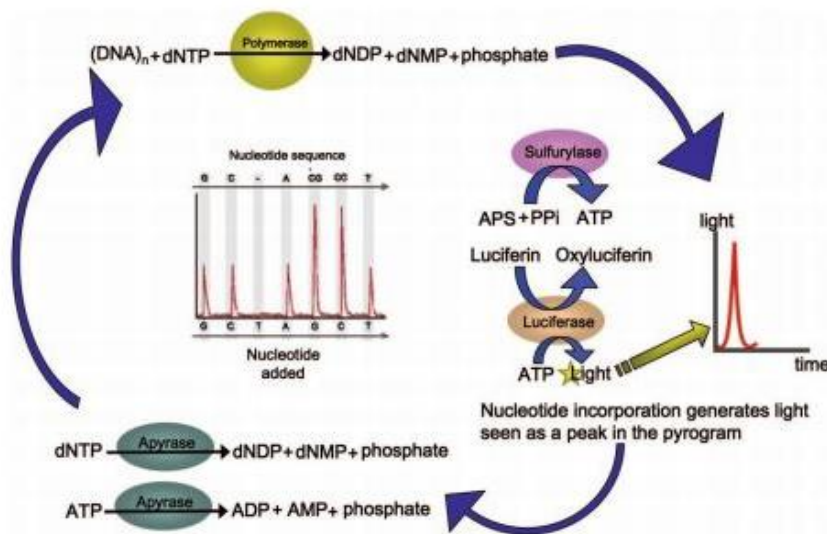
5.8.2 Pyrosequencing

Pyrosequencing, also known as sequencing by synthesis, is the unique approach to DNA sequencing (Ronaghi et al. 1996). It is based on the detection of pyrophosphate (PPi) release during DNA synthesis. The pyrophosphate combines with adenosine phosphosulfate and converts into ATP by enzymatic action of ATP-sulfurlyase. ATP provides energy for Luciferase to oxidize luciferin and generate a flash of chemiluminesence (Fig. 5.2). The emitted light can be recorded by an LCC camera as a peak which is called pyrogram. The light intensity is proportional to the amount of nucleotides added. The excess nucleotides are degraded by the presence of apyrase. Pyrosequencing allows the analysis of genetic variation, genotyping and microorganism identification etc (Petrosino et al. 2009; Ronaghi 2001). It offers to estimate SNP allele frequencies in pooled sample population. To exclude the polymorphic nature of pathogenic variant identified in a family segregating microcephalic primordial dwarfism, allelic quantification in a cohort of 286 samples was carried out by pyrosequencing.

Primer design for pyrosequencing: The Pyromark Assay Design (PSQ) Software v.1.0.6 (Qiagen, Hilden, Germany) was used to design variant-specific PCR and pyrosequencing

oligonucleotide primers. The forward primer was labelled with Biotin at the 5′ site (Appendix 7).

Preparation of PCR products for pyrosequencing: After optimizing the PCR amplification conditions of the target DNA,10 µM of sequencing primer (0.36 µl for one reaction) was added to 1x annealing buffer (11.6 µl for one reaction). High-purity water was added to the PCR samples to make the final volume of 40 µl and then mixed with equal volume of mixture of 3 µl streptavidin Sepharose beads and 37 µl of binding buffer. This mixtures was shaken for 6-10 minutes at room temperature to allow biotinylated PCR product capture on streptavidin-coated beads. Next, 80 µl mixture of PCR product and streptavidin sepharose was added to a 12 µl solution (sequencing primer +annealing buffer) with the help of the PyroMark Q96 Vacuum Prep Workstation. Hereafter mixture was incubated at 85 °C for 2 minutes to denature the template and cooled at room temperature. In last, parameters were selected according to manufacturer's instructions on a PSQ HS96A instrument (Qiagen, Hilden, Germany) and run the machine. Pyrosequencing was carried out using PyroMark Gold Q96 Reagents (Qiagen, Hilden, Germany).The pyrosequencing data were analysed with the Pyro Q-CpG V.1.0.9 analysis software (Qiagen, Hilden, Germany).

Figure 5.2: Schematic representation of the pyrogram and biochemical reactions involved in the generation of the light during pyrosequencing. Adopted from (Petrosino et al. 2009).

Solutions used for pyrosequencing

Streptavidin sepharose (Amersham)

Binding Buffer-pH7.6 (10 mM Tris-HCL; 2 M NaCl; 1 mM EDTA; 0.1% Tween™ 20)

Denaturation solution-0.2 M NaOH

Washing buffer-pH 7.6 (10 mM Tris-Acetate)

1X Annealing buffer (1XAB)-pH 7.6 (20 mM Tris-Acetate and 2 mM Mg-Acetate)

Enzyme-DNA Polymerase, ATP Sulfurylase, Luciferase and Apyrase

Substrate- Adenosine 5' phosphosulfate (APS) and Luciferin

5.8.3 Whole-exome sequencing

The latest scientific advancements in next generation sequencing (NGS) technologies have transformed genetic studies (Mardis 2008). NGS may involve the enrichment of certain genomic areas, either a linkage interval, or all transcribed sequences within the interval, or, more recently, the entire exome (protein-coding genes). The human exome is 1% of the total human genome with roughly 30 million bp (Ng et al. 2009). Exome sequencing is based on capturing by a probe hybridization method which uses the entire set of exome as a sequencing target (Gnirke et al. 2009). This methodology (Fig 5.3) has been successfully used at CCG to identify mutations in rare recessive families. In the present study, whole-exome sequencing was performed in 10 MCPH families (MCP101, MCP104, MCP105, MCP118, MCP121, MCP123, MCP125, MCP131, MCP165 and MCP171) and a family with microcephalic primordial dwarfism (MCP68). DNA samples (1 µg) of one patient per family with unknown cause of disease was used for sequence capture. The established protocols were employed starting with the generation of the sequencing library before the enrichment step. Fragments were enriched using the SeqCap EZ Human Exome Library v2.0 kit (Roche NimbleGen) and subsequently, sequencing was performed on an Illumina HiSeq 2000 using a paired-end 2 × 100-bp protocol. For the processing and analysis of the data, the inhouse pipeline Varbank (http://varbank.ccg.uni-koeln.de) was used at CCG. The sequence variants were prioritized for their likelihood to cause the disease phenotype. Variants of the highest priority were further validated by Sanger sequencing and tested for co-segregation in all family members.
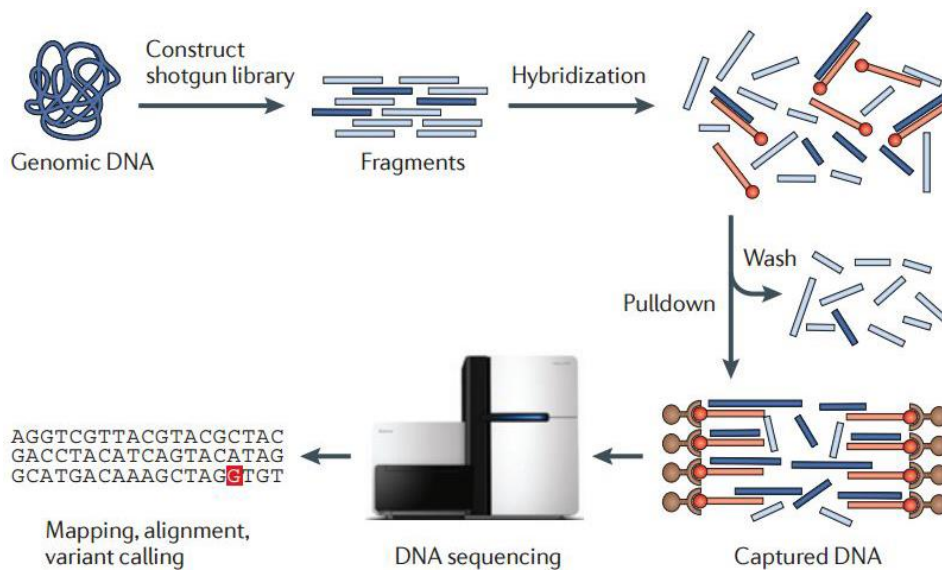
Figure 5.3: Schematic of workflow for exome sequencing. This image is taken from (Bamshad et al. 2011).

5.9 RNA extraction

Peripheral blood was collected in PAXgeneTM Blood RNA tubes to isolate total cellular RNA from whole blood. RNA was extracted by using PAXgene Blood RNA kit (PreAnalytiX) according to the manufacturer's protocol (version, 2001). 1.5 µl of well mixed RNA sample was loaded onto Nanodrop ND-1000 spectrophotometer (ThermoFisher Scientific) and absorbance measured at 260 nm. RNA samples were diluted according to the downstream applications.

5.9.1 Reverse transcriptase PCR (RT-PCR)

To synthesized cDNA from RNA, Superscript II reverse transcriptase enzyme (Invitrogen) was used. In brief, 1 µg of total RNA was mixed with 1 µl dNTPs (10 mM), 1 µl Oligo d(T)18-Primers (50 µM) and filled up to 13 µl with nuclease free water in a microcentrifuge tube. This mixture was incubated them for 5 minutes at 65 °C to remove secondary structures of the RNA. After incubation, 4 µl 5X First-Strand Buffer and 2 µl DTT (0.1 M) are added to the mixture and incubated at 42 °C for 2 minutes. Finally, 1 µl SuperScript RT (200 U/µl) was added and centrifuge shortly. The reaction mixture were taken to thermal cycler and incubated at 50 °C for 50 minutes and 70 °C for 15 minutes. The samples were stored at -20 °C until use.

5.9.2 Quantitative real time PCR

For quantitative real-time PCR, 3 µl of 1:3 diluted cDNA was used as template. Primers (appendix), GAPDH sequences, were run in triplicate, using an Applied Biosystems 7900 Fast Real-time PCR System (Applied Biosystems Inc., Norwalk, CT, USA) in 384-well microtiter plates. The 15 µl PCR reaction comprised of 3 µl cDNA template, 0.3 µM of reverse primer, 0.3 µM of forward primer, 0.3 µM Rox and 7.5 µl 2X qPCR Master Mix (KAPA Biosystems). PCR was performed as follow: (i) 95 °C for 5 min; (ii) 95 °C for 3 sec and 60 °C for 1 min; (iii) repeat 40 times stage ii and followed by a melting curve analysis at 60 °C to 95 °C with 2.3 °C increment per second at each step. Threshold cycle (Ct) values were calculated using the 7900 Fast System SDS Software (Applied Biosystems Inc.), and further statistical calculations were performed using Microsoft Excel (Microsoft Corporation, Bellevue, WA, USA). For the quantification of the relative expression of the PLK4 mRNA, GAPDH was used as the endogenous control for normalization. Primers sequence are given in appendix 7.

5.10 Cell culture and cell lines

HeLa cell line (Scherer et al. 1953), and human fibroblasts established from wild-type and patient fibroblast were cultured in Dulbecco´s Modified Eagle's Medium (DMEM, PAA supplemented with 10% fetal bovine serum (FBS, Biochrom), L-Glutamine (PAN Biotech) and

antibiotics (Penicillin/Streptomycin, 50 µg/ml Penicillin, 50 µg/ml Streptomycin, PAN Biotech). Cells were grown in a humidified atmosphere of 95% air and 5% CO_2 at 37°C.

5.10.1 Establishment of patient primary fibroblasts

Primary dermal human fibroblasts were obtained by skin punch biopsy from an individual (MCP68-7) affected by microcephalic primordial dwarfism. The biopsy was immediately placed in sterile transport tube in 20 ml DMEM supplemented with antibiotics (50 mg/ml gentamycin (Sigma-Aldrich-G3632, USA), 75 µg/ml amphotericin B (GIBCO) and 50 µg/ml of each penicillin and streptomycin (PAN Biotech). Sample was shipped to the CCG in cool packs. Further experiments were carried out at institute of biochemistry, Cologne under sterilized conditions using laminar flow cabinet (Thermo Scientific). Tissue sample was cleaned in 7-10 ml tincture of iodine (Betaisodona, Povidon-Iodine) in a falcon tube to disinfect all the infectious agents. The tissue was incubated in three different falcon tubes containing 15 ml PBS mixed with 50 mg/ml gentamycin (Sigma-Aldrich-G3632, USA), 250 µg/ml amphotericin B (GIBCO) and antibiotics (Penicillin/Streptomycin, each of 10 IU/mL, PAN Biotech) for 5, 10 and 20 minutes respectively. The tissue was then incubated overnight with Dispase II (1.5 U/ml, Roche) diluted in PBS at 4 °C. After removing epidermis, the dermis was transferred to small petri dish (60x15mm) with 10 ml culture medium. The Petri plate was incubated at 37 °C. Medium was changed every three to four days. Fibroblasts outgrowth was checked every 4-5 days under the simple microscope (Olympus). After one week, primary fibroblast were observed making ring around the tissue. Cells were grown till the ~80% confluence and the tissue pieces were transferred to other plates.

5.10.2 Cultivation of human fibroblasts

Cells were grown in different plates and confluent cells were harvested by trypsinization: culture medium was removed; cells were washed once with 1XPBS; 3 ml pre-warmed trypsin (Sigma) was added to the 15cm^2 dish; incubated at 37 °C for 8-10 minutes; 10 ml of culture medium was added to stop the reaction; centrifuged at 1100 rpm for 5 minutes; culture medium was aspirated; cell contents were either plated in new dishes or were frozen at -80 °C for long period. In frozen case, the cell pellet was dissolved in total of 1 ml culture medium containing 100 µl of FBS and 100 µl of DMSO. After dissolving pellet, the medium was transferred to the 1.2 ml cryopreservation tube (VWR International). The tube was placed in Mr. Frosty and immediately transferred to -80 °C refrigerator.

5.10.3 Fibroblast synchronization at G2 phase

Patient and control cells were treated with 9 µM RO-3306 (Sigma-Aldrich, SML0569) for 22 hrs. RO-3306 has been shown to inhibit CDK1/cyclin B1 and CDK1/cyclin A and reversibly

arrests cells at the G2/M border of the cell cycle (Vassilev et al. 2006). The patient fibroblast were treated for 22 hours and the compound was then washed away to release the arrest. Cells were fixed and blocked at 20, 40 and 60 minutes for centriole quantification. In case of mitotic index, cells were fixed at 45 minutes after removal of G2/M block. The cells were then proceeded for immunofluorescence. The method of fixation for immunofluorescence is described below in section 5.11.

5.10.4 *PLK4* construct and transfection

 Full-length of human PLK4 (pcDNA3.1/3x myc-A, Plasmid #41165) was obtained from Add gene (Habedanck et al. 2005). This plasmid was used as template to generate enhanced green fluorescent protein (EGFP)-PLK4 wild-type plasmid. The PCR fragments were then cloned into pEGFP-C3 (BD Biosciences Clontech, 6082-1). Site-directed mutagenesis was performed to produce the mutant (PLK4 c.2811–5C>G) plasmid using a Quick-change XL mutagenesis kit (Stratagene, La Jolla, CA) according to manufacturer's protocol. PLK4 expression vectors were transfected into HeLa cells using 1mg/ml polyethylenimine (PEI, Polysciences). The high cationic charge density of PEI (Boussif et al. 1995) condenses the DNA into small positively charged complexes which are endocytosed by the cell through charged mediated interaction and subsequently the DNA is released into the cytoplasm (Sonawane et al. 2003). The DNA and PEI were mixed in the ratio of 1:3 in DMEM and incubated for 10-15 minutes at room temperature. HeLA cells were then transfected with wild-type GFP-PLK4 and mutant GFP-PLK4. After 48 hours of transfection cells were processed for immunofluorescence.

5.10.5 Determination of growth

The 1×10^5 of patient (passage 2) and control primary fibroblasts (passage 5) were seeded per well of six-well plate. Cells were grown in normal culture medium as described above. For analysis, cells were trypsinized and counted every 24 hour for 6 days using a Neubauer-Improved chamber. The counted cells were recorded for final results.

5.11 Immunofluorescence and microscopy

Human fibroblasts were grown either on 1000 µg/ml fibronectin (Sigma-Aldrich) coated or non-coated 12 mm glass coverslips (Menzel-Gläser). Coverslips were placed in each wells of either six-well or twenty four-well cell culture plates. Approximately 1×10^5 cells were seeded per well. 1-2 ml culture medium was added to each well. Cells were processed for immunofluorescence as follows. The media was aspirated and cells were incubated shortly with PBS. Cells were fixed with 3% PFA for 15-20 minutes at room temperature. Fixation agent was removed and cells were washed four times with PBS or tubulin stabilization buffer (TSB). Cells were

permeabilised with 3% BSA, 1% TritonX-100 in PBS for 3-4 minutes at room temperature. For tubulin staining, cells were fixed using prechilled absolute methanol for 5 min at −20 °C without permeabilization and all the washing steps were performed with tubulin stabilization buffer (TSB). After incubation with 1X PBS or TSB, the cells were blocked for 30 minutes with blocking buffer (1x PBG: PBS containing 5% BSA and 0.45% fish gelatin). After blocking, primary antibodies were diluted in blocking buffer and incubated over night at 4 °C or ≥1h at room temperature. Primary antibody solution was aspirated and the cells were incubated 3 times in PBS. After this, the cells were incubated with secondary antibody for 30-60 minutes at room temperature protected from light. Cells were then treated with 1X PBS three times each for 5 minutes in dark condition. Finally, the cells were mounted on glass slides with Gelvatol. Dried the slides at 37 °C for 1 hour or at room temperature for 2-3 hours. Images were taken with a confocal laser microscope (Leica, LSM TCS SP5). Image processing was done with Leica LAS AF Lite software and Adobe Photoshop (CS6).

Overexpression assay: After 48 hour transfection, HeLa cells were fixed as described in the previous section. Primary antibodies Cyclin A and Centrin-2 were diluted in IX PBG and incubated overnight at 4 °C. The mo-Cyclin A antibody (Abcam ; 1:1000) was used to indicate cell-cycle position (Maridor et al. 1993) while rab-centrin2 (Santa Cruz Biotechnology, sc-27793-R; 1:500) was used to quantify centriole number and localization (Salisbury et al. 2002).

Centriole quantification at mitotic phase: Primary antibodies used were mo-centrin 3 (Novus biologicals, H00001070-M01; 1:500) and rat monoclonal anti α-tubulin 1:20 (Kilmartin et al. 1982).

Mitotic index: Patient and control cells were fixed with pre-chilled methanol and stained with mo-histone H3 phosphorylated at Ser10 (Cell signaling, 9706; 1:200 dilution) and rat monoclonal anti α-tubulin (Y/L1/2; 1:20).

Immunostaining of gamma tubulin: The cells were fixed and permeabalized with 100% prechilled methanol for 5 minutes at −20 °C. The cells were blocked and proceed for immunofluorescence as described previously. The mouse- γ-tublin antibody (Sigma; GTU-88 1:300) was used as centrosomal marker.

PLK4 signal quantification at centriole: The patient and wild-type fibroblast were stained with γ-tubulin and PLK4 (dianovoa, Germany; DIA-444) antibodies (Cizmecioglu et al. 2010). The cells were treated with 10 µM MG132 (Calbiochem®, 474790-1MG) for 5 hours before fixation. The MG-132 is a robust proteasome inhibitor that is cell-permeable (Elliott et al. 2003). The PLK4 signal was quantified using Volocity software and AlphaEaseFCTM software (version 4.0.0) (Alpha Innotech, San Leandro, CA, USA).

5.11.1 Solutions used in Immunocytochemistry

Gelvatol PBG (pH 7.4)

Polyvinyl alcohol (87%-89%, Sigma P8136)

Glycerol

Ionized water

0.2 MTris-HCl (pH 8.5)

centrifugation (5000 g, 15 min),

2.5% Diazabicyclooktan (DABCO, Sigma Aldrich, D2522)

Aliquot-storage: –20 °C

10X Hanks buffer

137 mM NaCl

5 mM KCl

1.1 mm Na2HPO4

0.44 mM KH2PO4

5 mM glucose

4 mm NaHCO3

1XTubulin stabilization buffer (pH 6.8)

10X Hank's buffer

1 mm MES

2mM EGTA

2mM MgCl2

10x Triton X-100 (25ml)

Triton X-100	2.5 ml
1x PBS	22 ml

0.5% Triton X-100 (50ml)

10x Triton X-100	2.5 ml
1x PBS	47.5 ml

Store at 4 °C

3% paraformaldehyde (pH 6.1, 100 ml)

paraformaldehyde powder	3g
1X PBS	80ml
0.4 M NaOH (to adjust pH 6.1)	
deionized H2O (final volume)	

PBS (pH 7.2)

10 mM NaCl

10 mM KCl

32 mM KH2PO4

16 mM Na2HPO4

Note: IxTSB was used instead of 1X PBS for microtubule staining.

5.12 FACS analysis

Primary fibroblasts derived from a patient MCP68-7 carrying a mutation in *PLK4* along with wild type were used for the apoptotic assay. The cells were treated with PBS and trypsinized with StemPro Accutase Cell Dissociation Reagent (A1110501, Life Technologies). The cells were pelleted for 5 minutes at 1100 rpm. For each FACS reaction 1×10^6 cells were

resuspended in Annexin V binding buffer (422201, Biolegend, San Diego, CA). After incubation with APC Annexin V (640919, Biolegend) and 7-AAD (420403, Biolegend), for 15 minutes at room temperature, the apoptotic cells were analyzed immediately by using the BD FACSAira III cell sorter (BD Bioscience USA), and data were analyzed using DIVA 6.0 (BD Biosciences) software.

5.13 Preparation of protein lysates from mammalian cells

Cells were scrapped with ice cold 1X phosphate buffered saline (PBS). The cells were spun for 5 minutes at 14,000 rpm, with a temperature of 4 °C in a centrifuge. PBS was drained and the cell pellet was resuspended in cold lysis buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% NP-40, 0.5% Na-desoxycholat, 0.1% SDS and Proteinase Inhibitor Cocktail (PIC, Sigma) supplemented with protease inhibitors 1 mM DTT, 1 mM Benzamidine and 100 mM PMSF. The amount of lysis buffer was added usually 2-4 fold volume of the cell pellet. To fragment the DNA, the lysed suspension was passed several times through 0.4x19 mm needles (27Gx3/4", Nr.20, BD Microlane TM3) attached to 1 ml syringe. The lysate was placed on ice for 30 minutes and then centrifuged at 14,000 rpm for 15 minutes. The volume of the pellet was estimated. The sample was denatured in SDS sample buffer at 95 °C for 5 minutes. After denaturation, the sample was either frozen or processed further.

5.14 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were resolved using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The SDS polyacrylamide gels consisted of a running and a stacking gel. The percentage of the resolving and stacking gel contained 10-15% and 4% acrylamide respectively. The gels were polymerized for 30-40 minutes at room temperature. Isopropanol was used to smoothen the top of resolving gel. SDS-PAGE was performed with the Biorad Mini-PROTEAN® Gel Chamber. For all buffers and solutions see the section 5.15. 10-40 µl of protein lysate and 4 µl of Page Ruler Plus Prestained Protein Marker (Fermentas; 26619) were loaded and the gels ran in 1X SDS PAGE running buffer first at 60 V for the stacking gel and then the resolving gel at 100-120 V. Gels were stained with Coomassie-Brilliant-Blue R 250 solution for 30 minutes at room temperature and after that destained in Coomassie destaining solution. The gels were then used for western blots.

5.14.1 Solution used for SDS-PAGE

5X SDS-sample buffer

0.125M Tris/HCl (pH 6.8)

20% Glycerol

1X SDS-running buffer (pH 8.3)

25 mM Tris-base

192 mM Glycine

4% SDS	0.1% SDS
10% ß-mercaptoethan	
Bromophenol blue	

Coomassie Blue staining solution (500ml)

0.25% Brilliant-Blue R250 (Sigma)

50% Ethanol absolute

10% Acetic acid

Destain solution

10% Ethanol

7% Acetic acid

5.15 Western blot

After proteins separation through SDS-PAGE, wet blotting was performed to transfer the proteins from gels to a nitrocellulose membrane (PROTRANR, Germany).The wet blot were set up from anode to cathode as sponge, two layers of whatman filter paper, the membrane, the gel, another two layers of whatman filter paper and sponge. All these were clamped tightly together after ensuring no air bubbles had formed between the gel and the membrane. Thereafter complete sandwich was submerged in transfer buffer. The blot was run at 4 °C overnight, and an electrical field of 15 V was applied. Next day, nitrocellulose membrane was blocked with Tris-buffered saline (TBS) containing 5 % dry milk for 1 hour at room temperature with shaking. Membrane was incubated with primary antibody diluted in IX TBS for 1 hour at room temperature or at 4 °C overnight with shaking. After primary antibody incubation, membrane was washed with 1X TBS three times each for 5 minutes. The membrane was incubated with corresponding appropriate horseradish peroxidase-conjugated secondary antibody for 1 hour at room temperature with shaking. Membrane was then washed three times for 5 minutes each with 1X TBS. Afterward, the membrane was incubated with 5 ml of Enhanced ChemiLuminescence (ECL) reagent for 1 minute. The membrane was exposed to an X-ray film in dark room and developed by using developer and fixer solution. The same membrane was stripped with 0.4 M NaOH for 15 minutes and washed twice with TBS. After washing step, the membrane was blocked and incubated with the new primary and secondary antibodies in the same way as described.

5.15.1 Solution used for SDS-PAGE

Wet blot buffe	**10X Tris-buffered saline (TBS)**
39 mM Glycine	1 M Tris/HCl (pH 7.4)

48 mM Tris-HCl, pH 8.0   5 M NaCl

20% Absolute ethanol

1X Tris-buffered saline-Tween 20 (TBS-T)

10% 10X TBS buffer

1% Tween 20 (Sigma)

ECL detection solution

p-Cumaric acid (0.09 M in DMSO)	89 µl
Luminol (0.25 M in DMSO)	200 µl
1M Tris/HCl (pH 8.5)	2 ml
Water (dH2O)	18 ml
30% H_2O_2	6.1 µl

5.16 List of used antibodies for immunofluorescence and western blotting

| Antibody	Working concentration	Manufacturer
Primary antibodies for immunofluorescence		
Mouse monoclonal γ-tubulin	1:300	Sigma, GTU-88
Mouse monoclonal centrin 3	1:500	Novus Biologicals, H00001070M01
Mouse monoclonal pH3	1:200	Cell signaling, 9706
Mouse Cyclin A	1:1,000	Abcam
Rabbit polyclonal PLK4 antibody	1:100	Abcam, ab137398
Rabbit polyclonal centrin 2	1:500	Santa Cruz Biotechnology, 27793-R
Rat monoclonal anti Y/L1/2	1:20	(Kilmartin et al. 1982)
Secondary antibodies for immunofluorescence		
Alexa Fluor® 568 goat anti mouse Ig	1,000	Life Technologies, A11004
Alexa Fluor® 647 goat anti-mouse IgG	1:1,000	Life Technologies, A31573
Alexa Fluor® 488 goat anti-mouse IgG	1:1,000	Life Technologies, A11001
Alexa Fluor® 488 goat anti-rat IgG	1:1,000	Life Technologies, A11006
Alexa Fluor® 568 goat anti-rabbit IgG	1:1,000	Life Technologies, A11011
Primary antibodies for western blotting		
Mouse polyclonal PLK4 antibody	1:200	Abnova, DIA-444
Rat monoclonal anti Y/L1/2	1:1	(Kilmartin et al. 1982)
Secondary antibodies for western blotting		
Anti-mouse IgG peroxidase conjugated	1:10,000	Sigma, A4416
Anti-rat IgG peroxidase conjugated	1:10,000	Sigma, A5795
DNA staining agent		
4',6-Diamidino-2'-phenylindole (DAPI)	1:5,000	Sigma, D9542

6. Results

6.1 Autosomal recessive primary microcephaly (MCPH)

6.1.1 Homozygosity mapping in MCPH families

In an effort to find novel genetic causes of primary microcephaly, twenty nine consanguineous families segregating this disorder were genotyped. Initially, affected individuals from six families (MCP101, MCP104, MCP105 MCP106, MCP117 and MCP118) were scanned for homozygous genomic regions using polymorphic microsatellite markers spanning the first seven loci reported to cause primary microcephaly, i.e. MCPH1 to MCPH7 loci. Of the six families, five showed homozgosity at known MCPH loci except MCP104. Thereafter array-based genome mapping was performed in thirteen MCPH families to identify homozygous regions shared between the affected individuals. These families also included four families which were tested with STR markers before (MCP104, MCP105, MCP106 and MCP120). Genome-wide linkage analysis revealed four known MCPH loci in 12 of the 13 families. Homozygosity mapping was followed by exome sequencing and Sanger sequencing of candidate genes. In addition, *ASPM* was sequenced in 14 additional families, according to the literature, mutations of *ASPM* are a very common cause of congenital microcephaly among families from Northern Pakistan (Roberts et al. 2002). The linkage status of all studied families is summarized in Table 6.1.1 and Fig 6.1.1.

Table 6.1.1: Linkage status of families with primary microcephaly.

	MCPH1	MCPH2	MCPH3	MCPH5	Excluded Families	Total
No. of families	2	1	2	23	1	**29**
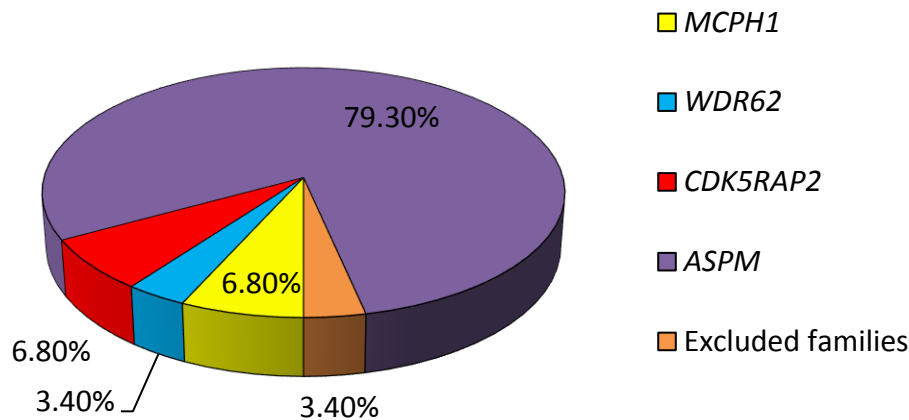
Figure 6.1.1: A pie chart illustrating the linkage status of the MCPH families. Among 29 primary microcephaly families, 79.3% (purple colored) were linked to the MCPH5 locus, 6.8% (yellow) with MCPH1, 3.4% (blue) with MCPH2, 6.8% (red) with MCPH3, 3.4% (marigold) were excluded with all known MCPH loci.

6.1.2 MCPH1 linked families

Two primary microcephaly families, namely MCP118 and MCP125, were linked to the MCPH1 locus which was originally mapped to chromosome 8p22-pter within a 13-cM region between markers D8S1824 and D8S1825 (Jackson et al. 1998).

6.1.2.1 Family MCP118

The MCP118 family is a large consanguineous family that was ascertained from the Punjab province of Pakistan (Fig 6.1.2). The family consists of seven affected individuals; two of these died in the first decade of their life due to unknown reason, while the rest of the patients aged from18-44 years. All the affected individuals had moderate intellectual disability and reduced head circumference (-6SD to -7SD). All the patients had impaired ability to communicate and to learn. There was no history of seizures and epilepsy. DNA samples from three affected members of the family were genotyped using Axiom™ Genome- Wide CEU 1 array (Affymetrix). Parametric linkage analysis was performed and mapped the peaks on chromosome 8, 11, 18 and 22 with maximum possible LOD score (ALLEGRO) of 3 (Fig 6.1.3). The underlying homozygous regions were investigated manually with the NCBI and UCSC genome databases. The linkage region on chromosome 8 contains *MCPH1*, a known MCPH gene (encoding microcephalin). This region maps between markers AX-11648363 (position: 2,408,965 bp; hg19) and AX-11328718 (position: 6,404,947 bp; hg19) and results in 3.9 Mb linkage region. For mutation detection, all 14 exons and exon–intron boundaries of *MCPH1*

(MCPH1-001, ENST00000344683) were PCR amplified. PCR failed to amplify exons 1-2 in all affected individuals while PCR products were generated for normal individuals, suggesting the presence of a homozygous deletion. In parallel, whole-exome sequencing was also performed on one patient (MCP118-5) to exclude variants in other known MCPH genes. Data were analyzed by our in-house exome analysis pipeline (Varbank v.2.1) that successfully detected the homozygous deletion of first two exons in *MCPHI* gene (Fig 6.1.6). Due to the limitation of sequencing coverage, we followed by performing genome-wide CytoScan® HD array (Affymetrix) in one affected individual (MCP118-6) (Fig 6.1.4). Analysis of genotyping data demonstrated a drop off hybridization signal intensity for the 164 kb DNA segment. The flanking probes S-3YKOB (position: 6,246,392 bp, hg 38) and C-6BVSB (position: 6,410,707 bp; hg38) were found in the last undeleted region. This deletion was absent in the large cohort of healthy individuals assessed from DGV (database of genome variant). The homozygous deletion affects the first two exons and the 5′ upstream region of *MCPH1*. To further map the precise breakpoints, primers across the deletion was designed (appendix 4). These PCR primers generated ~590 bp product (Fig 6.1.5-a). This PCR amplicon was Sanger sequenced which redefined the coordinates predicated by CytoScan array. To further confirm, we sequenced 5′-end and 3′-end of breakpoint by designing one primer of the pair in the deleted region. A PCR product was amplified for healthy individuals of the family while in case of affected individuals; no PCR product was obtained (Fig 6.1.5-b, c). The new refined coordinates of the break point are 6,246,012 bp to 6,410,262 bp (cytoband 8p23.1). The microdeletion in the MCP118 family comprises of the first 2 exons and extends ~160 kb upstream to the 5′ end of the *MCPH1* (Fig 6.1.7).

Family MCP118
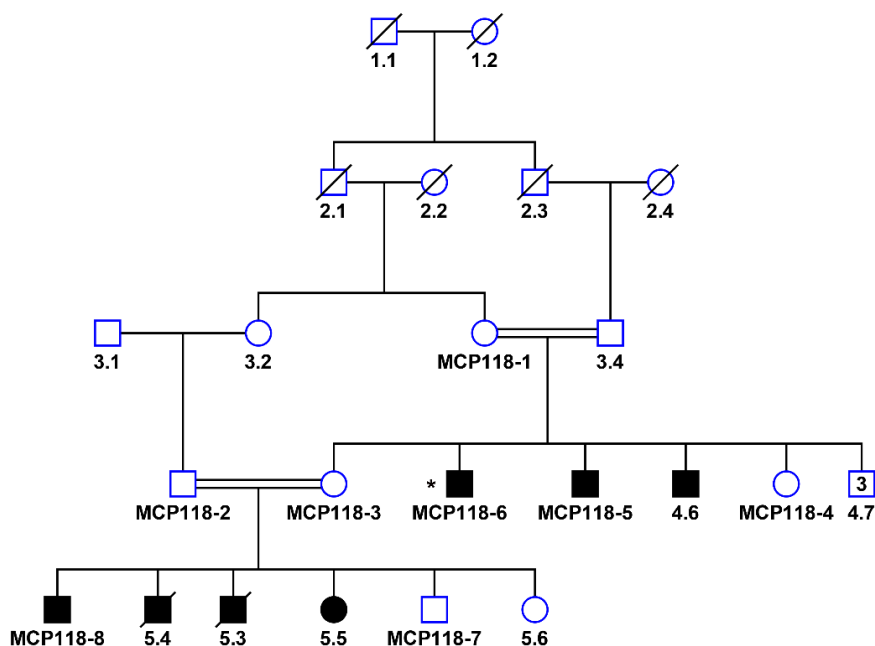
Figure 6.1.2: Pedigree of family MCP118 with seven afflicted sibs and two healthy sibs along with their parents, showing a recessive mode of inheritance. Filled symbols represent affected individuals. A double line represents consanguinity. The slashed denote the deceased individuals. An asterisk represents the subject analyzed by CytoScan array.

(a)   **(b)**
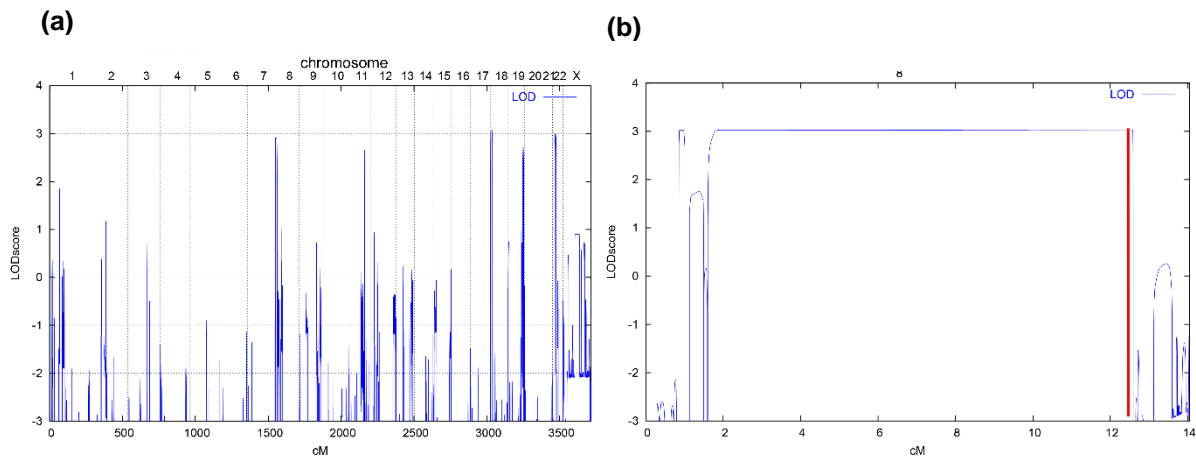
Figure 6.1.3: (a) Genome-wide linkage data demonstrates high peaks on chromosomes 8, 11, 18 and 22. The highest LOD score (ALLEGRO) value was calculated z ~3. The underlying regions for these peaks are 3.9 Mb on chromosome 8 (2,408,965-6,404,947 bp; hg19), 1.7 Mb on chromosome 11 (110,347,929-112,094,487 bp; hg19), 3.4 Mb chromosome 18 (903,640-4,354,096; hg19), 1.9 Mb on chromosome 22 (25,275,523-27,198,719 bp; hg19). **(b)** The LOD plot on chromosome 8 shows the linkage region with position of *MCPH1* gene as red line, 0 refers to 2.3 cM.
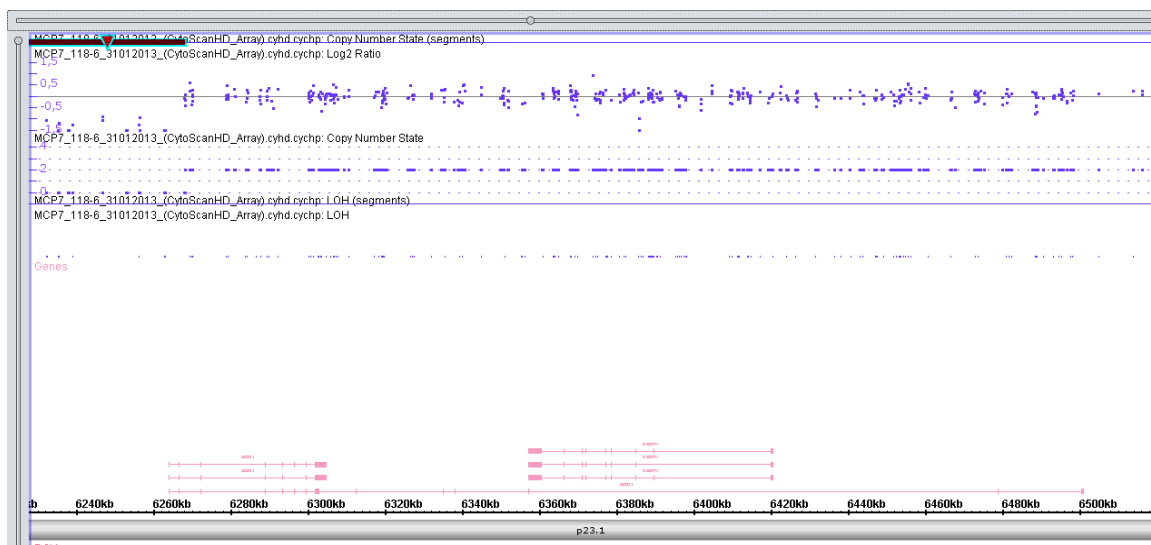

Figure 6.1.4: The region of *MCPH1* copy number loss is highlighted by dark red line on top and evidenced by the reduced log2 fluorescence ratios. Each dot shows a log2 ratio of the hybridization signals of patient versus gender-matched control for each probe in the *MCPH1*. The LOH region is shown below the copy number state plot. The horizontal line at bottom represents the physical position of the oligonucleotide probes on 8p23.2-8p23.1 (hg 19), with distal p-arm probes (telomeric) on the left end of the plot and the proximal p-arm probes (centromere) on the right end.
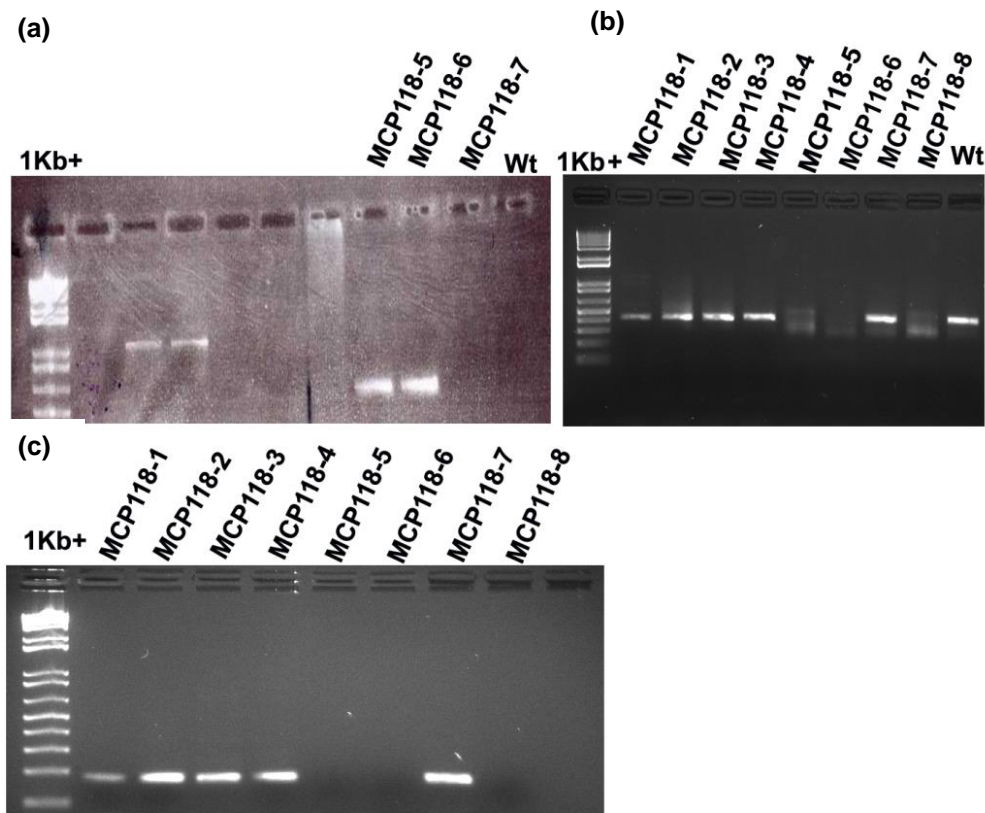
Figure 6.1.5: Representative gel images of the PCR products of junction fragments to validate the break point in *MCPH1* locus. **(a)** PCR product amplified with primers across the break point. Amplification was only detected for patients (~590 bp) **(b)** 3′- junction amplification: PCR was performed with one primer designed at the flanking region of microdeletion and with the other primer within the deleted region. Only DNA of healthy individuals of the family was amplified. **(c)** 5′- junction amplification: PCR product amplified with primers designed at the flanking region of microdeletion and with the other primer within the deleted region. Amplification was only detected for healthy individuals of the family. PCR products underwent electrophoresis in a 2% agarose gel and were stained by ethidium bromide. 1 Kb plus DNA ladder is used to estimate the molecular size of amplified DNA.
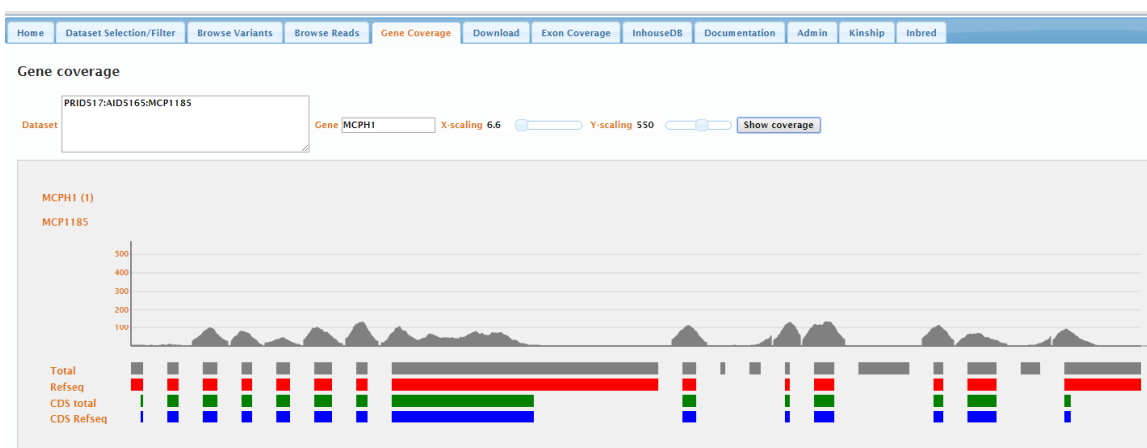
Figure 6.1.6: Snap shot shows no coverage for first two exons in *MCPH1*. The mean coverage for the sample was 70x, 96 % of sequences had at least 10x coverage and 85 % of the sequences were covered at least 30x.

48
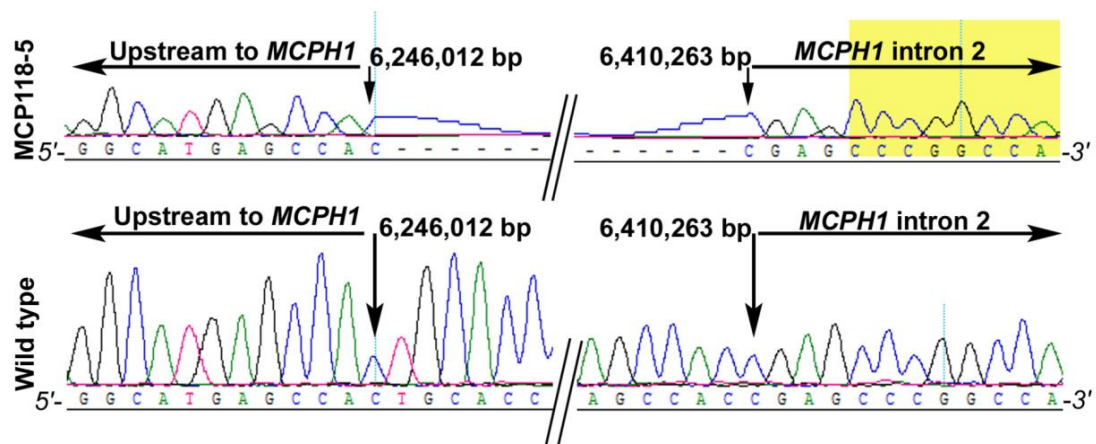
Figure 6.1.7: The electropherogram shows the breakpoints in MCP118. Sequence coordinates are listed according to ensemble hg GRCh38 (December, 2013). The wild type sequence is shown below deletion. The arrows are pointing to the last undeleted nucleotides at both upstream 5' end and 3' end. In total, a region of 164,250 bp is deleted in *MCPH1*.
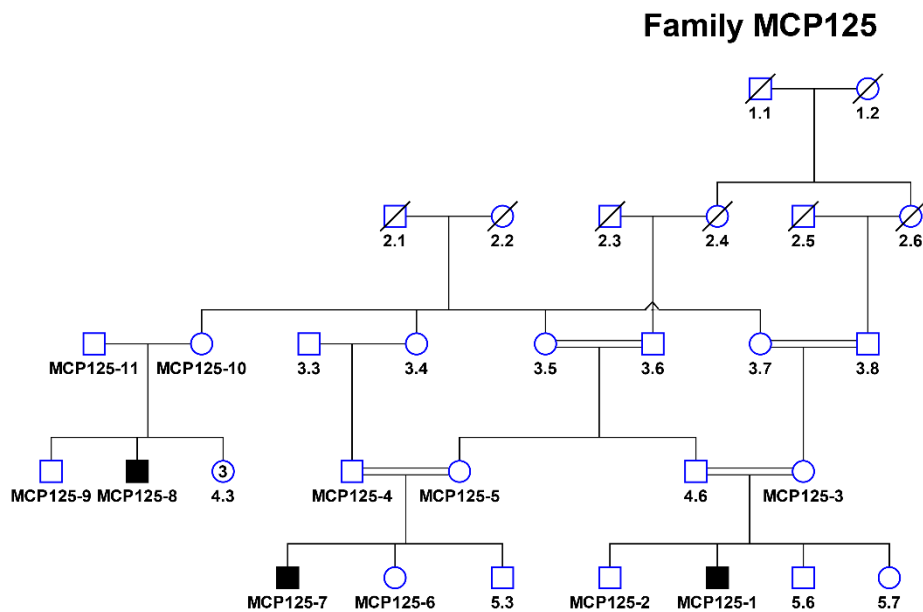
6.1.2.2 Family MCP125

Family MCP125 is a large consanguineous family of Baloch origin, ascertained from District Tando Allahyar, Sindh, Pakistan. The DNA samples were available from 11 individuals including three patients (Fig 6.1.8-a). The affected individuals had normal height and reduced head circumference (-6 SD to -7 SD). They had mild intellectual disability and limited self care skills. All patients are challenged in learning and communication. There was no history of epilepsy.

A genome-wide linkage analysis of the family was done by genotyping all affected individuals using Affymetrix Axiom™ Genome-Wide CEU 1 arrays. Due to the computational constraints, the large family was spilt into three subfamilies and coded as second degree of consanguinity. The HLOD and LOD scores were calculated with a reduced panel of 20K SNPs. Model-based linkage analysis was performed and we detected seven peaks on chromosome 1, 3, 6, 8, 12 and 15 with maximum multipoint LOD scores (ALLEGRO) of z ~5.4. Next, to identify homozygous region shared by all three affected individuals, these chromosomes were analyzed again with high resolution by selecting maximum markers from the Axiom array. Five homozygous regions that were 6.8 Mb on chromosome 3 (51,754,424-58,561,548 bp; hg19), 13.8 Mb on chromosomes 6 (167,510-14,056,150 bp; hg19), 3.7 and 9.3 Mb on chromosome 8 (2,011,098-5,782,057 bp and 6,323,132-15,674,115 bp; hg19) and 4 Mb on chromosome 12 (74,146,967-78,181,072 bp; hg19) were detected, shared by at least two individuals. Further calculations were carried out to look for the regions where all three cases had identical

haplotypes. There were many small homozygous regions but large regions of 3 Mb and 2 Mb were found on chromosome 6 and 8, respectively. The haplotypes were only constructed for chromosome 8 which contain *MCPH1* (6,264,113-6,501,140 bp; hg19), a known gene for primary microcephaly. We tried to sequence all 14 exons and exon–intron boundaries of the *MCPH1* (MCPH1-001, ENST00000344683). PCR failed to amplify exons 1-11 in all affected individuals while PCR products were obtained for normal individuals. At the same time, whole-exome sequencing was performed on the DNA of one patient (MCP125-7). Snapshot from the interface of our in-house pipeline varbank (v.2.1) shows no coverage for the first eleven exons (Fig 6.1.12). These observations were taken as evidence for the presence of homozygous deletion. For CNV analysis, a genome-wide CytoScan® HD array was used in one affected individual (MCP125-7) (Fig 6.1.10). Genotyping data were analyzed in the same way as performed for the sample MCP118-6 and we detected a homozygous deletion of ~577kb that covered the upstream region and first eleven coding exons of *MCPH1*. On careful inspection of the genotyping data, probes C-7JUTI (position: 5,911,493 bp; hg38) and C-6UQKI (position: 6,492,876 bp; hg 38) were found to flank the deletion.

In similar vein like in family MCP118, to map the exact breakpoint boundaries, PCR primers were selected in the undeleted region (based on the CytoScan array). These PCR primers generated ~1030 bp product which was sequenced through Sanger sequencing (Fig 6.1.11-a, b, c). The Sanger sequencing redefined the above coordinates of Cyto Scan array and clearly delineated the breakpoint at position 5,913,147 bp to 6,490,742 bp (Fig 6.1.13). In MCP125, the microdeletion removed the first 11 coding exons and ~493 kb upstream in the 5′ region Fig 6.1.13). No exonic point mutation of *MCPH1* was identified by Sanger sequencing of the parents.
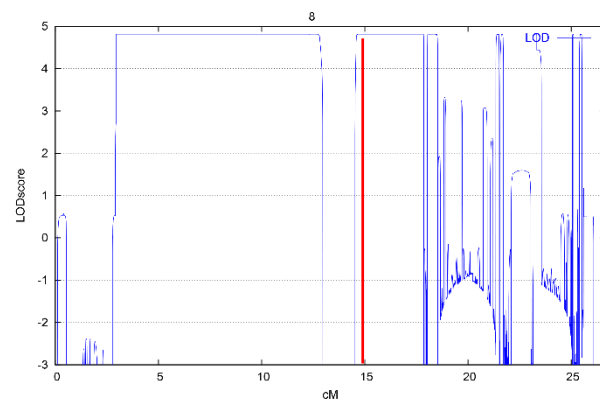
(a)

Family MCP125

(b)
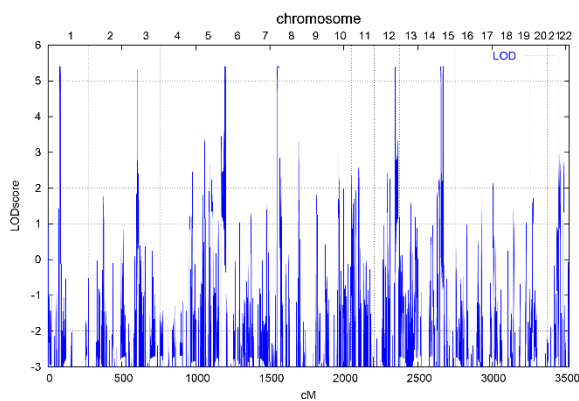
(c)

Figure 6.1.8: (a) Pedigree of family MCP125. An asterisk represents subject analyzed by CytoScan array. (**b)** Genome-wide scan was generated based on the genotypes of three affected siblings. The highest LOD score values were obtained for regions on chromosomes 3, 6, 8 and 12 with maximum values of z~ 3.8. **(c)**The LOD plot on chromosome 8 shows the linkage region with position of *MCPH1* as red line, 0 refers to 0.26 cM.
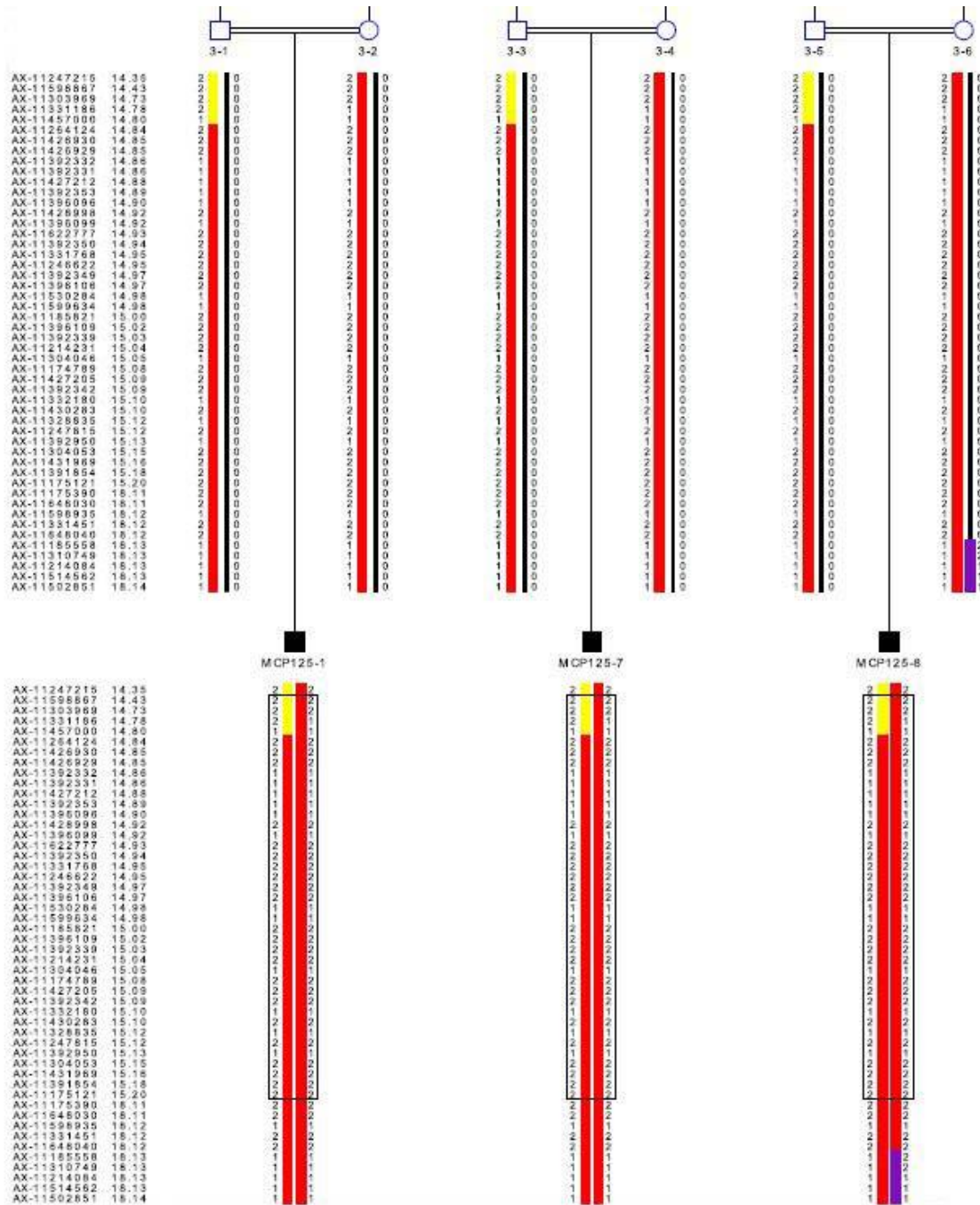
Figure 6.1.9: Haplotype reconstruction from the genotyping of three affected individuals of family MCP125 (Axiom ™ Genome-wide Array (Affymetrix). The markers AX-11457000 (Position: 6,334,558 bp; 14.8 cM; hg 19) and AX-11185558 (position: 8,388,493 bp; 18.1 cM; hg19) are the flanking markers for the shared homozygous region on chromosome 8. The markers within the *MCPH1* are shown in the box.
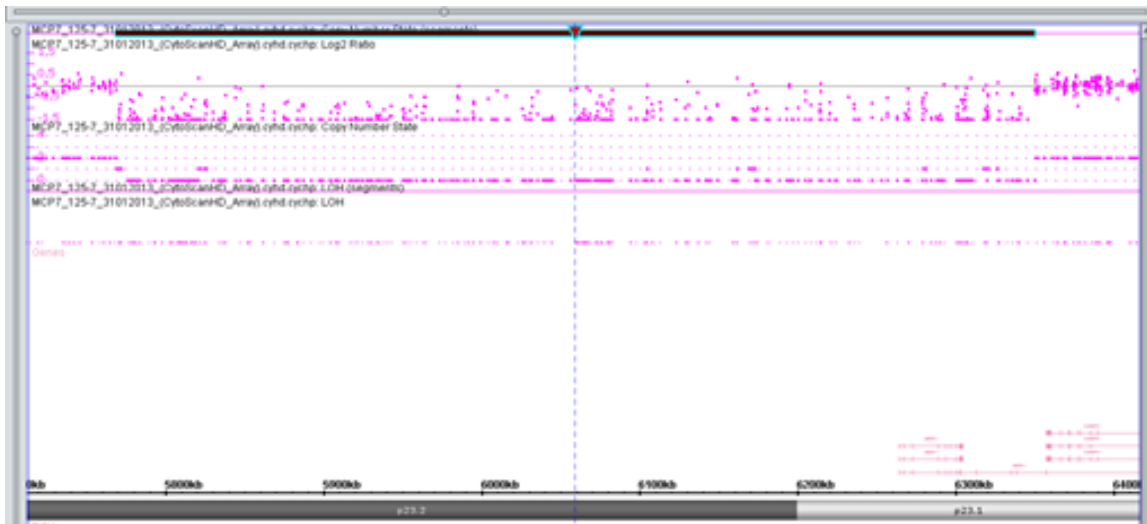
Figure 6.1.10: The region of *MCPH1* copy number loss is highlighted by dark red line on top. Each dot shows a log2 ratio of the hybridization signals of patient versus gender-matched control for each probe in the *MCPH1*. The LOH region is shown below the copy number state plot. The horizontal line at bottom represents the physical position of the oligonucleotide probes on 8p23.2-8p23.1 (hg19), with distal p-arm probes (telomeric) on the left end of the plot and the proximal p-arm probes (centromeric) on the right end.
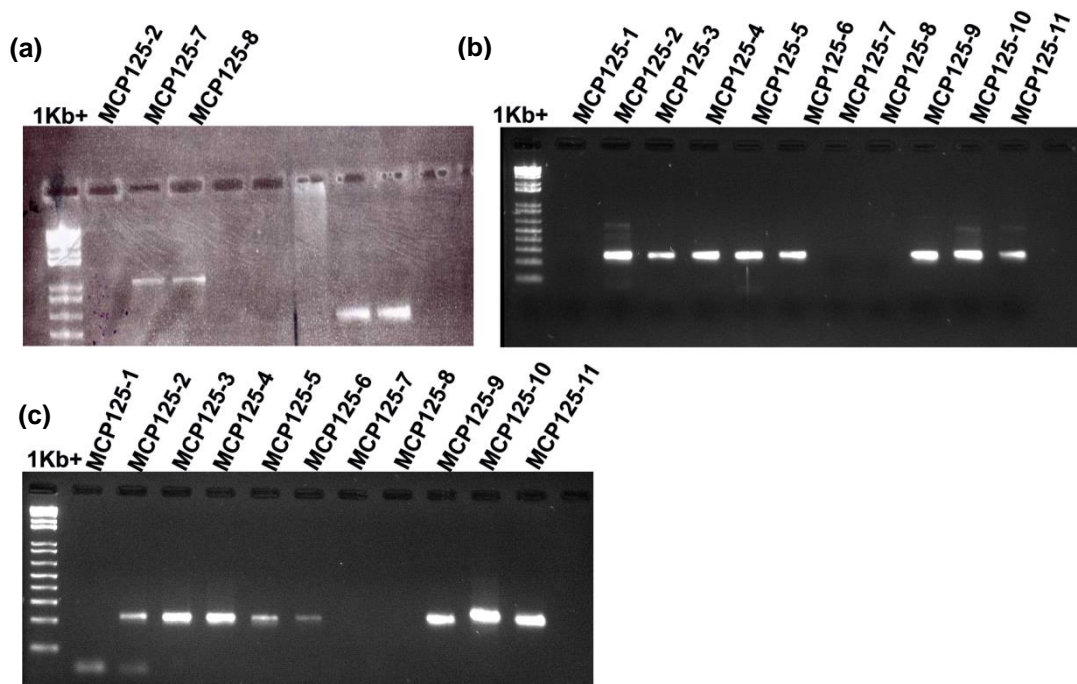
Figure 6.1.11: Representative gel images of the PCR products of junction fragments to validate the break point in *MCPH1* locus. **(a)** PCR product amplified with primers across the break point. Amplification was only detected for patients (~1030 bp). **(b)** 3′- junction amplification: PCR was performed with one primer designed at the flanking region of microdeletion and with the other primer within the deleted region. Only DNA of healthy individuals of the family was amplified. **(c)** 5′- junction amplification: PCR product amplified with primers designed at the flanking region of microdeletion and with the other primer within the deleted region. Amplification was only detected for healthy individuals of the family. PCR products underwent electrophoresis in a 2% agarose gel and were stained by ethidium bromide. 1 Kb plus DNA ladder is used to estimate the molecular size of amplified DNA.
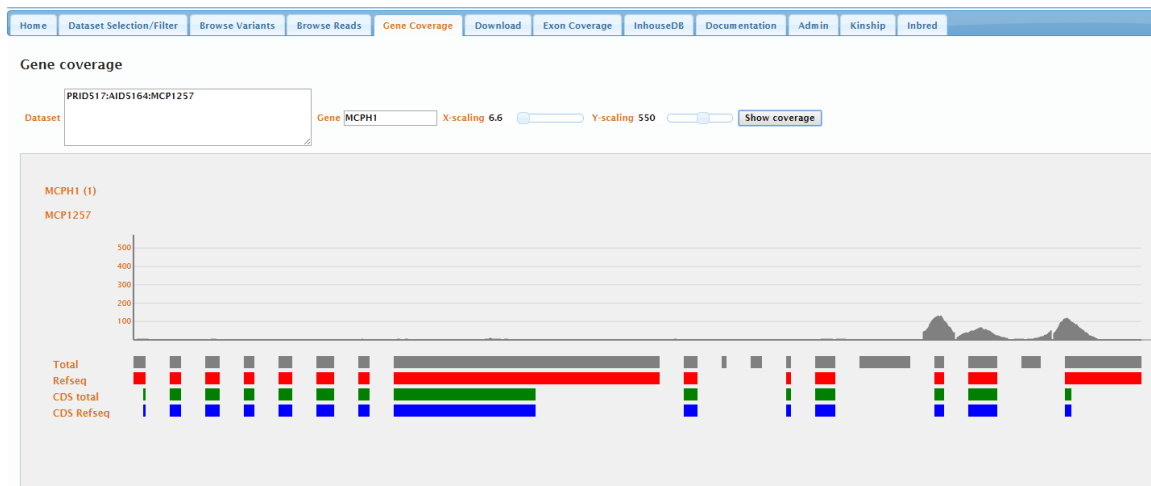
Figure 6.1.12: Snap shot shows no coverage for first 11 exons in *MCPH1*. Mean coverage for the sample was 65x with 95 % of target sequences being covered at least 10x and 83 % of the sequences at least 30x.
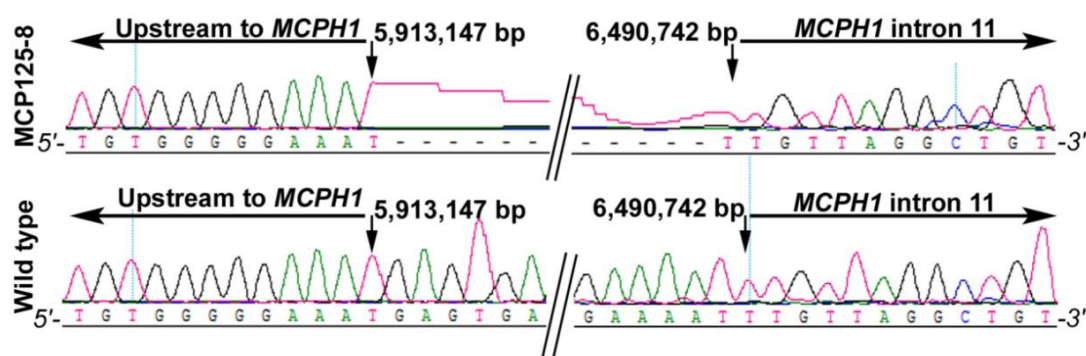

Figure 6.1.13: Sanger sequencing chromatograms showing the *MCPH1* genomic breakpoint and corresponding wild type sequence from 3' and 5' ends. The exact positions of the breakpoint are marked with arrows. Numbers are based on the nucleotide position according to the human genome assembly GRCh38 coordinates (December, 2013). In total, a region of 577, 594 bp is deleted in *MCPH1*.

6.1.3 MCPH2 linked families

Family MCP129 was linked to the MCPH2 locus which is located on chromosome 19q13.12. (Roberts et al. 1999). It is a large consanguineous pedigree with four affected individuals (Fig 6.1.14). This family was recruited from District Jamshoro in the Sindh province of Pakistan. The evaluation of the affected individuals was carried out at a local hospital. In all patients, microcephaly was noticed at birth and their head circumferences were between -5 SD and -7SD. The patients had profound intellectual disability and speech impairment. The patients showed aggressive behavior; one case MCP129-1 had hypotonia in the lower limbs. All affected individuals were genotyped using Affymetrix Axiom™ Genome-Wide CEU 1 arrays.

Multipoint linkage analysis (ALLEGRO) was performed and we got a significant LOD score on chromosome 19. Next we selected all markers on chromosome 19 and calculated parametric LOD scores of z~3.9 (Fig 6.1.15). The underlying homozygous linkage region of 20.6Mb is flanked by markers AX-11666108 (position: 17,872,092 bp; 42.8 cM; hg19) and AX-11668721(position: 38,486,149 bp; 63.6 cM; hg19). Map Viewer tool of NCBI and UCSC genome browser were used for the candidate gene identification. *WDR62* was concluded as the only candidate gene, based on its association with the phenotype in the previous reports. On subsequent Sanger sequencing of the *WDR62*, a known missense mutation with a single nucleotide transversion (c.332G>C) (ENST00000401500) in exon 3 was identified that resulted in the substitution of arginine by threonine (p.Arg111Thr) (Fig 6.1.16). The disease allele segregated faithfully within the family. All the four affected cases were found to be homozygous for the mutation and parents were heterozygous. These result support an autosomal recessive mode of inheritance.
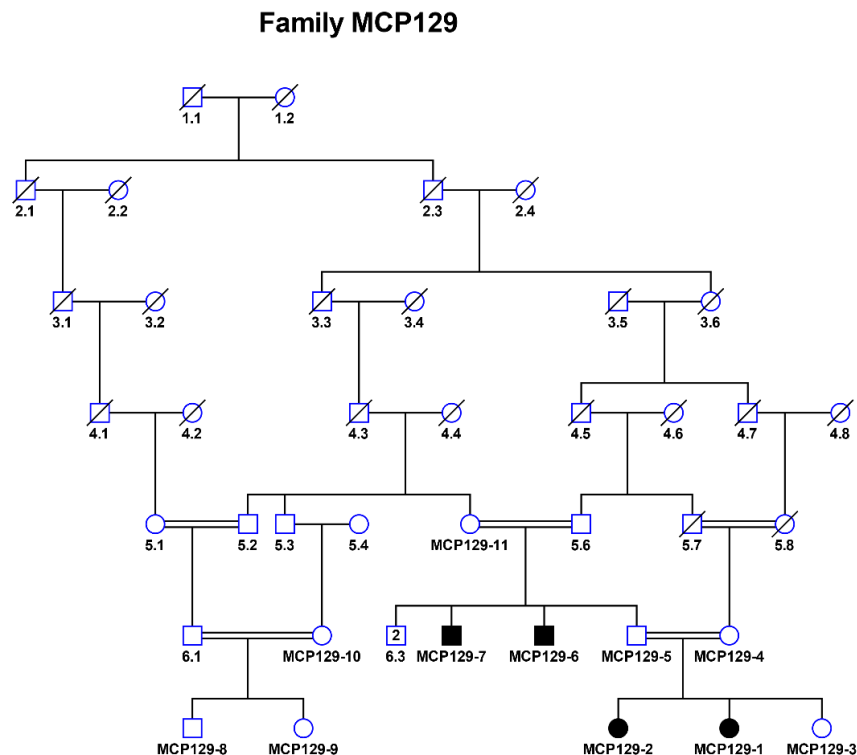
Figure 6.1.14: Pedigree of family MCP129 with autosomal recessive primary microcephaly. All affected individuals were genotyped for genome wide linkage analysis.
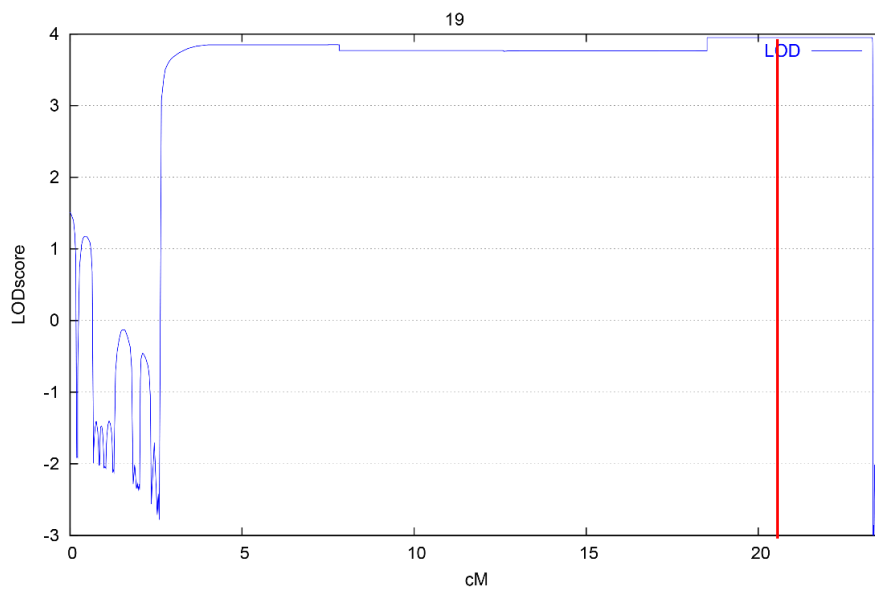
Figure 6.1.15: The LOD plot on chromosome 19 shows the linkage region with position of *WDR62* as red line, 0 refers to 40.2 cM.

MCP129-7, homozygous mutant, c.332 G>C

MCP129-11, heterozygous parent, c.332 G>C
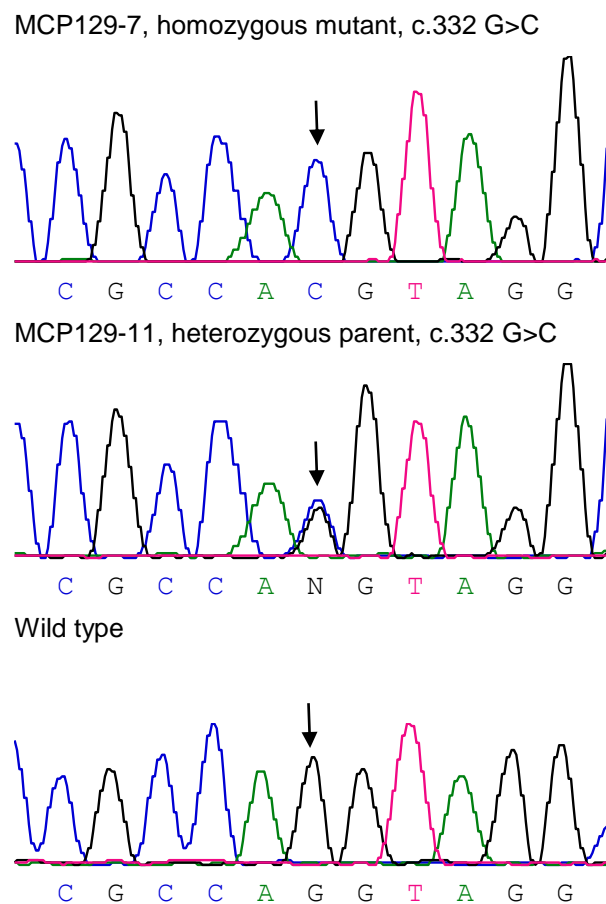
Wild type

Figure 6.1.16: Sanger traces for mutation c.332 G>C in exon 3 of *WDR62*. Chromatogram of representative homozygous affected individual (MCP129-7), heterozygous individual (MCP129-11) and homozygous individual for wild type allele. Arrow shows site of mutation.

6.1.4 MCPH3 linked families

MCPH3 locus was first mapped on chromosome region 9q34 in a large consanguineous Pakistani family. A ~12 cM region of homozygosity was identified defined by the markers cen-D9S1872- D9S159-tel (Moynihan et al. 2000). Two families MCP105 and MCP121 showed linkage to the MCPH3 locus.

6.1.4.1 Family MCP105

Family MCP105 is consanguineous family with two affected individuals. This family was recruited from District Khushab, Punjab, Pakistan. The pedigree (Fig 6.1.17) consists of two affected individuals that are born to first cousin parents. MCP105-1 aged 12 years with head circumference -8 SD while MCP105-2 aged 16 years with -9 SD. The patients had mild intellectual disability but had no other neurological symptoms or skeletal anomalies. All the available individuals from the family were genotyped using microsatellite markers around the first seven MCPH loci. The affected individuals demonstrated homozygosity at MCPH3 locus. The haplotype data are not shown here. To confirm further, both affected members were genotyped using Axiom ^{TM} Genome-wide Array (Affymetrix). Genome wide homozygosity mapped peaks on chromosomes 2, 9, 14, 17 and 20 with no statistically significant LOD scores. Next, to check if this finding is consistent with the linkage map of microsatellite markers, maximum markers on chromosome 9 were selected and calculated the multipoint LOD score (ALLEGRO) of z~1.5. *CDK5RAP2* was found to reside within the small linkage interval (122,801,186-124,152,939; hg19) on chromosome 9 (Fig 6.1.17). The haplotypes were constructed by inspection the homozygous markers that spans 1.3 Mb region (Fig 6.1.18).

To find the causal variant and exclude variants in other known MCPH genes, DNA sample of one patient MCP105-1 were subjected to whole-exome sequencing. Exome analysis detected a homozygous novel nonsense mutation p.Arg1372* in *CDK5RAP2* at nucleotide position c.4114C>G (ENST00000349780) in exon 27 (Fig 6.1.23). This variant was corroborated by Sanger sequencing and segregation of the disease allele in all available individuals of the family. The parents were found to be heterozygous for the mutation and patients were homozygous (Fig 6.1.19).

## Family MCP105

(a)

(b)
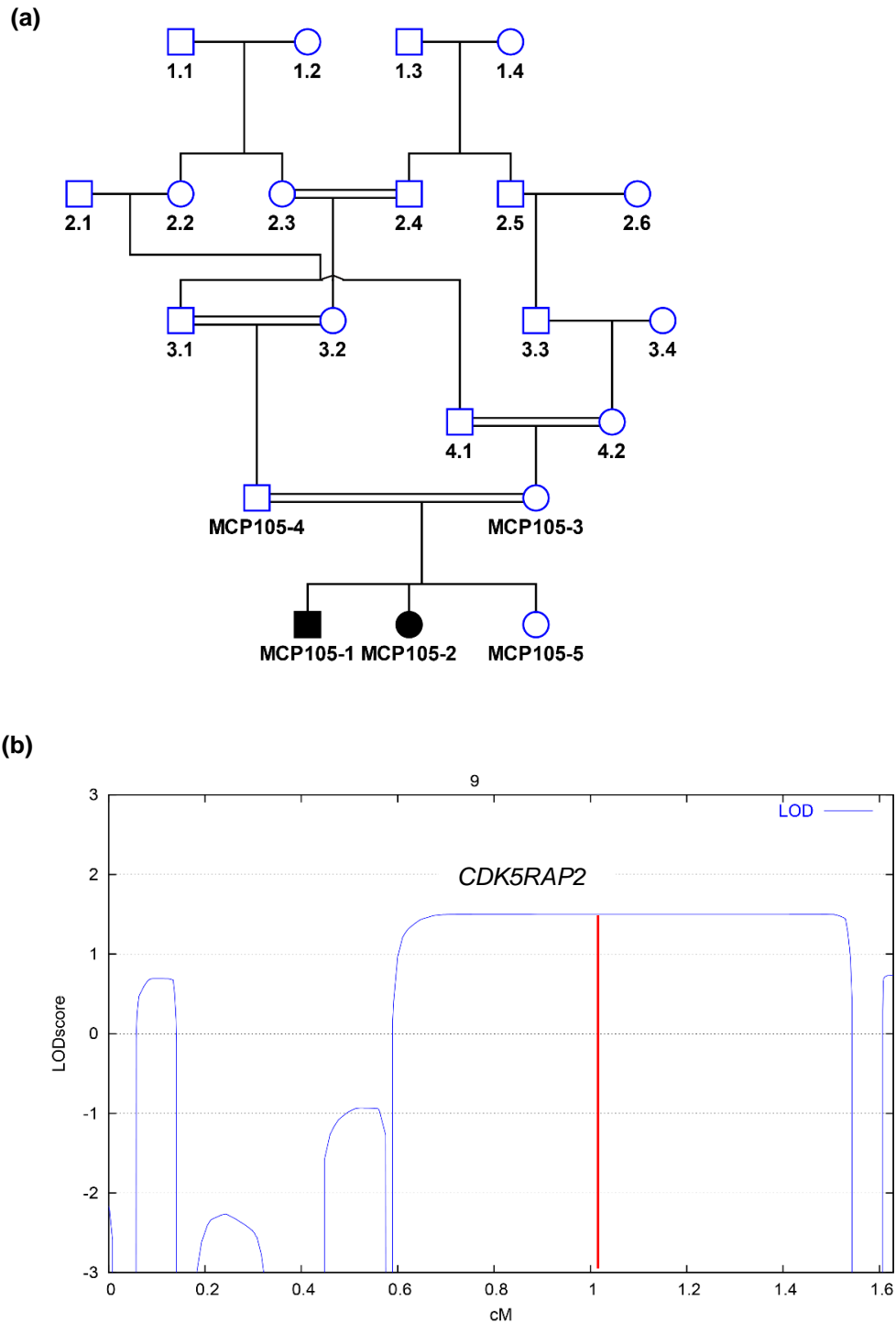
Figure 6.1.17: (a) Pedigree of family 105 with autosomal recessive primary microcephaly. **(b)** Results of parametric linkage analysis showing LOD scores on chromosome 9 with the position of *CDK5RAP2*, shown as vertical red line. The coordinate of 0 (cM) in the figure corresponds to the 126.8 cM.
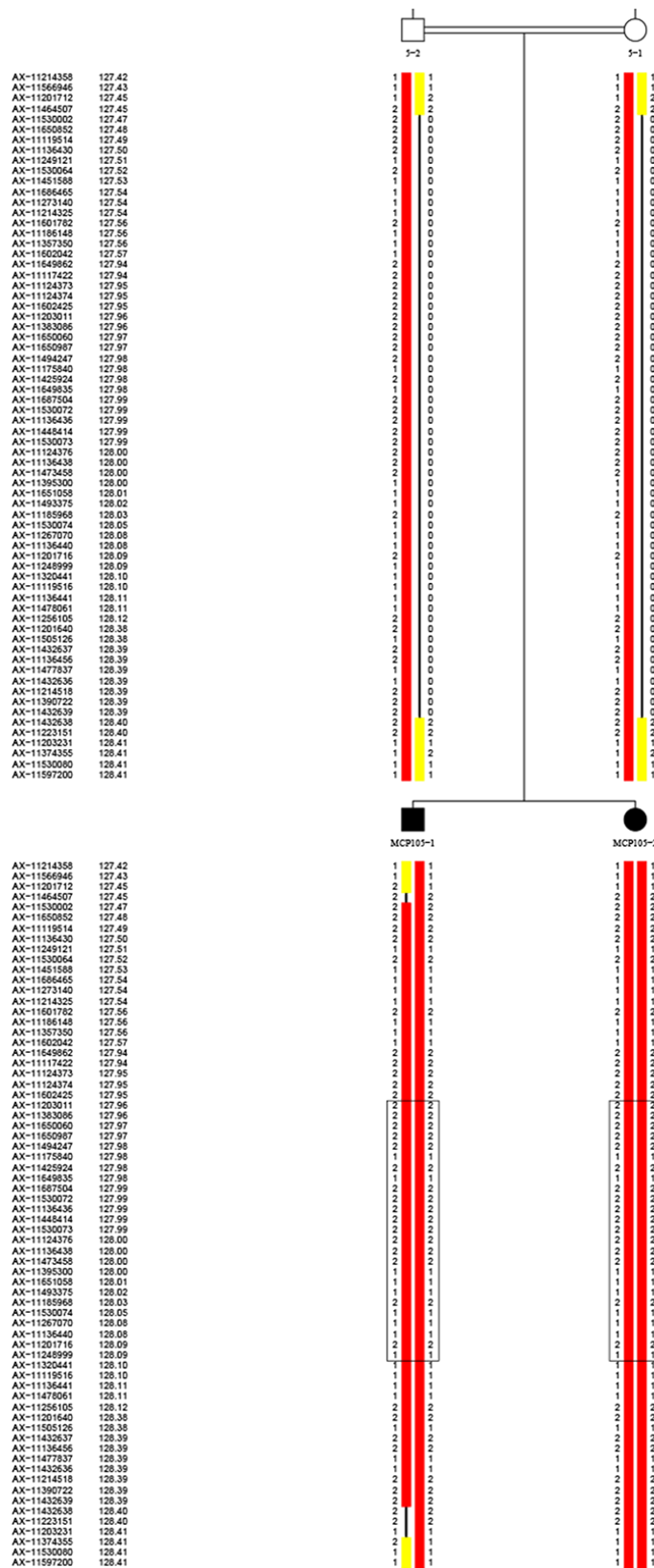
Figure 6.1.18: Haplotypes were constructed by inspection from the genotyping of Axiom ™ Genome-wide Array (Affymetrix). The flanking markers AX-11464507 (position: 122,801,186 bp; 127.4 cM; hg19) and AX-11432638 (position: 124,152,939 bp; 128.4 cM; hg19) spans 1.3 Mb homozygous region. 25 homozygous markers are located within *CDK5RAP2* which is shown as box.
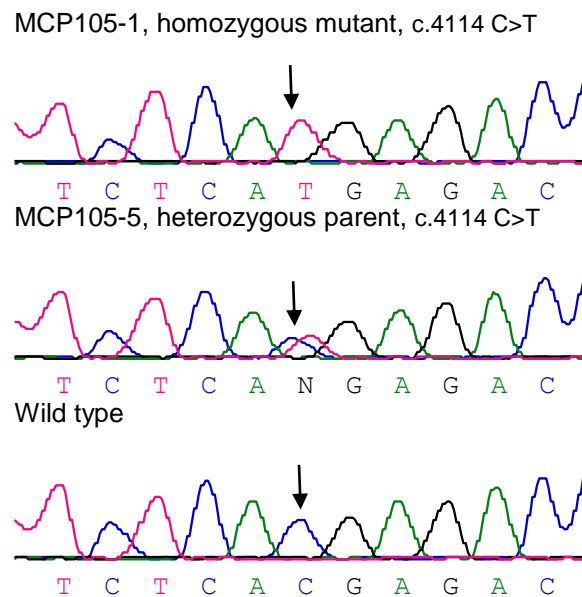
Figure 6.1.19: Sanger traces for mutation c.4114 C>T in exon 27 of *CDK5RAP2*. Chromatogram of representative homozygous affected individual (MCP105-1), heterozygous individual (MCP105-5) and homozygous individual for wild type allele. Arrow shows site of mutant nucleotide.
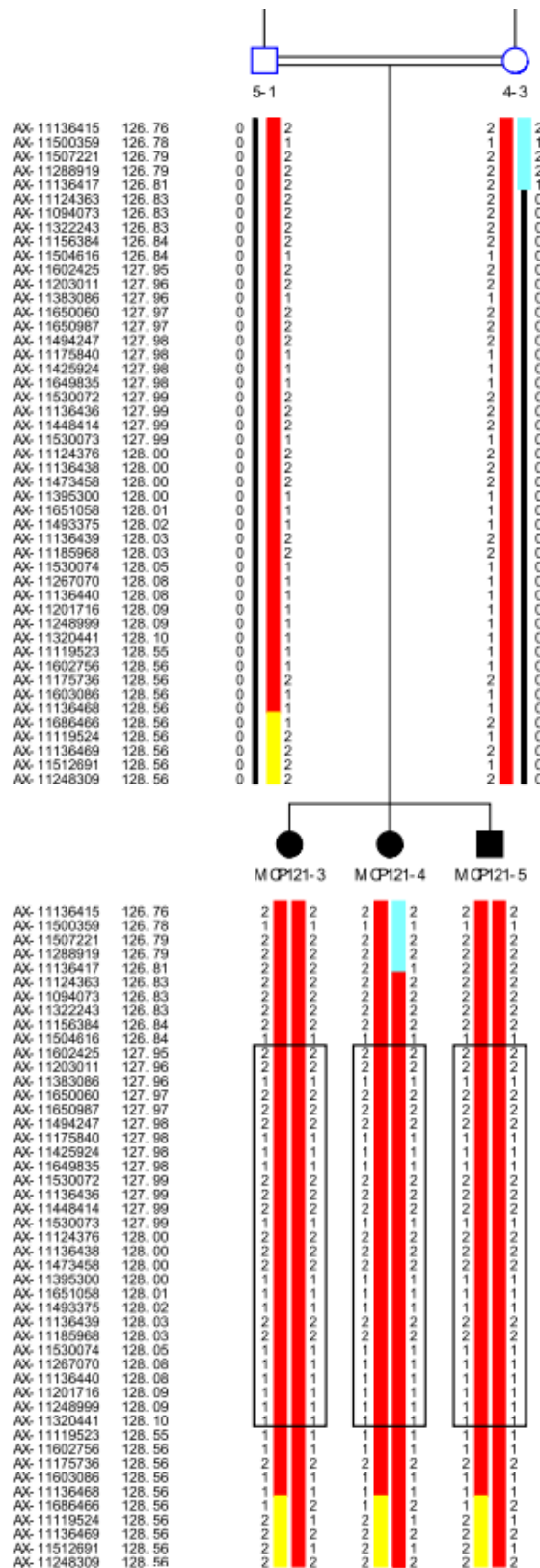
6.1.4.1 Family MCP121

Family MCP121 is a consanguineous family with three affected individuals (Fig 6.1.20-A). This family was recruited from District Tando Muhammad Khan, Sindh, Pakistan. The microcephaly was noted at birth, and head circumference is ranged from -5 SD to -7 SD at the time of sampling. The patients had mild intellectual disability without any facial dysmorphism or other neurological deficit. The age of the patients varied between 7 to 17 years.

All affected cases were genotyped using the Axiom ^{TM} Genome-wide Array (Affymetrix). Parametric linkage analysis was performed for recessive mode of inheritance with 20K SNPs and detected peaks on chromosome 9 and 21 with maximum LOD scores of z~3.3 (Fig 6.1.20-b). The underlying homozygous regions spanned 2.1 Mb (122,473,586-124,647,279 bp; hg19) on chromosome 9 and 8.0 Mb on chromosome 21 (10,727,679-18,743,733 bp; hg19). The candidate region on chromsosme 9 harbours a *CDK5RAP2*, a gene already described in patients with primary microcephaly (Fig 6.1.20-c,Fig 6.1.21). The identification of the disease locus was followed by whole-exome sequencing of one affected (MCP121-4) and filtering rare variants in known MCPH genes. A homozygous novel nonsense mutation was identified in *CDK5RAP2* at nucleotide position c.1279 in exon 12 (c.1279C>T, ENST00000349780) (Fig 6.1.23). The C-to-T transition is predicted to result in a premature termination codon (p.Arg427*). Sanger sequencing validated the presence of the homozygous mutation in the two other affected siblings and a heterozygous state for the parents and two healthy sibs (Fig 6.1.22).

Family MCP121

(a)

(b)

(c)
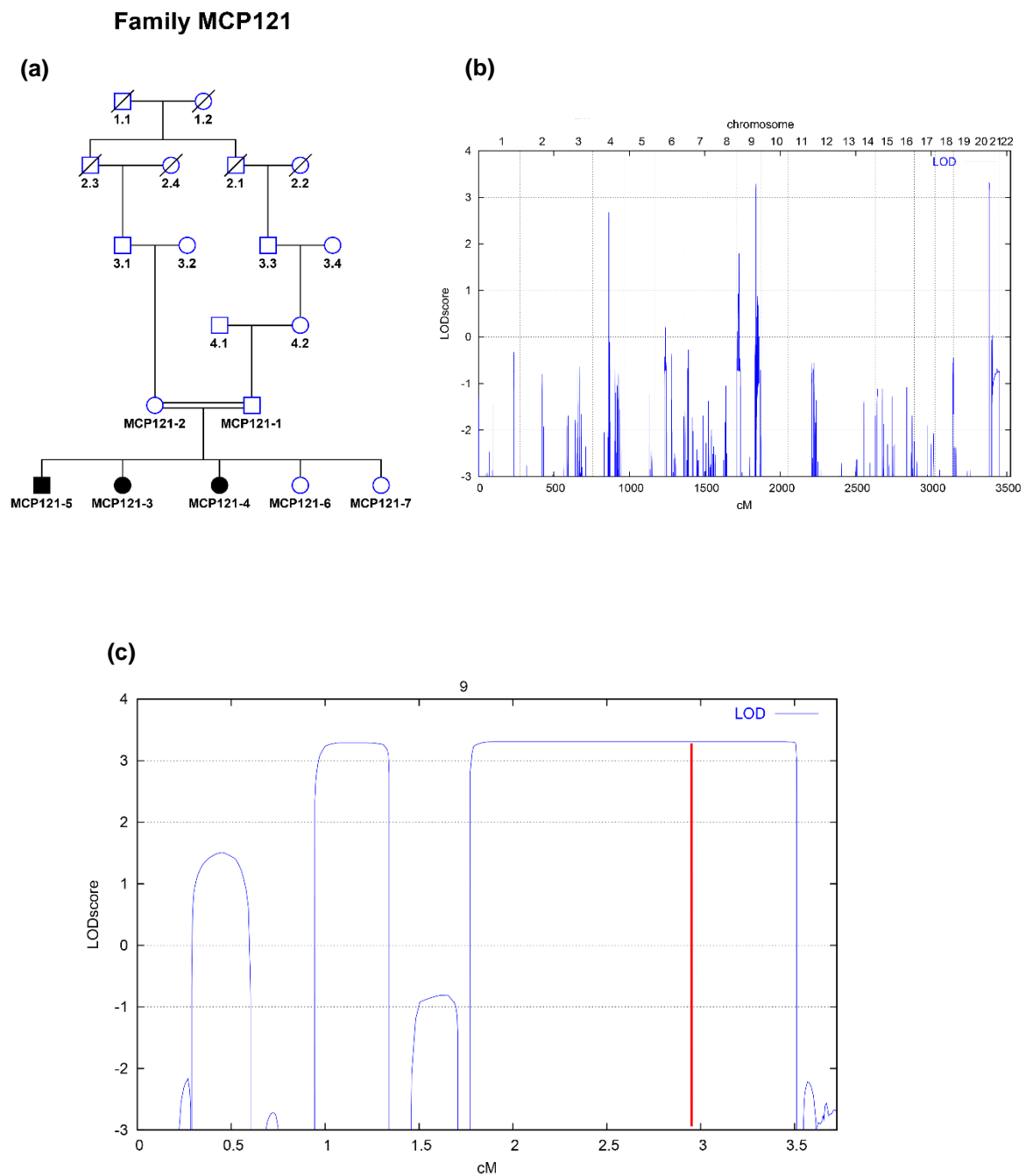
Figure 6.1.20: (a) Pedigree of family MCP121 with autosomal recessive primary microcephaly. Filled symbols represent affected individuals. Empty symbols denote non affected individuals, and slashed symbols represent deceased individuals. **(b)** Genome wide scan shows two peaks on chromosome 9 and 21, with maximum LOD scores of z ~3.3. **(c)** Results of parametric linkage analysis showing LOD scores on chromosome 9 with the position of *CDK5RAP2*, shown as vertical red line. The coordinate of 0 (cM) in the figure corresponds to the 125.0 cM.

Figure 6.1.21: Haplotype reconstruction from the genotyping of three affected individuals of family MCP121 (Axiom TM Genome-wide Array (Affymetrix). AX-11136417 (Position: 122,473,586 bp; 126.8 cM; hg19) and AX-11686466 (position: 124,647,279 bp; 128.5 cM; hg 19) markers are the flanking markers for the shared homozygous region. 25 markers are located within *CDK5RAP2*, shown as enclosed box.

MCP121-1, homozygous mutant, c.1279C>T

MCP121-2, heterozygous parent, c.1279C>T
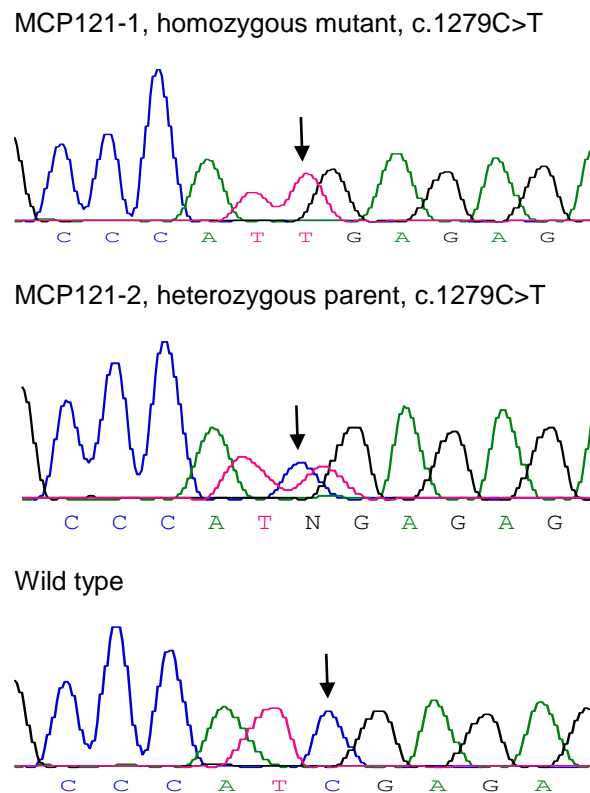
Wild type

Figure 6.1.22: Sanger traces for mutation c.1279 C>T in exon 12 of *CDK5RAP2.* Chromatogram of representative homozygous affected individual (MCP121-5), heterozygous individual (MCP121-2) and homozygous individual for wild type allele. Arrow shows site of mutation.
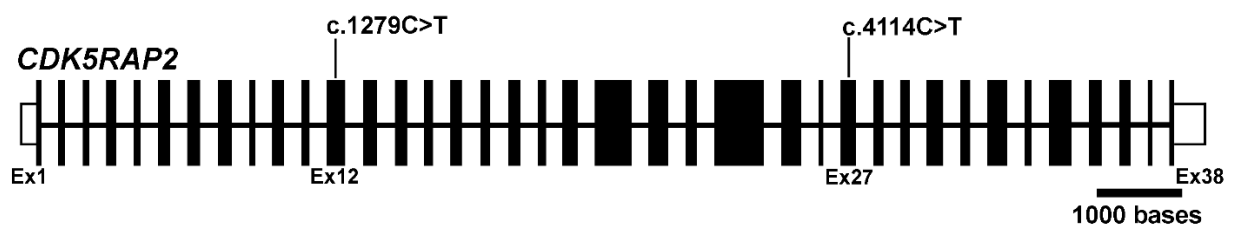

Figure 6.1.23: Schematic representation of the exon/intron structure of *CDK5RAP2*. Exons (vertical lines) are drawn according to scale whereas introns (straight lines) are arbitrary. Both nonsense mutations in exon 12 and 27 of the *CDK5RAP2*, respectively, are shown. These are novel mutations identified in the current study.

4.1.5 MCPH5 linked families

MCPH5 is the fifth locus that was mapped at 1q25-q32 by homozygosity mapping in a Turkish family (Jamieson et al. 2000). *ASPM* (abnormal spindle–like microcephaly associated) is the mutated gene at this locus (Bond et al. 2002). In total, 9 families demonstrated linkage to the MCPH5 locus through homozygosity mapping using polymorphic microsatellite markers and SNP arrays. Multipoint linkage analysis (ALLEGRO) was used to calculate LOD scores. These families were further subjected to Sanger sequencing.

6.1.5.1 Sanger sequencing of *ASPM*

ASPM contains 28 exons with an open reading frame of 10,434 bp. All coding exons and exon-intron boundaries of *ASPM* (ENST00000367409) were bidirectionally sequenced in the index patient of each linked family. In addition, targeted Sanger sequencing were performed in 14 MCPH families with no prior linkage mapping (appendix 6). These families were ascertained largely from North-West part of Pakistan. Taken together, Sanger sequencing revealed nine novel and three known mutations in 23 families (Fig 6.1.24). Interestingly, 13 families carried the already reported founder mutation p.Trp1326* in *ASPM* (Table 6.1.2). Moreover, two families harbored compound heterozygous novel mutations (Fig 6.1.24 and Fig 6.1.34). Each of these mutations segregated faithfully within the family. All mutations identified in *ASPM* are summarized in Table 6.1.2.
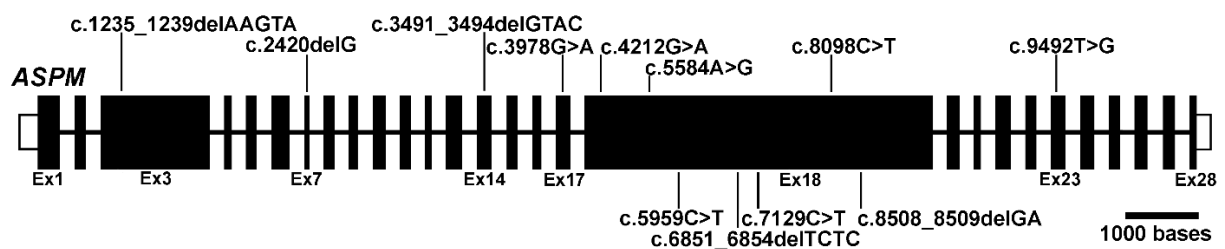
Figure 6.1.24: Graphical representation of *ASPM*. Mutation identified in the current study are shown here. All homozygous mutations are shown above the scheme and below the compound heterozygous mutations.

Family MCP101
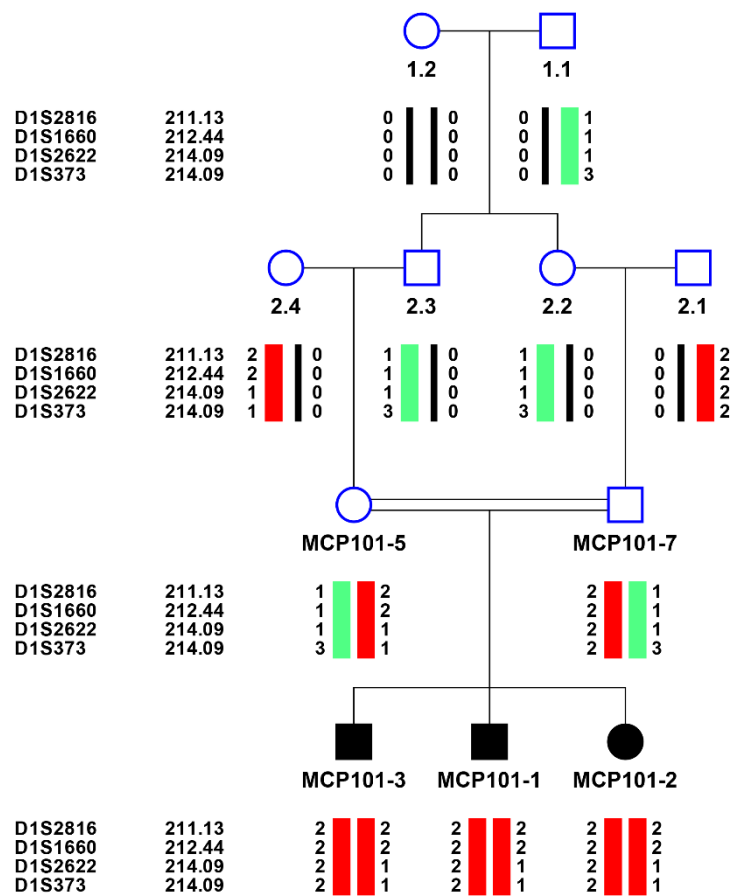
Figure 6.1.25: Pedigree with haplotypes of the MCPH5 locus for the family MCP101. The affected individuals showed homozygosity at MCPH5 locus at chromosome 1. *ASPM* resides in the underlying homozygous region (chr1: 196,650,545-201,391,058 bp; hg19).

Family MCP106

(a)

(b)
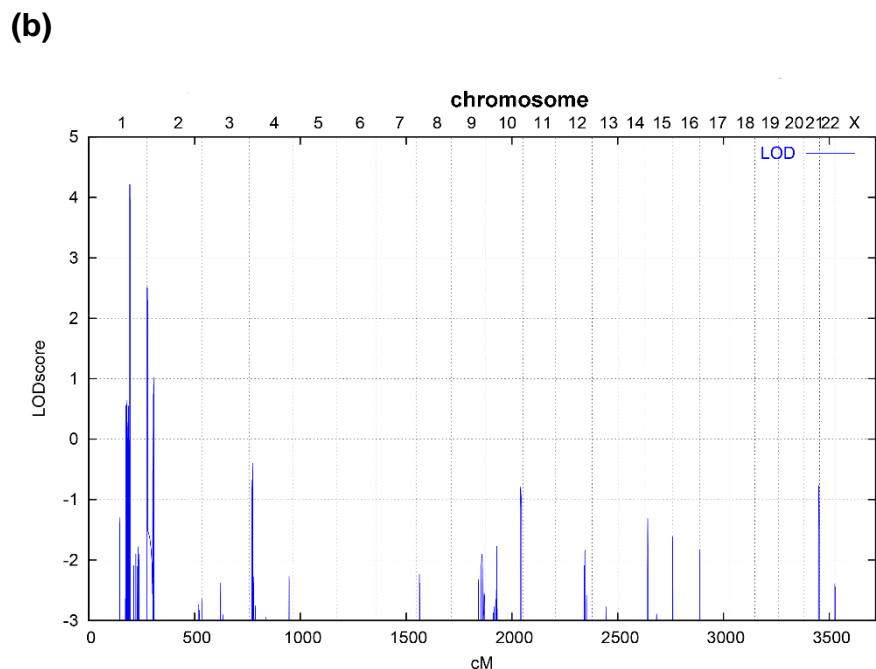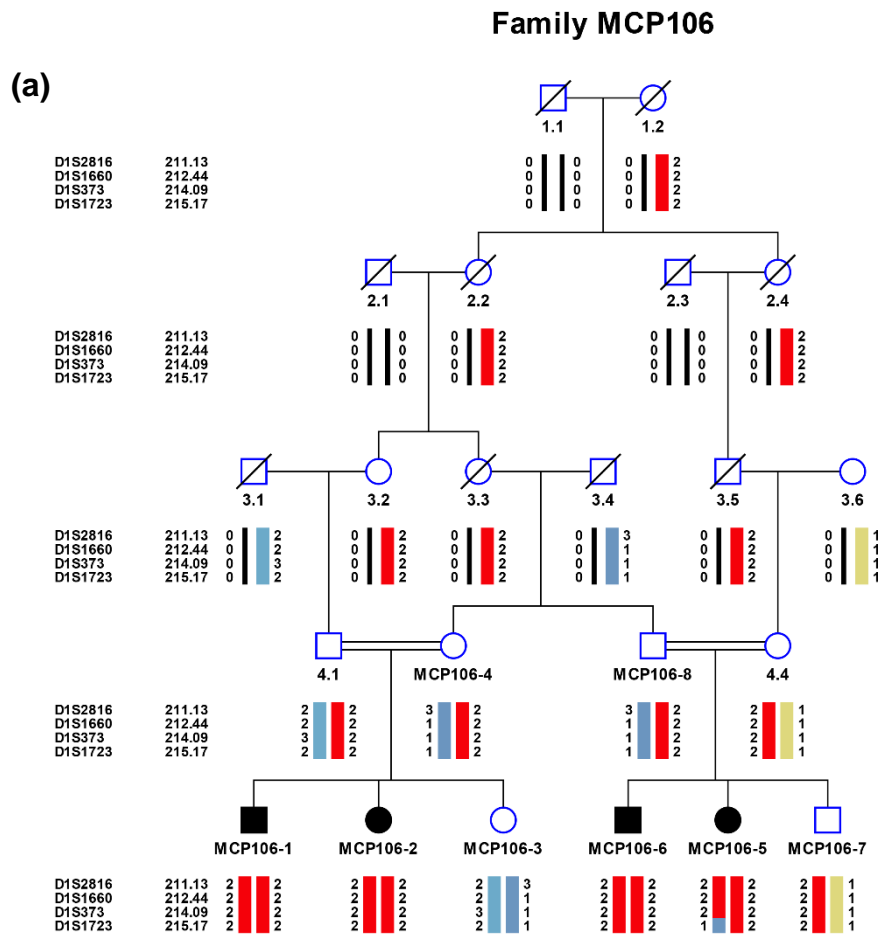
Figure 6.1.26: (a) Simplified pedigree with haplotypes of the MCPH5 locus for the family MCP106. Haplotypes was constructed by using algorithm in Allegro v1.2c and was plotted using Haplopainter. Marker names and position on genetic map are displayed on the left side of the pedigree. **(b)** Array-based mapping (Affymetrix Axiom genome-wide array) was also performed for genome wide linkage analysis. Parametric linkage analysis demonstrated a peak on chromosome 1 with the LOD score of z=4.21.The underlying homozygous region of 4.3 Mb (Chr1: 195,989,351-200,303,492 bp; hg19)

contains *ASPM*. Only three STR markers are included in the linkage region of the array based linkage map.
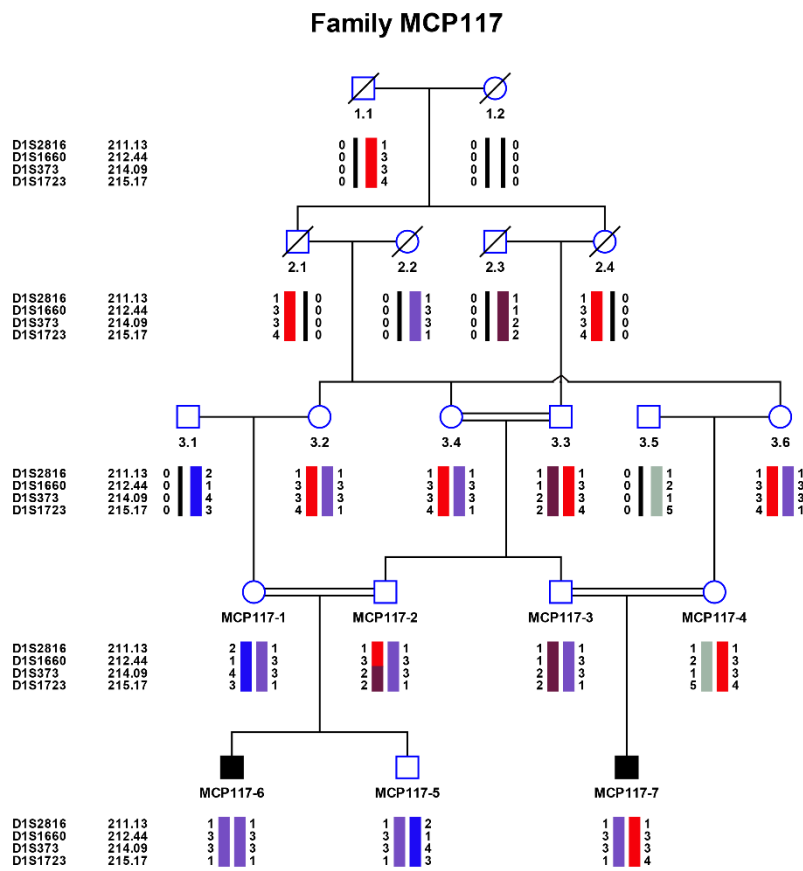
Figure 6.1.27: Pedigree with haplotypes of the MCPH5 locus for the family MCP117. Haplotypes was constructed by using algorithm in Allegro and was plotted using Haplopainter. Marker names and position on genetic map are displayed on the left side of the pedigree. The underlying homozygous region (Chr1: 195,989,351-200,303,492 bp; hg19) contains *ASPM*.

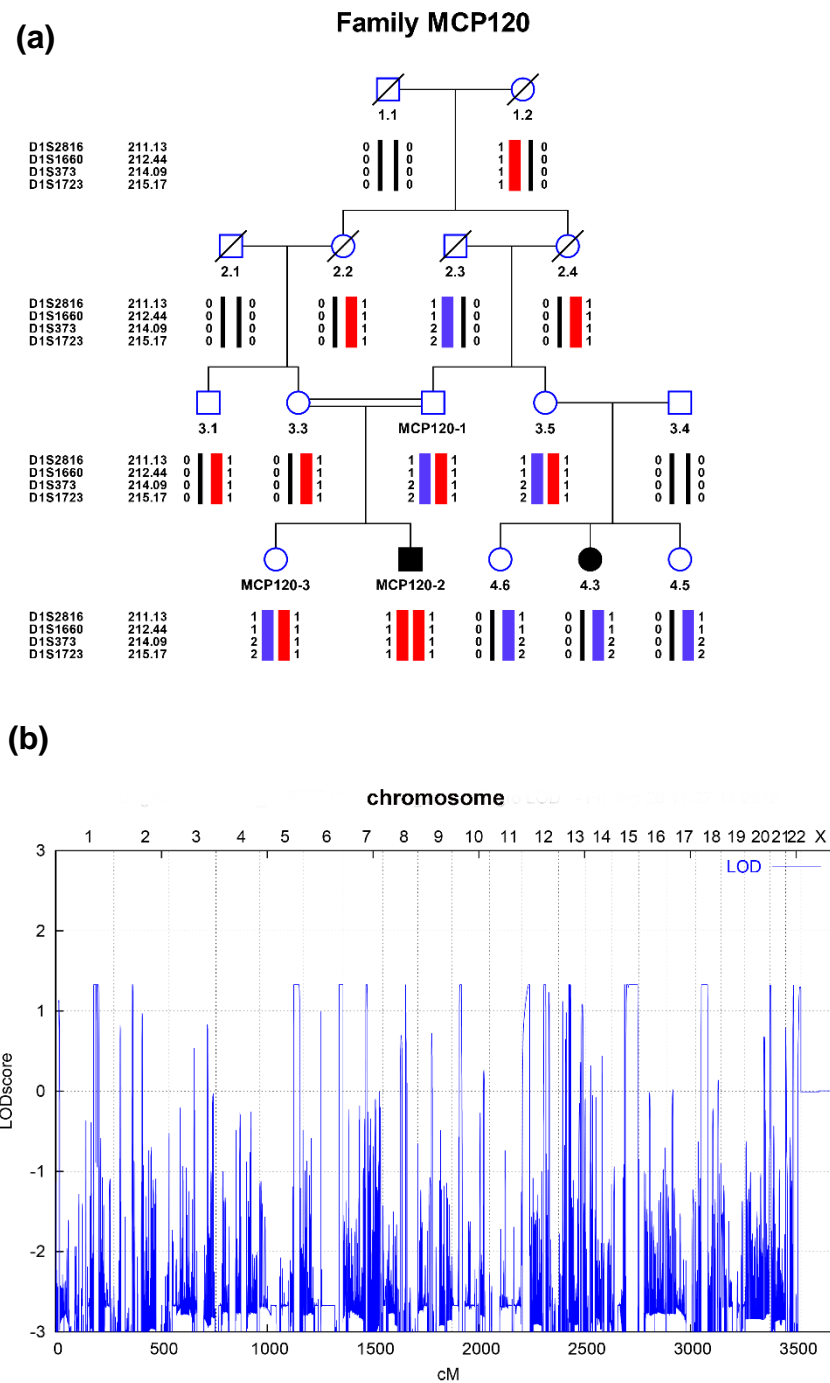**Figure 6.1.28 :(a)** Pedigree with haplotypes of the MCPH5 locus for the family MCP120. Individuals MCP120-1, MCP120-2 and MCP120-3 were genotyped using STR marker. The affected individual demonstrates homozygosity at MCPH5 locus. **(b)** Array based mapping was also performed in the same family and the parametric linkage analysis resulted in 16 peaks with maximum LOD score (ALLEGRO) of z~1.32. *ASPM* resides within the linkage interval on chromosome1 (178,034,744-203,596,318 bp; hg19).Affymetrix Axiom^{TM} genome-wide array was used for genotyping.

(a)

(b)
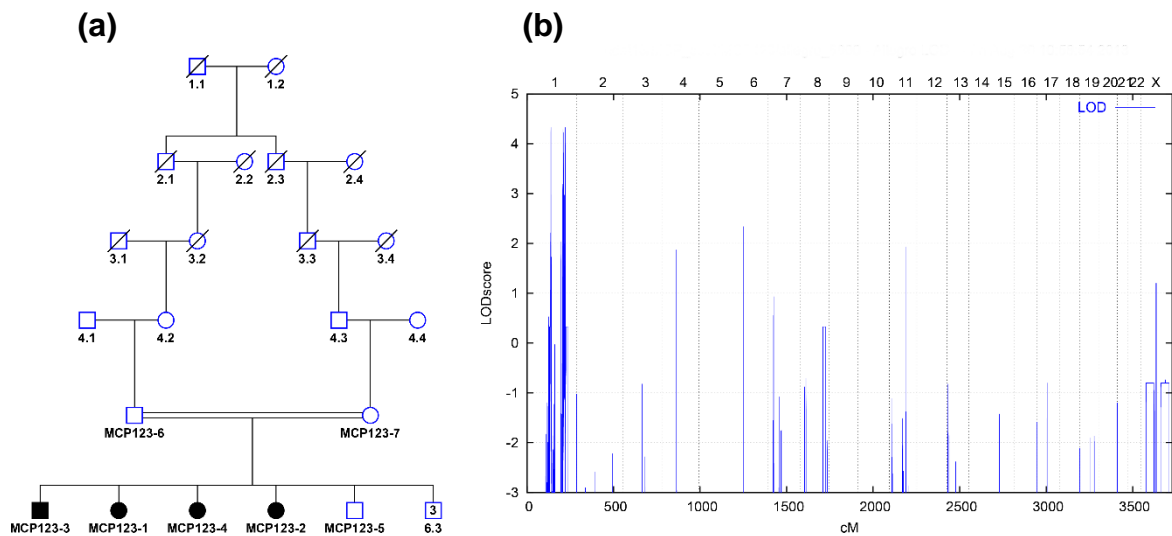
Figure 6.1.29: (a) Pedigree and **(b)** graphical view of the genome-wide linkage data of family MCP123. Parametric linkage analysis resulted in a maximum LOD score of z~4.3 on chromosome 1. Four small linkage regions were detected on chromosome 1. Four small linkage regions were detected on chromosome 1assuming homozygosity for the ASPM locus. *ASPM* resides out of the linkage regions. Interestingly, compound-heterozygou mutations were identified in affected individuals after sequencing of *ASPM*. Illumina HumanCoreExome-12 v1.1 Bead Chip was used for genotyping.

(a)

(b)
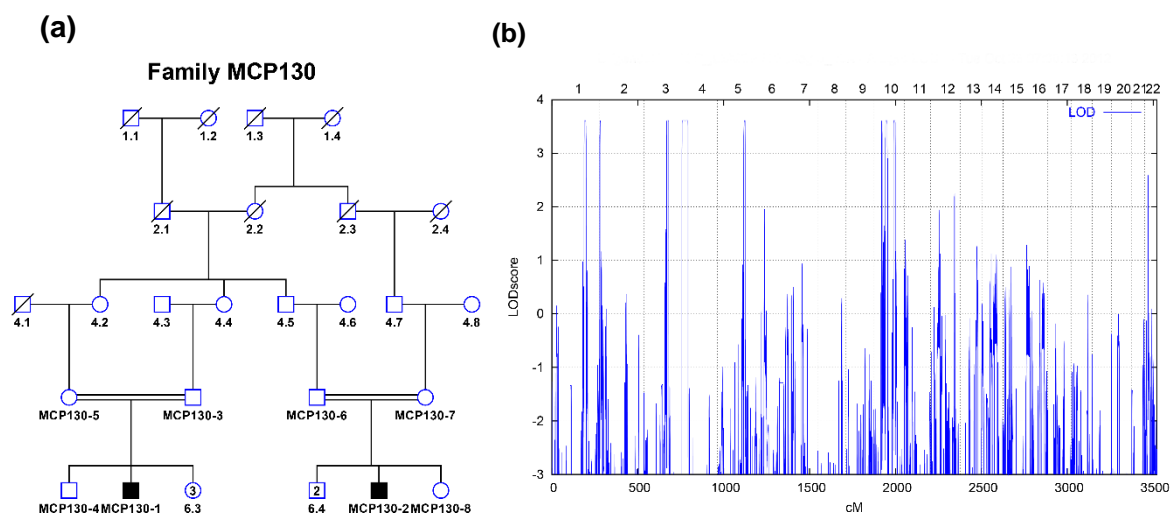
Figure 6.1.30: (a) Pedigree of family MCP130 **(b)** Genome wide scan based on the genotypes of two affected individuals. Parametric score with recessive mode of inheritance was calculated and detected peaks on chromosomes 1, 2,3,4,5 and 10 with maximum values of 3.6. *ASPM* resides within the linkage interval (Chr1:188,707,774-201,392,063 bp; hg19). Affymetrix Axiom TM genome-wide array was used for genotyping.
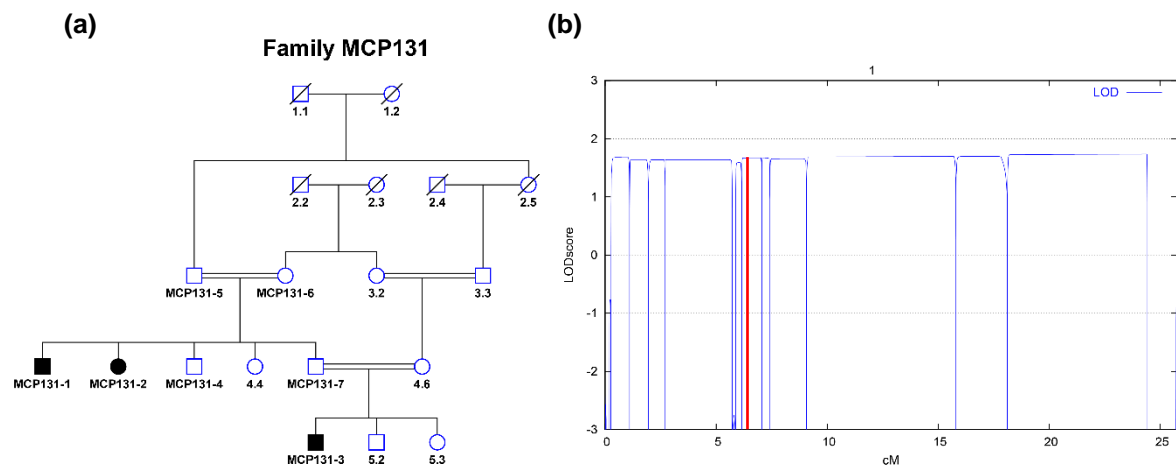
Figure 6.1.31: (a) Pedigree of family 131 with autosomal recessive primary microcephaly. **(b)** Results of parametric linkage analysis showing LOD scores (ALLEGRO) on chromosome 1 with the position of *ASPM,* shown as vertical red line. The coordinate of 0 (cM) in the figure corresponds to the 187.5 cM. Affymetrix Axiom TM genome-wide array was used for genotyping.
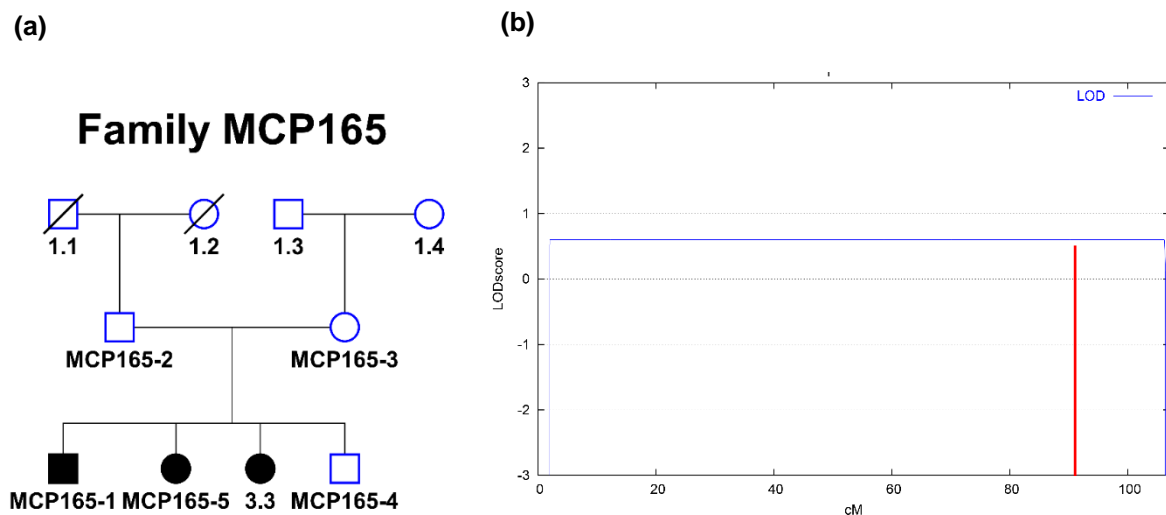
Figure 6.1.32: (a) Pedigree of family 165 with autosomal recessive primary microcephaly. **(b)** Results of parametric linkage analysis showing maximum theoretical LOD scores z=0.5 on chromosome 1 with the position of *ASPM*, shown as vertical red line. The coordinate of 0 (cM) in the figure corresponds to the 116.4 cM. Illumina HumanCoreExome-12 v1.1 BeadChip was used for genotyping.

(a)
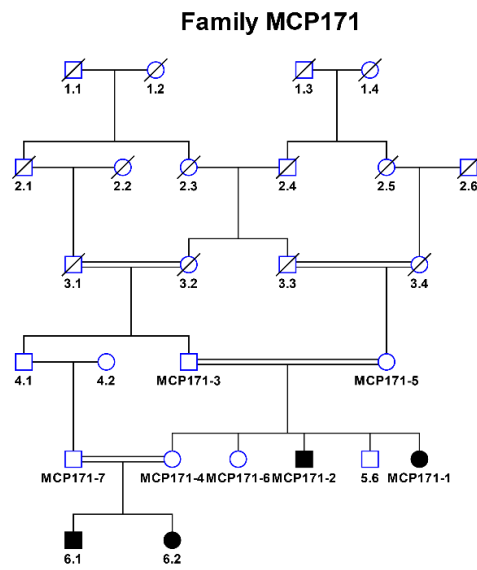
Family MCP171

(b)
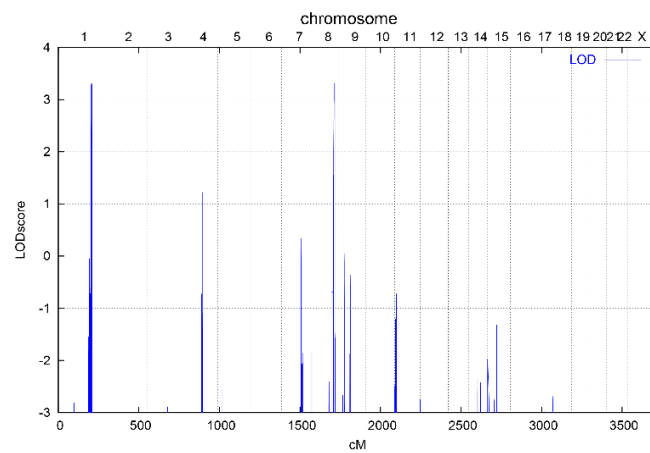
Figure 6.1.33: (a) Pedigree of the family MCP171. **(b)** Genome wide scan of multipoint parametric linkage analysis (ALLEGRO) that demonstrated peaks on chromosomes 1 and 8 with maximum LOD score of z ~3.3. *ASPM* resides within the linkage interval on chromosome 1 (193,038,360-199, 997,778 bp; hg19). Illumina HumanCoreExome-12 v1.1 BeadChip was used for genotyping.

Table 6.1.2: List of mutations identified in *ASPM*

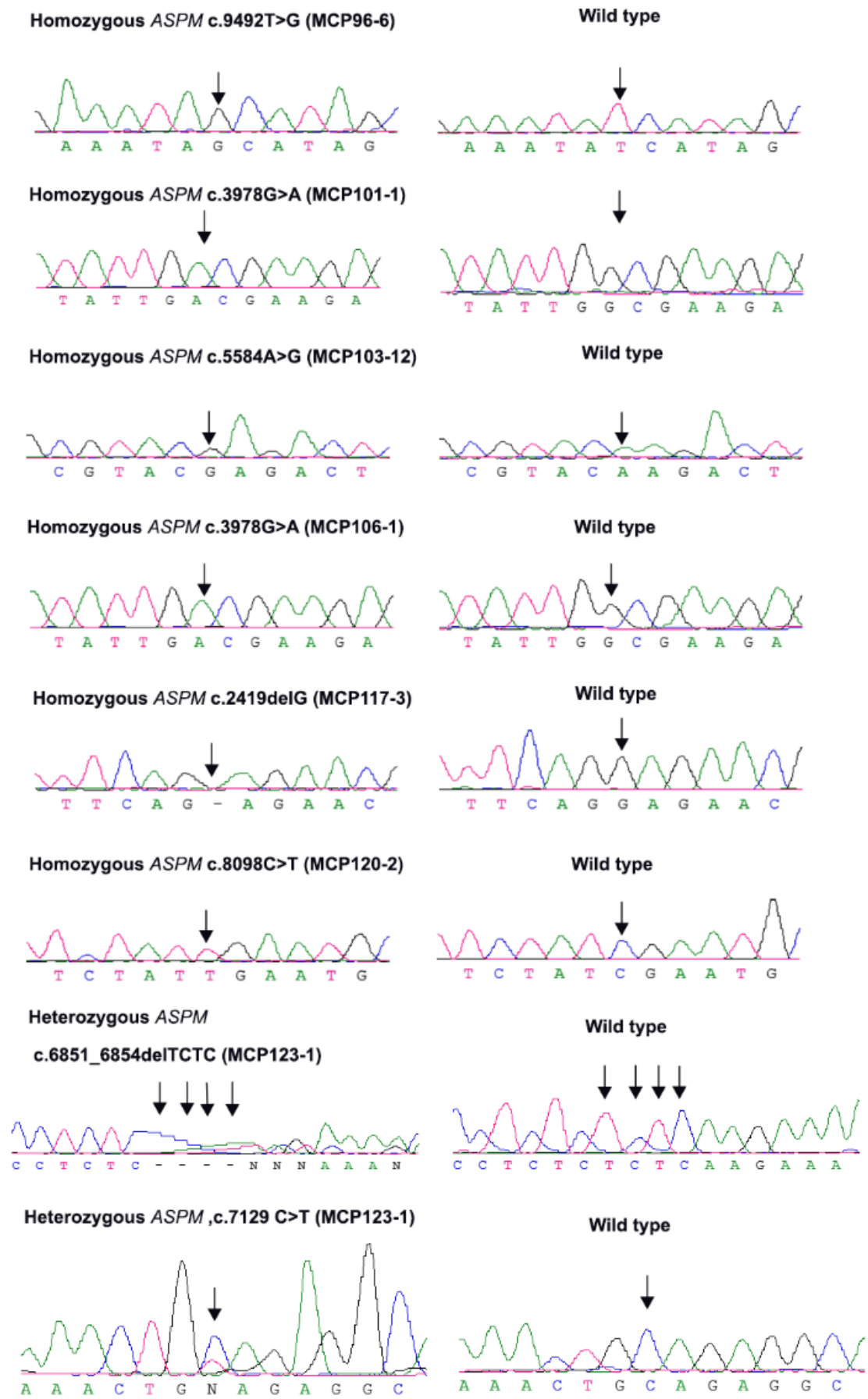| S. No	Family ID	Gene	Exon	cDNA mutation	Protein Mutation	Status
1	MCP96	*ASPM*	23	c.9492T>G	p.Tyr3164*	known
2	MCP101*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
3	MCP103*	*ASPM*	18	c.5584A>G	p.Lys1862Glu	known
4	MCP106*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
5	MCP117*	*ASPM*	7	c. 2419del G	p.Gly807Glu fs*7	novel
6	MCP120	*ASPM*	18	c.8098C>T	p.Arg2700*	novel
7	MCP123	*ASPM*	18	c.6851_6854delTCTC c.7129C>T	p.Leu2285Argfs*6 p.Gln2377*	novel novel
8	MCP130	*ASPM*	18	c.4212G>A	p. Trp1404*	novel
9	MCP131*	*ASPM*	3	c.1232_1236delGTAAA	p.Lys412Thrfs*5	novel
10	MCP135*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
11	MCP136*	*ASPM*	14	c.3491-3494delGTAC	p.Arg3491Leufs*15	novel
12	MCP137*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
13	MCP138*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
14	MCP139*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
15	MCP140*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
16	MCP141*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
17	MCP143*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
18	MCP165	*ASPM*	18	c.8508_8509delGA c.5959C>T	p.Lys2837Metfs*34 p.Gln1987*	novel novel
19	MCP171	*ASPM*	18	c.5959C>T	p.Gln1987*	novel
20	MCP174*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
21	MCP175*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
22	MCP176*	*ASPM*	17	c.3978G>A	p.Trp1326*	known
23	MCP177*	*ASPM*	17	c.3978G>A	p.Trp1326*	known

Families marked with asterisk* had Pashtun ethnicity including MCP137 and MCP138, originate from Afghanistan and currently living as refugees in Khyber Pakhtunkhwa, province in Pakistan.

Homozygous *ASPM* c.9492T>G (MCP96-6)

Wild type

A A A T A G C A T A G

A A A T A T C A T A G

Homozygous *ASPM* c.3978G>A (MCP101-1)

T A T T G A C G A A G A

T A T T G G C G A A G A

Homozygous *ASPM* c.5584A>G (MCP103-12)

Wild type

C G T A C G A G A C T

C G T A C A A G A C T

Homozygous *ASPM* c.3978G>A (MCP106-1)

Wild type

T A T T G A C G A A G A

T A T T G G C G A A G A

Homozygous *ASPM* c.2419delG (MCP117-3)

Wild type

T T C A G - A G A A C

T T C A G G A G A A C

Homozygous *ASPM* c.8098C>T (MCP120-2)

Wild type

T C T A T T G A A T G

T C T A T C G A A T G

Heterozygous *ASPM*

c.6851_6854delTCTC (MCP123-1)

Wild type

C C T C T C - - - - N N N A A A N

C C T C T C T C T C A A G A A A

Heterozygous *ASPM* ,c.7129 C>T (MCP123-1)

Wild type

A A A C T G N A G A G G C

A A A C T G C A G A G G C

73

Homozygous *ASPM* c.4212G>A (MCP130-1)

C A T T G A C G T G C

Wild type

C A T T G G C G T G C
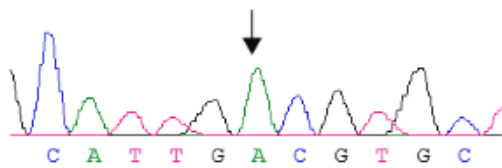
Homozygous *ASPM*

c.1235_1239del AAGTA(MCP131-3)

T G T A C - - - - - C A T T A T

T G T A A A G T A C C A T T A T

Homozygous *ASPM* c. 3978G>A (MCP135-1)

Wild type

T A T T G A C G A A G A

T A T T G G C G A A G A
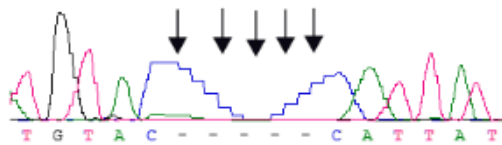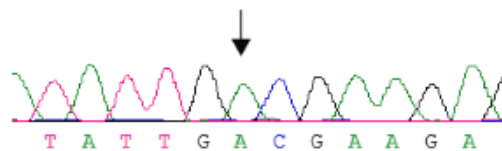
Homozygous *ASPM* c.3491_3494delGTAC (MCP136-2)

Wild type

G T C A G - - - - C T A C T

G T C A G C G T A C T A C T

Homozygous *ASPM* c.3978G>A (MCP137-11)

Wild type

T A T T G A C G A A G A

T A T T G G C G A A G A
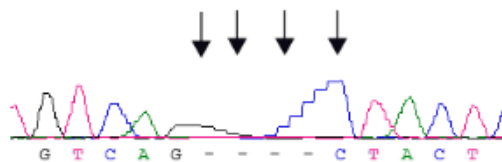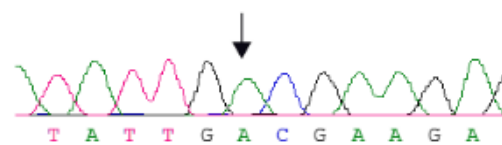
Homozygous *ASPM* c.3978G>A (MCP138-3)

Wild type

T A T T G A C G A A G A

T A T T G G C G A A G A

Wild type

Homozygous *ASPM* c.3978G>A (MCP139-6)

T A T T G A C G A A G A

T A T T G G C G A A G
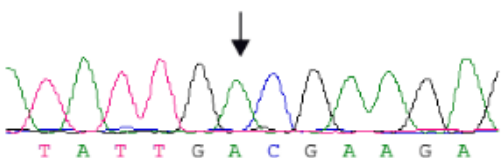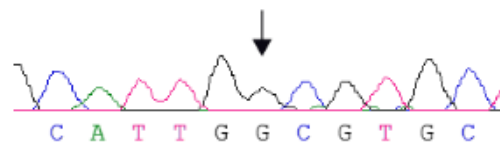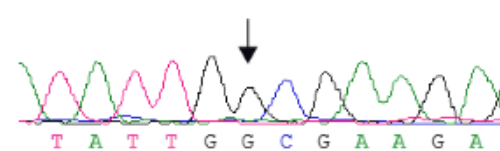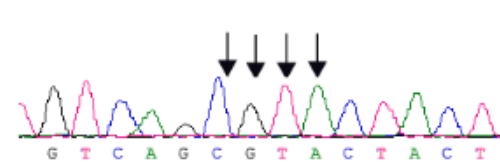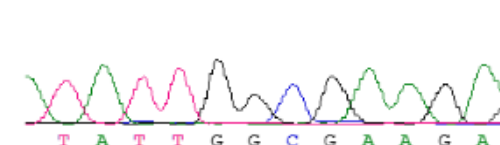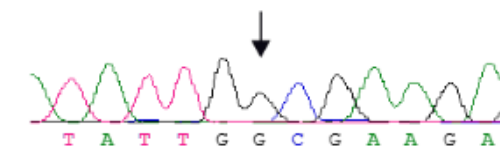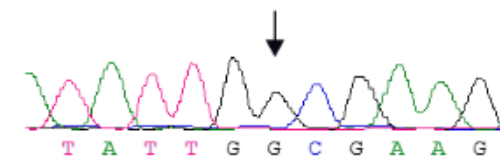
Homozygous *ASPM* c.3978G>A (MCP140-1)
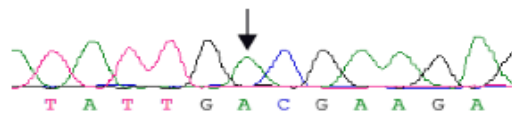
Wild type

Homozygous *ASPM* c.3978G>A (MCP141-3)

Wild type
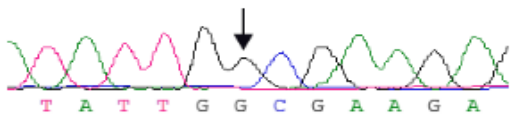
Homozygous *ASPM* c.3978G>A (MCP143-8)
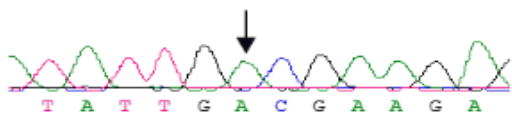
Wild type
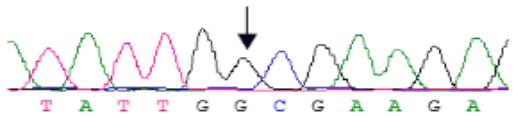
Heterozygous (MCP165-1)

ASPM c.8508_8509delGA
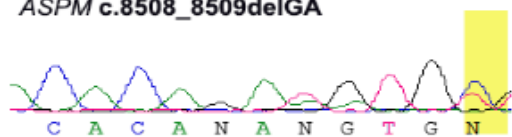
Wild type
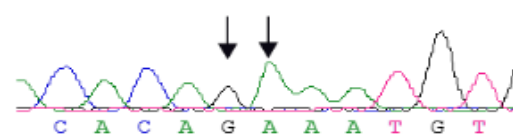
ASPM c.5959C>T (MCP165-1)
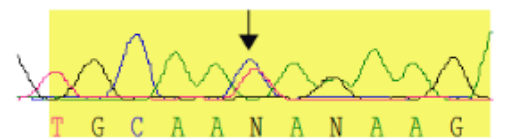
Wild type
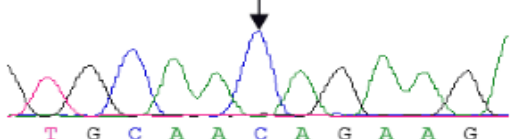
Homozygous *ASPM* c.5959C>T (MCP171-4)
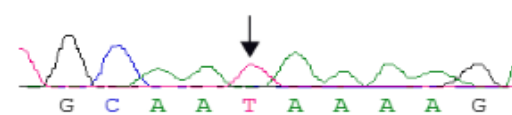
Wild type
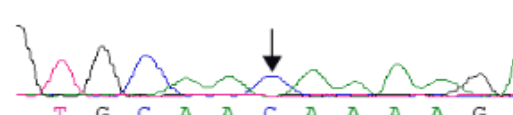
Homozygous *ASPM* c.3978G>A (MCP174-2)
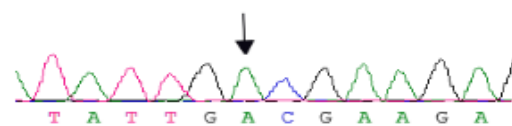
Wild type
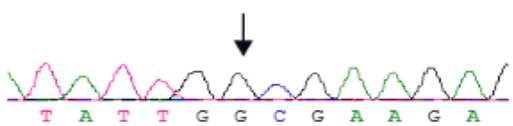
Homozygous *ASPM* c.3978G>A (MCP175-1)
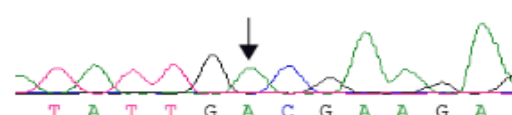
Wild type
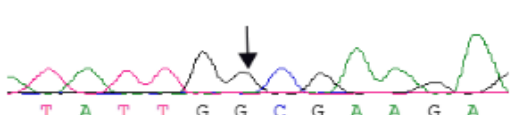
Homozygous *ASPM* c.3978G>A (MCP176-2)

Wild type
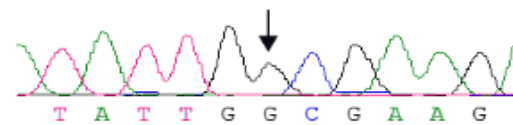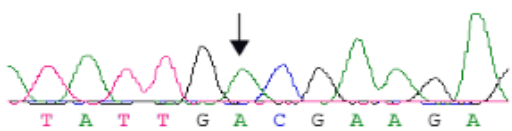
75
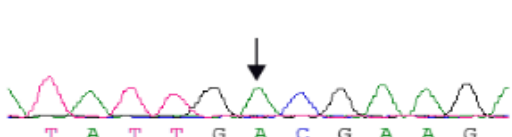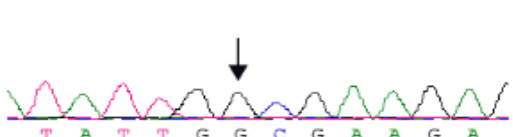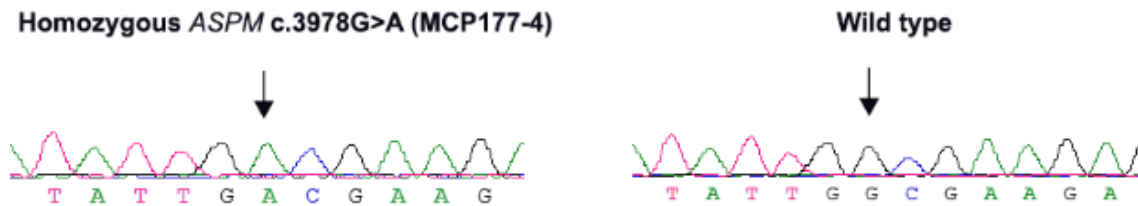
Figure 6.1.34: Sanger traces of families with *ASPM* mutations. Representative chromatogram shows different mutations. The mutant trances are displayed at the left while wild type (references sequence) is on the right. Black arrows indicate site of mutation.

6.1.6 Excluded families

PCR-based genotyping and linkage analysis through different arrays resulted in the exclusion of one family.

6.1.6.1 MCP104

Family MCP104 (Fig 6.1.35) is a consanguineous family ascertained from the Punjab province, Pakistan. This family has three affected individuals (2 males and 1 female). Both the parents are unaffected, and inheritance in this family was suggested to be autosomal recessive. Head circumference among affected individual was -7 to -9 SD below the population mean. The patients had mild intellectual disability and good sense of self care skills. There were no skeleton anomalies and no history of seizures or epilepsy. As a first investigation, the family was tested for linkage to the first seven MCPH loci using polymorphic microsatellite markers. No MCPH locus was found to have homozygous haplotypes for the affected individuals. We next performed Whole-genome linkage analysis by Illumina HumanCoreExome-12 v1.1 Bead Chip. DNA samples from two affected, both parents and one healthy sib were genotyped. Parametric linkage analysis was performed with reduced panel of 25 K SNPs and mapped the peaks on chromosome 3, 7, 15 and 17 with maximum possible LOD score of z ~2.23 (Fig 6.1.36). These linkage regions defined a candidate region of 2.5 Mb on chromosome 3 (63,340,627–65,817,691 bp; hg19), a candidate region of 6.0 Mb on chromosome 7 (149,630,727–155,664,686 bp; hg19), a candidate region of 2.2 Mb on chromosome 15 (23,236,972–25,413,558 bp; hg19) and a candidate region of 3.9 Mb on chromosome 17 (72,646,040–76,613,675 bp; hg19). These regions were further explored for the candidate gene selection through UCSC genome browser and map viewer (NCBI). Genes were prioritized that are described or being involved in cell division, chromosome segregation and role in neurogenesis. There were a few possible candidate genes in these linkage regions, so we performed whole exome sequencing on one individual (MCP104-1). All variations were filtered in genes residing in the linkage intervals and no true pathogenic variant was detected in any gene. The underlying linkage regions were also analyzed for deletions or duplications

but no such event was detected. There is a possibility that this family has variation in the regulatory regions which are inaccessible to exome sequencing.

Family MCP104
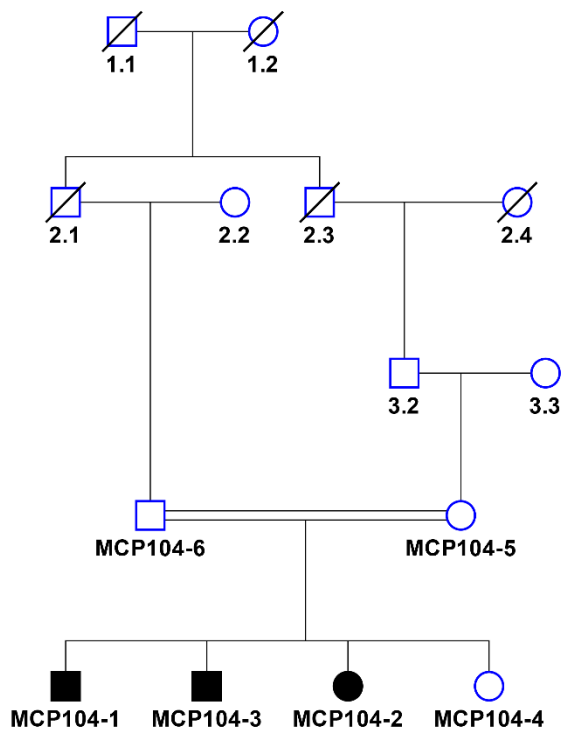
Figure 6.1.35: Pedigree of family MCP104 with autosomal recessive primary microcephaly.
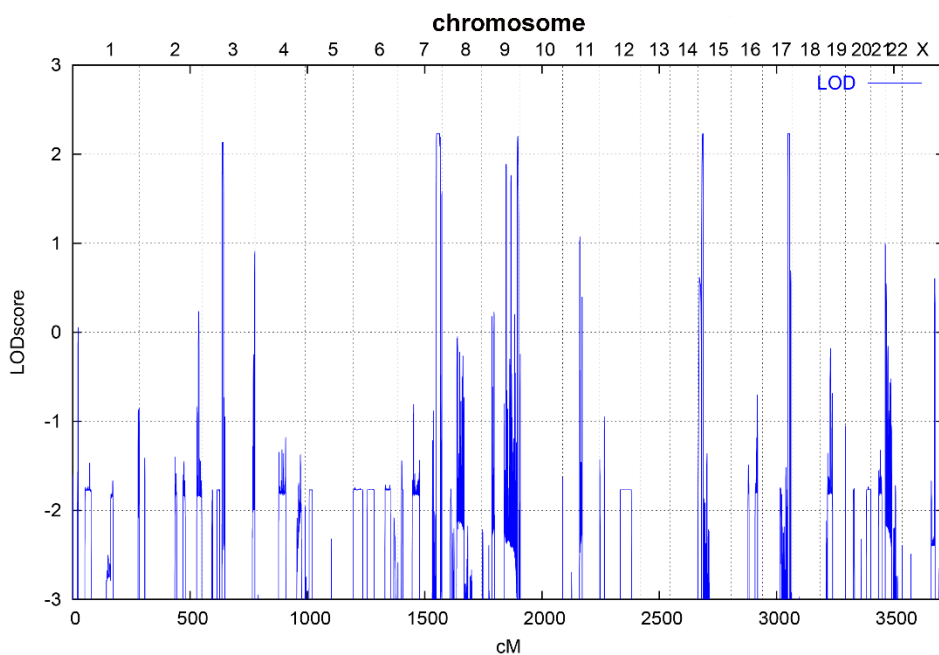
Figure 6.1.36: Genome-wide linkage scan of family MCP104. The highest LOD score values were obtained for regions on chromosomes 3, 7, 15 and 17 with maximum values of 2.23.

6.2 Microcephalic primordial dwarfism

A consanguineous family was ascertained with microcephalic primordial dwarfism from the Khyber Pukhoonkhwa (KPK), province of Pakistan. This family was named as MCP68.

6.2.1 Family MCP68

MCP 68 is a consanguineous family with seven affected individuals (Fig 6.2.1). The affected individuals exhibited profound congenital microcephaly (head circumference, from -7.44 SD to -15.14 SD) and substantial intrauterine growth retardation. The patients had markedly reduced height (-4.64 SD to -8.26 SD at current age), severe intellectual disability and speech impairment. Seizures were also reported in two patients MCP68-5 and MCP68-7. Aggressive behavior was evident in all cases, with marked delay in walking and evidence of hypertonia or contractures in lower limbs. In two cases MCP68-7 and MCP68-9, MRI (magnetic resonance imaging) findings suggested marked reduction in cortical size with simplified gyral folding. Brain stem and cerebellar were also reduced, with collection of increased cerebrospinal fluid (CSF), particularly in MCP68-9 in whom large interhemispheric arachnoid cysts were seen (Fig 6.2.2). On ophthalmological assessment ocular abnormalities were also observed e.g., microcornea, cataract, contractures and corneal opacity. In patient MCP68-9, one eye was completely absent (Fig 6.2.2-c). Additional clinical features were also observed in some patients, such as postaxial polydactyly in MCP68-5 and congenital heart disease in patient 4.8. No significant facial dysmorphsim or any typical facial gestalt was found in any patients apart from prominent microcephaly. Routine karyotyping was performed on one patient (MCP68-7) and no numerical or gross chromosomal aberrations were detected. The detail clinical data are summarized in the Table 6.2.1.

Family MCP68
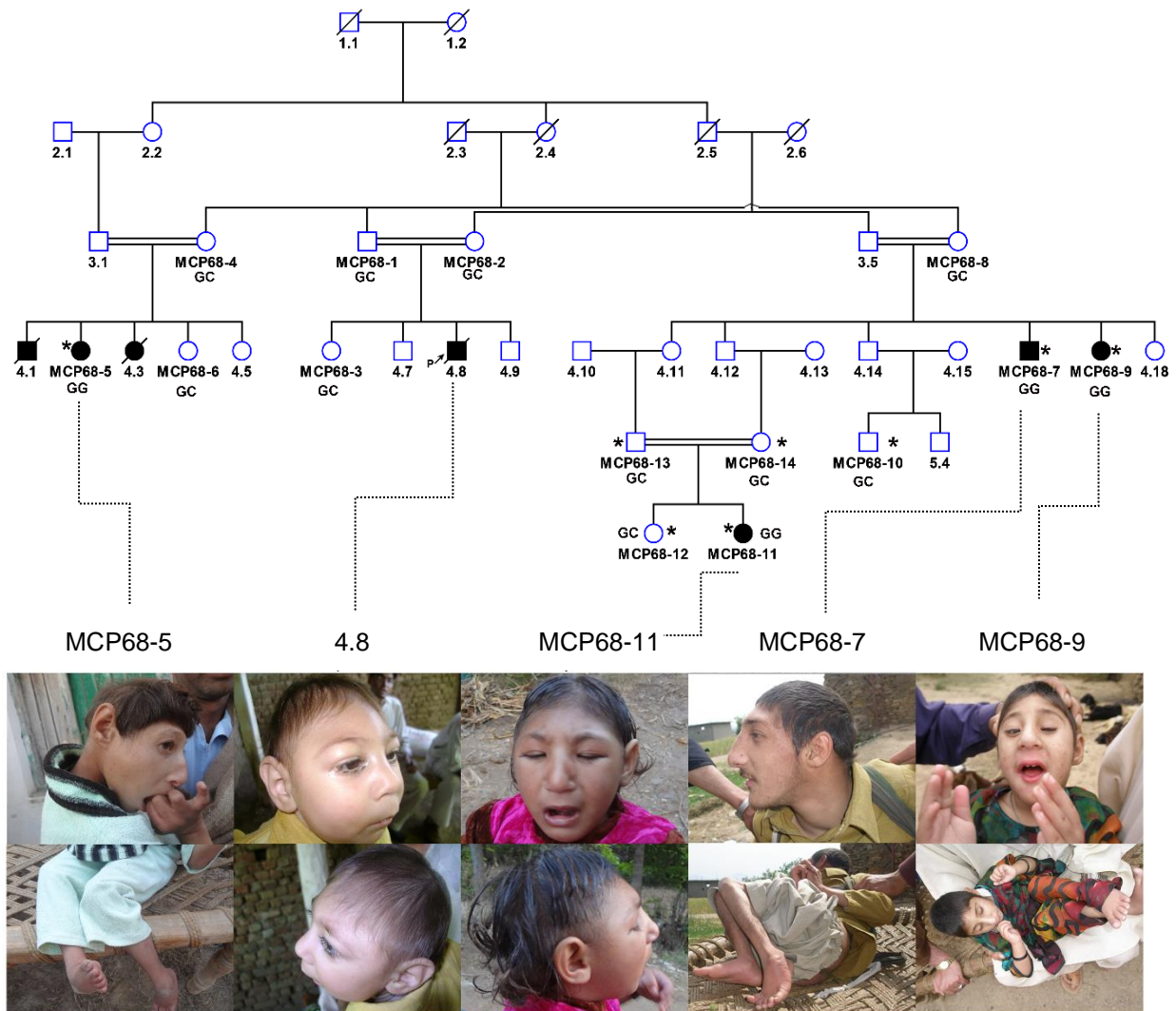
Figure 6.2.1: Pedigree and photographs of affected individuals of the family MCP68 with autosomal recessive microcephalic primordial dwarfism. It is large consanguineous family with seven affected individuals including three deceased ones. An asterisk indicates that the individual is genotyped. The double line shows the consanguinity and slashed subjects indicates the deceased individuals. The arrow shows the proband in third generation. He died at age 5.5 years due to heart problem (septal defect). Genotypes for the mutation is showed as GG for affected and GC for heterozygous individuals (Martin et al. 2014).

Table 6.2.1 Clinical summary of the patients with *PLK4* mutation in family MCP68

Individual	Sex	Age	At Examination		Intellectual Disability	Clinical features	Neuroimaging
			OFC/SD (cm)	Height/SD (cm)			
MCP68-7	M	12y	-7.44 (43cm)	-4.64 (115cm)	Profound. Can't sit unaided. No speech	Microcornea. Cataract (R). Contractures. Epilepsy	NA
MCP68-9	F	5y	-13.98 (35cm)	-6.9 (78cm)	Profound. Can't sit unaided. No speech.	Microphthalmia. Microcornea. Corneal opacity with cataract(R). Lumbar scoliosis	Microcephaly with simplified gyri. Large interhemispheric arachnoid cyst. Hypoplasia of pons and cerebellum. Partial agenesis of corpus callosum (short). Enlarged 4th ventricle
MCP68-5	F	10.5y	-13.46 (33cm)	8.26 (88cm	Very severe. Few words. Can't sit unaided.	Polydactyly – partial duplication right great toe. Talipes equinovarus.2/3 toe syndactyly (R+L). Contractures. Epilepsy	NA
MCP68-11	F	1.5y	-15.14 (30cm)	-8.07 (58cm)	Very severe. Can't sit unaided.	Microphthalmia. Microcornea. Vitreous inclusions. Contractures	Microcephaly with markedly simplified gyri. Hypoplasia of pons and cerebellum. Partial agenesis of corpus callosum (short). Venae galena cyst. Enlarged 4th ventricle.
4.8	M	0.66Y	-14.3 (29cm)	NA	Very severe Few words. Walked with support age 5.5y	Died age 5.5Y- congenital heart disease (septal defect)	NA
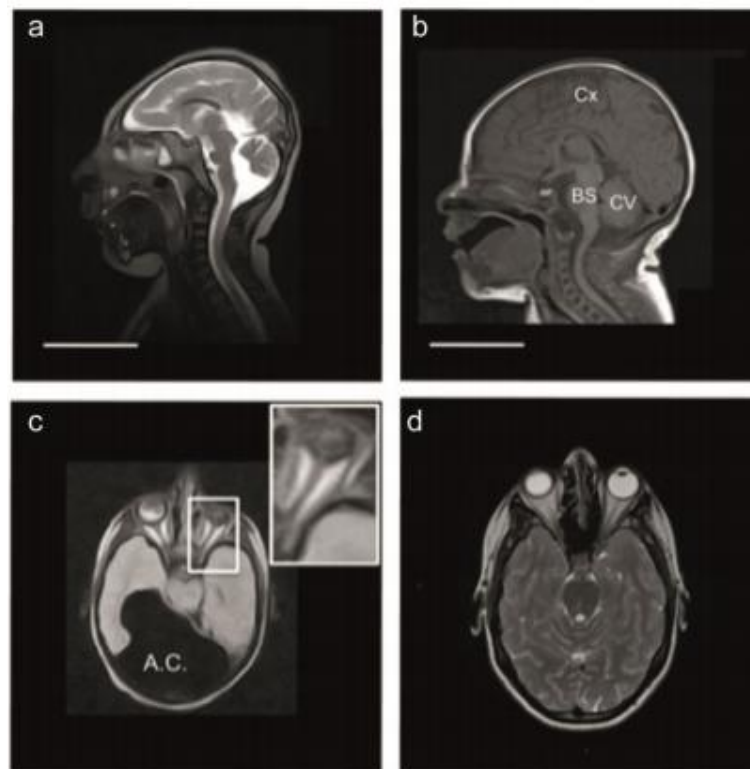
Figure 6.2.2: Structural abnormalities in the brain and eye of patients. Additional brain anomalies were detected through magnetic resonance imaging (MRI) of the brain. **(a)**Sagittal image of Patient MCP68-11, demonstrating severe reduction in cortical size, simplified gyral folding, thin brain stem and reduced cerebellar vermis size. **(b)** A normal age-matched control is shown for comparison: Cx Cortex, BS Brainstem CV: Cerebellar Vermis. **(c)** Structural eye abnormalities are evident in patient MCP68-9, with absent eye globe on the left (insert). In addition, a large region with signal intensity of cerebrospinal fluid (CSF) is evident within the cranium, consistent with a large interhemispheric arachnoid cyst (A.C.). **(d)** Axial sections of patient with normal anatomy for comparison. A, B, -T2 weighted images (CSF, white). C, D-T1 weighted images (CSF, black) (Martin et al. 2014).

4.2.1.1 Genome-wide mapping and new locus identification

Eight individuals that are marked by asterisk in the pedigree (Fig 6.2.1) were genotyped on an Illumina HumanCoreExome-12 v1.1 BeadChip. Genome-wide linkage analysis was performed with low resolution using 24,209 selected SNP markers, and mapped a peak on chromosome 4. Next, chromosome 4 was selected with maximum SNPs and multipoint analysis (ALLEGRO) LOD score was calculated that demonstrate a region of linkage on chromosome 4, with a maximum LOD score of z=4.7 (Fig 6.2.3). The underlying homozygous region defined a candidate region of 11.7 Mb (126,454,911-138,168,552; hg19) on chromosome 4. Haplotypes were constructed from homozygous SNPs, obtained from high-density SNP genotyping data. DNA segments identical among all affected are shown in the figure 6.2.4. Homozygous region was further explored for the candidate gene selection through UCSC genome browser and map viewer (NCBI). This region contains 56 annotated and predicted coding genes. Genes were scrutinized manually and none of these 56 genes were reported to associate with

microcephalic primordial dwarfism. Therefore, homozygosity mapping established a new disease locus in family MCP68 (Fig 6.2.3).
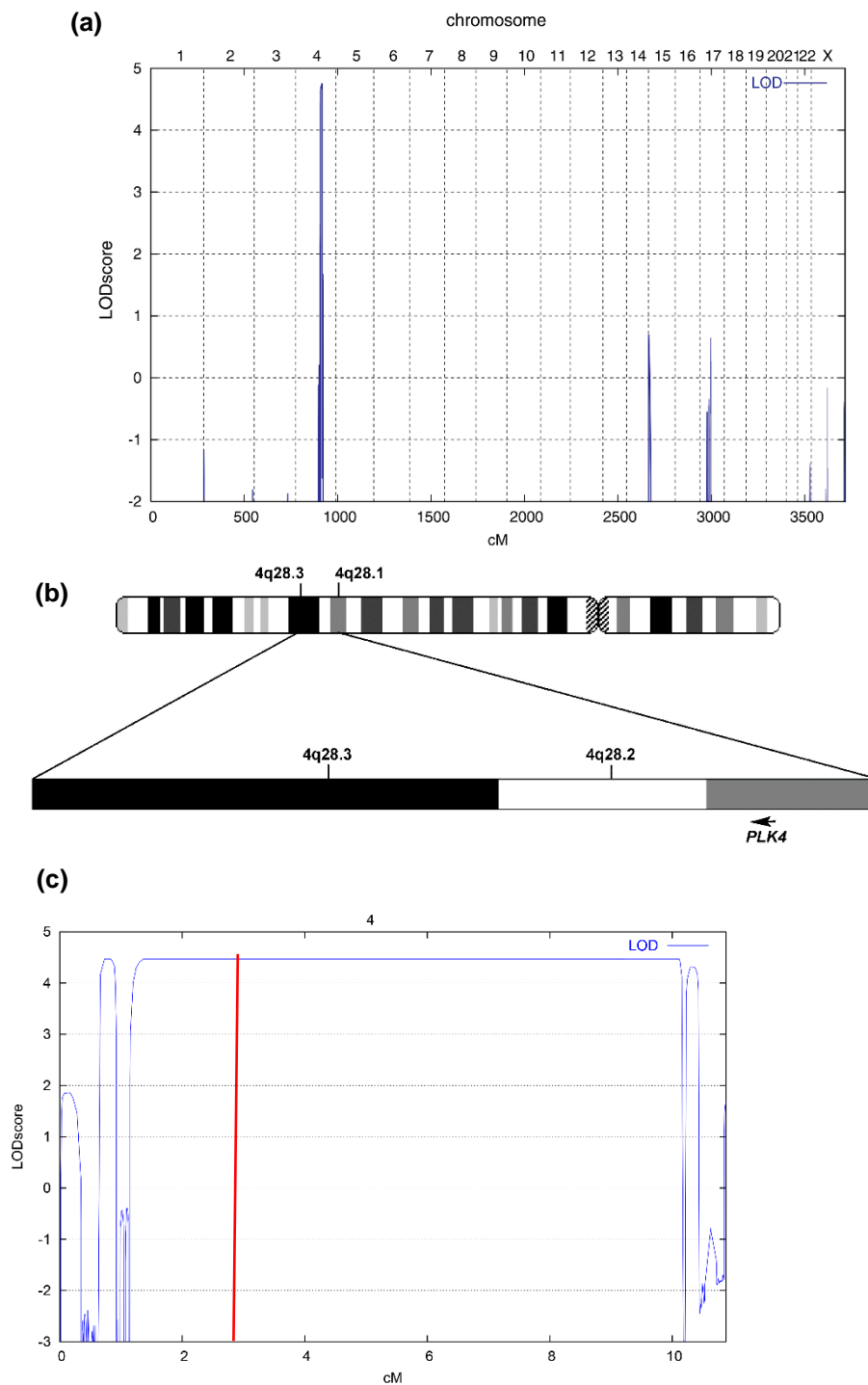
Figure 6.2.3: **(a)** Genome-wide scan of the family MCP68. **(a-b)** Homozygosity mapping identifies a novel locus for primordial dwarfism on chromosome 4 (4q28.3-4q28.2). Multipoint analysis (ALLEGRO) demonstrated high peak on chromosomes 4 with LOD score z=4.3. **(c)** Results of parametric linkage analysis showing LOD scores on chromosome 4 with the position of *PLK4* which is shown by red line. The coordinate of 0 (cM) in the figure corresponds to the 132.1 cM (Martin et al. 2014).
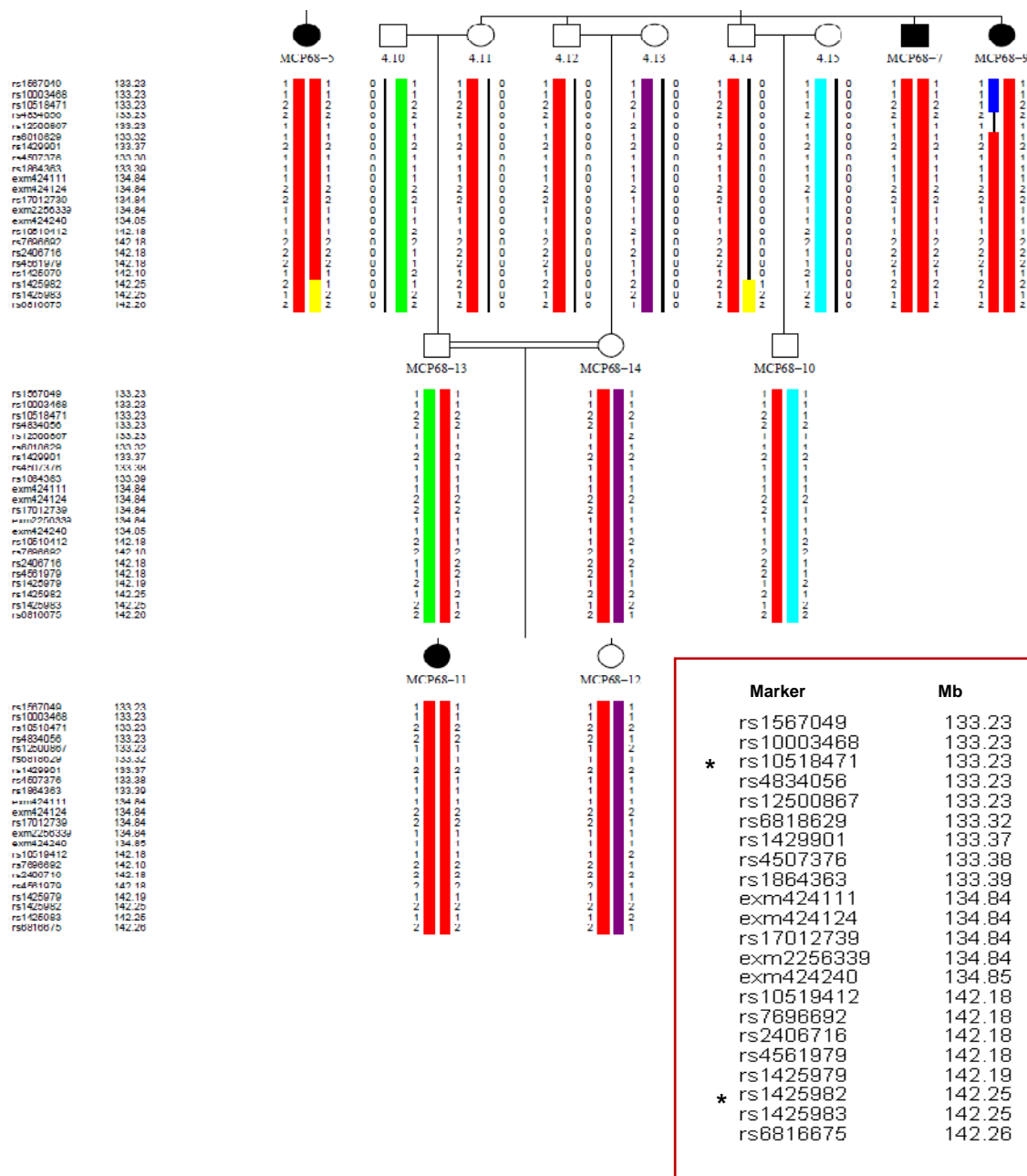
Figure 6.2.4: Haplotype reconstruction from Illumina beadchip SNP genotyping of eight individuals of family MCP68 across the chromosome 4 region containing *PLK4*. 1000 markers were deleted to show clearly the homozygous region. The homozygous markers were highlighted in the enclosed rectangle. Asterisk* denotes flanking markers, rs10518471 and rs1425982. The rs17012739 is located within *PLK4* gene.

6.2.1.2 Identification of a mutation in *PLK4*

After identification of the new disease locus, whole exome sequencing was performed of the affected boy MCP68-7. First variations were filtered in all known genes associated with primary microcephaly and microcephalic primordial dwarfism by using in-house exome pipeline (Varbank 2.1) but no deleterious mutation was found. Then, high quality unknown variants were prioritized in the linkage interval on chromosome 4, allowing filtering for homozygous and

compound heterozygous variations (dbSNP Build 135, 1000 Genomes Project database build 20110521 and the public Exome Variant Server, NHLBI ESP, Seattle, build ESP6500, In-house 511 exomes). Homozygous variations were identified in two different genes. One homozygous intronic mutation, c.2811–5C>G, detected in the *PLK4* (encoding Polo-like kinase 4) and the other mutation, c.1126_1132del(A)7ins(A)8, was found in *LARP1B* gene (La ribonucleoprotein domain family). Candidate genes were selected that are involved in cell division, chromosome segregation and embryonic development. *PLK4* was the most suited candidate gene based on the previous literature as it has important role in cell cycle by regulating centriole biogenesis. Meanwhile an independent study, described the second mutation in two additional families with phenotypes of MPD (Martin et al. 2014). The other gene, *LARP1B*, could be assumed not to be relevant for the phenotypes. This gene has no known function in pathways suggested for microcephalic primordial dwarfism. Furthermore, no variation in this gene was detected in other two independent patients with *PLK4* mutation (Martin et al. 2014). However, its potential role as a genetic modifier cannot be entirely excluded for the present family.

6.2.1.2.1 Validation of the mutation in *PLK4*

Capillary sequencing was performed to validate the variant c.2811–5C>G detected by exome sequencing in intron 15 of PLK4. DNA samples from all available MCP68 family members were bidirectionally sequenced. The affected individuals were homozygous for the mutation while parents and healthy sibs were heterozygous (Fig 6.2.5). This showed co-segregation of the mutation with the disease phenotype which are in accordance with the supposed recessive genetic model (Fig. 6.2.1). To investigate the disease allele frequency in the local population, 286 samples were analyzed mostly from the North West part of Pakistan through pyrosequencing (Fig. 6.2.6). No mutation was detected in any of these healthy control individuals. Moreover, this mutation was also not found in the 1000 Genomes Project and National Heart, Lung, and Blood Institute Grand Opportunity (NHLBIGO) Exome Sequencing Project (ESP) and in other public variant databases (Martin et al. 2014).

MCP68-7, homozygous mutant, c.2811–5C>G

C T T C A G T T A G G
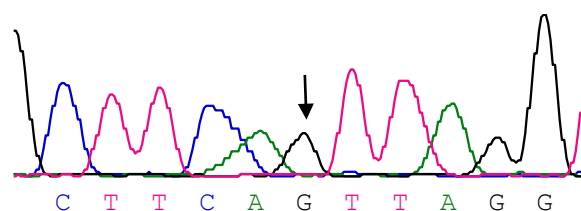
MCP68-13, heterozygous mutant, c.2811–5C>G

C   T   T   C   A   N   T   T   A   G   G

Wild type

C   T   T   C   A   C   T   T   A   G   G
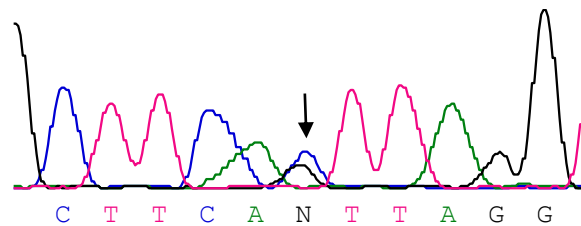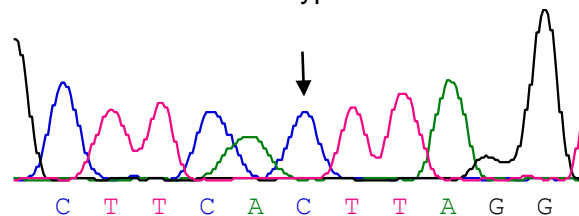
Figure 6.2.5: Sanger traces for mutation c.2811–5C>G in intron 15 of *PLK4*. Representative chromatogram for affected individual (MCP68-7), heterozygous parent (MCP68-13) and unrelated healthy control. Arrow shows site of mutation.
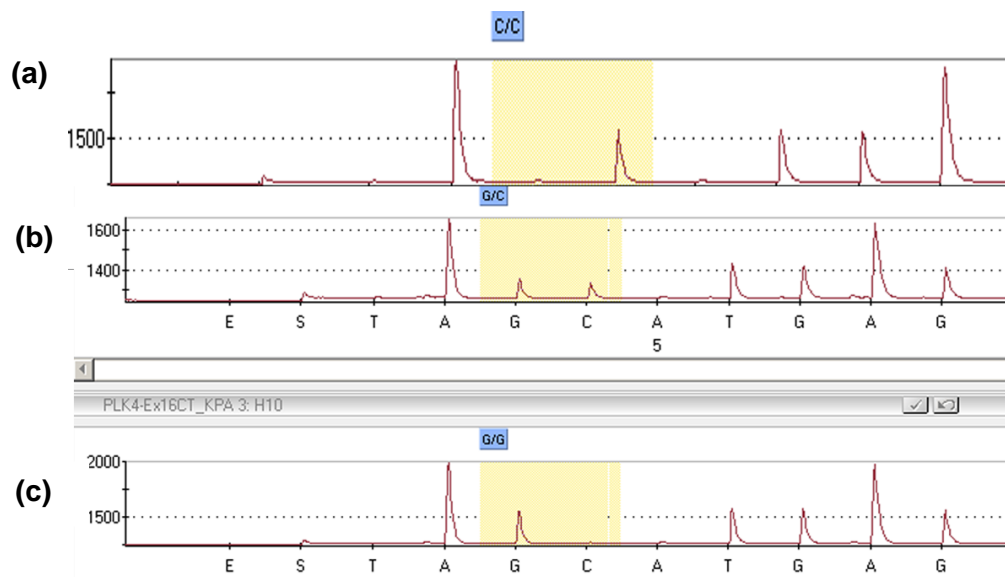
Figure 6.2.6: Pyrogram traces for the detection of transversion at position (g.127,898,434 bp, hg38) of intron 15 in *PLK4*. The disease allele frequency was estimated in the DNA pool of 286 healthy control individuals from Pakistan. (**a**, top) indicates no change for nucleotide (**b**, middle) shows the heterozygous for the mutation (MCP68-4, positive control). (**c**, below) demonstrates the homozygous mutant (MCP68-9, positive control).

6.2.1.3 Structure of the *PLK4*

The human *PLK4* is located at chromosome 4q28 (127,880,861-127,899,195; hg38). The official name of this gene is polo-like kinase 4. The *PLK4* is consist of 16 exons (Fig 6.2.7). It gives rise to 11 transcripts of variable length (only six are protein coding). The longest transcript with 3841 bases length codes for a protein of 970 amino acids.
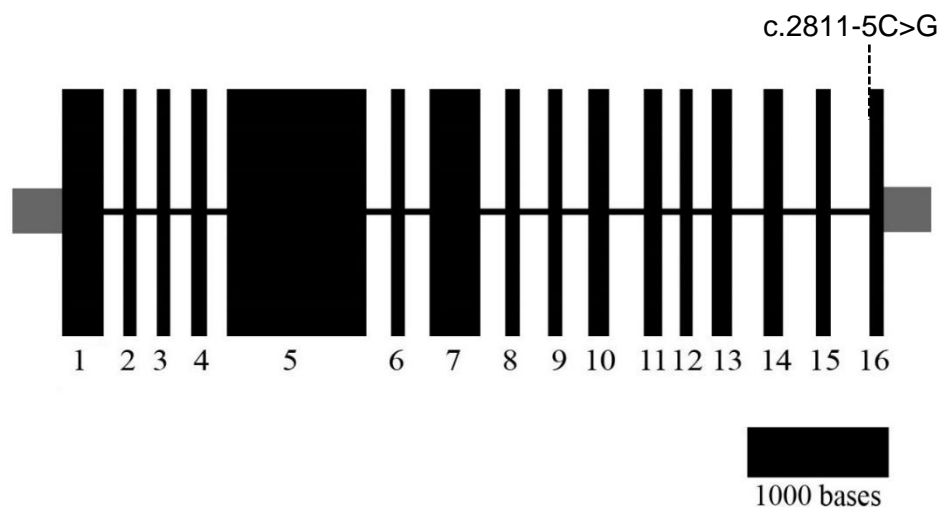
Figure 6.2.7: Schematic representation of the exon/intron structure of the human *PLK4*. Exons (vertical lines) are drawn according to scale whereas introns (straight lines) are arbitrary. The positions of coding exons (black), UTRs (grey) are indicated.The splice site mutation c.2811-5C>G is shown on intron 15.

6.2.1.4 Transcriptional consequences of the c.2811-5G>C *PLK4* mutation

The intronic variant c.2811-5G>C (Fig 6.2.8) was further followed at the transcript level to see how it affects the pre-mRNA splicing. The mutation was characterized at the mRNA level and its biological consequences predicted by in silico analysis.

6.2.1.4.1 In silico prediction of the alteration of the splicing

To evaluate the splice site prediction, sequences from genomic position (Chr4: 127,896,801-127,898,558 bp; hg38), spanning exon 15 and intron 16 of *PLK4*, were submitted to the online splice-site-prediction algorithms (http://www.cbs.dtu.dk/services/NetGene2/). Bioinformatics predicted the generation of a new splice acceptor site in intron 15, incorporating 4 bp of intronic sequence into exon16 of the *PLK4* transcript (Fig 6.2.7).
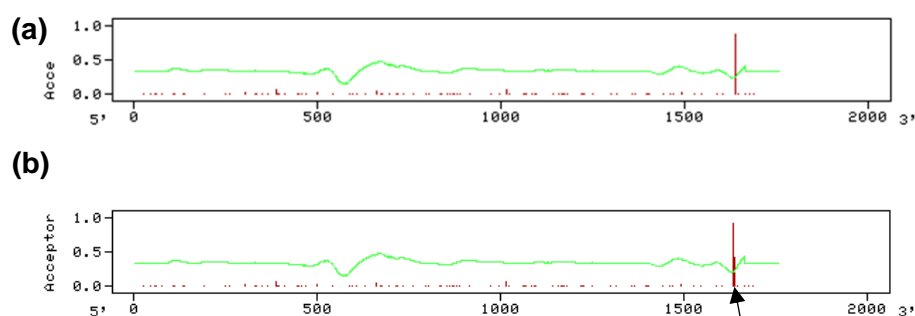
Figure 6.2.7: Schematics showing splice-site prediction of the altered 16th exon/intron junction. **(a)** Upper panel represents the normal acceptor site prediction. **(b)** Lower panel indicates the new predicated site (arrow).

6.2.1.4.2 Reverse transcription polymerase chain reaction (RT-PCR)

To assess the influence of the predicted splice site, RT-PCR amplifications were carried out targeting both exons 15 and 16 in total RNA extracted from the blood of patients and unrelated healthy control individuals. Upon Sanger sequencing of the amplified cDNA, it was detected that 4 bases, TTAG, from intron 15 were now added to the mutant transcript which resulted in a shift of the reading frame and caused a premature stop at the protein level (p.Arg936Serfs*1) while the control showed a normal transcript (Fig 6.2.8-L). In order to validate these findings and to exclude the presence of remaining wild-type transcripts in patients, capillary electrophoresis was performed on fluorescently labeled RT-PCR products. The labeled fragment sizes corresponded to 299 bp for the wild type and 303 bp for the homozygous mutant carrying the c.2811-5C>G mutation. Capillary electrophoresis revealed that the mutant transcript was 4 bp longer in size (303) than wild-type (299) (Fig 6.2.8-R). No residual wild-type transcript was detected in the homozygous c.2811-5C>G patient.
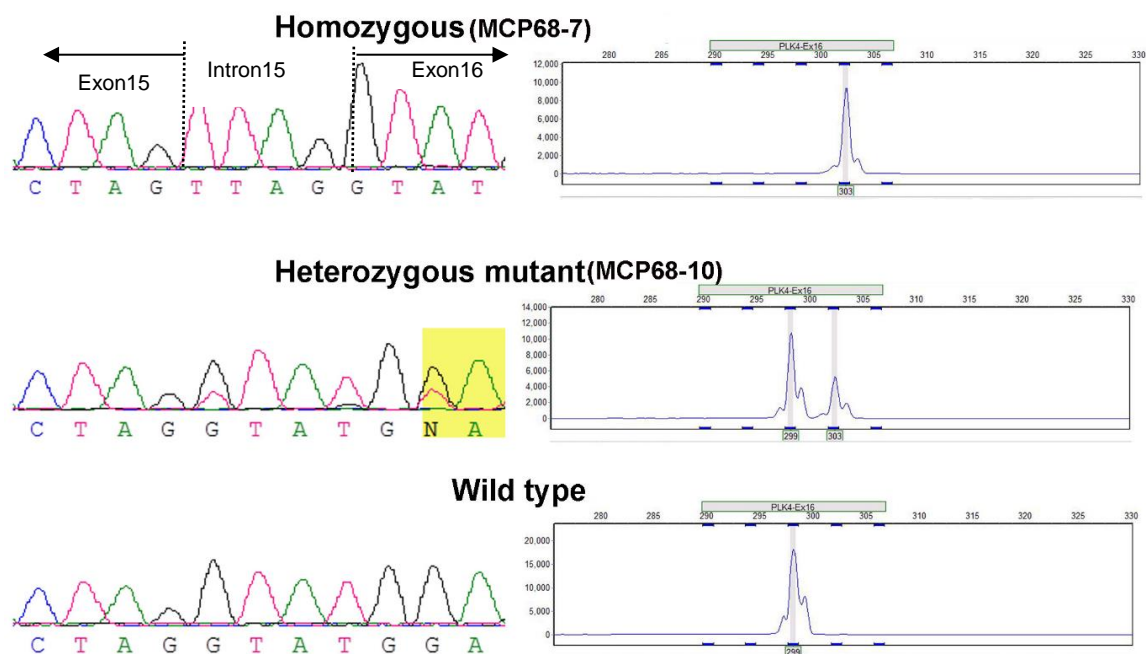
Figure 6.2.8: cDNA sequencing chromatogram of the exon15–exon16 junction of *PLK4* amplified by RT-PCR from patient and control RNA. The c.2811-5C>G mutation creates a new splice acceptor site which change the reading frame by incorporating 4 bases (TTAG) in the transcript of *PLK4*. **Left (L)**, electropherograms of RT-PCR products from patient MCP68-7, heterozygote MCP68-10 and control subjects. **Right (R)**, Genescan fragment sizing of RT-PCR products on ABI capillary sequencer, confirms that the mutant transcript is 4 bp greater in size than wild-type. No residual wild type transcript was detected in the homozygous c.2811-5C>G patient (Martin et al. 2014).

6.2.1.4.3 Quantitative PCR for gene expression analysis

Quantitative real-time PCR was performed with cDNAs from the blood of patient MCP68-7 and his heterozygous mother MCP68-8, along with one control samples. The relative expression of the target gene *PLK4* compared to control level was calculated and we found that *PLK4* transcript levels were substantially reduced in the patient (Fig 6.2.9).
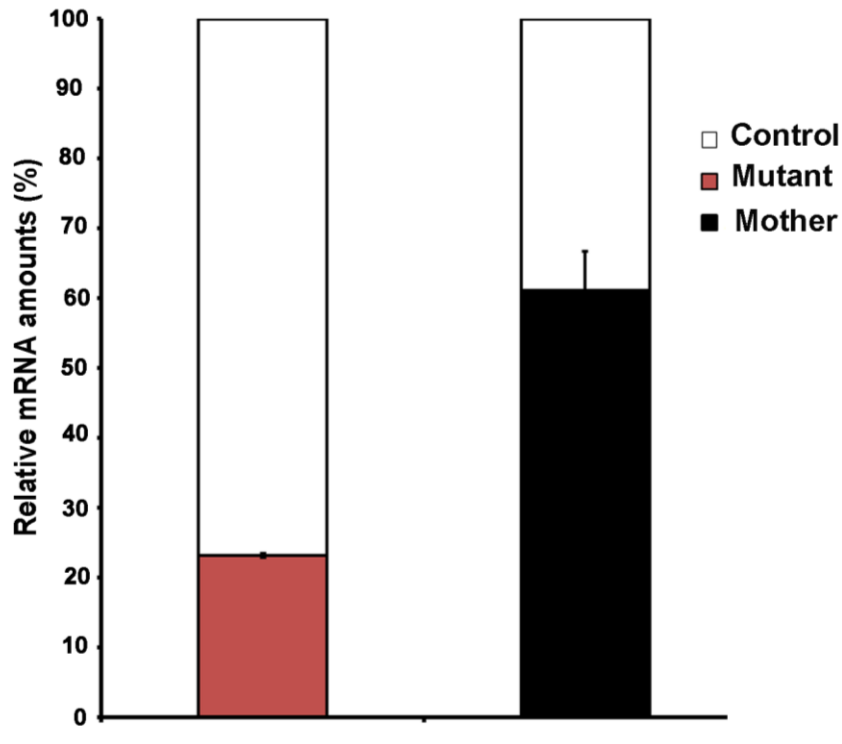
Figure 6.2.9: Quantitative Real-Time PCR data for mRNA of patient MCP68-7 and his unaffected mother MCP68-9. PLK4 mRNA amounts were quantified relative to a control sample, which was taken as 100%. In comparison to the control, levels of functional PLK4 transcript were reduced to 23% in patient and 61% in the mother (n=2, Error bars represent SD).

6.2.1.5 Consequences at the protein level of the c.2811-5G>C *PLK4* mutation

In order to investigate the impact of the splice site mutation (c.2811-5G>C) on PLK4 structure, a computer program was used to construct a 3-dimensional model. The mutation was also characterized at the cellular level in patient MCP68-7 primary cells.

6.2.1.5.1 Effects of the mutation on PLK4 structure

The splice site mutation (c.2811-5G>C, p.Arg936Serfs*1) leads to the loss of 34 amino acids at the C-terminus. The C-terminal polo box domain (PB3) of PLK4 consists of 79 amino acids (aa 884 to 962). PB3 exists as a domain-swapped homodimer, composed of twisted

antiparallel ß sheets and a C-terminal helix (Jana et al, 2012). The crystal structure for the mouse Plk4 PB3 domain is available (PDB 1MBY). It shows 97% homology with the human PB3 domain. To predict how the mutation affects the structure of PB3, the advanced program WebLab ViewerPro was used. From the three-dimensional structure it can be easily recognized that the deletion of the last 34 amino acids destabilizes the homodimer by weakening the interaction between the two molecules (Fig 6.2.10). The analysis was carried out by Dr. W. Höhne.
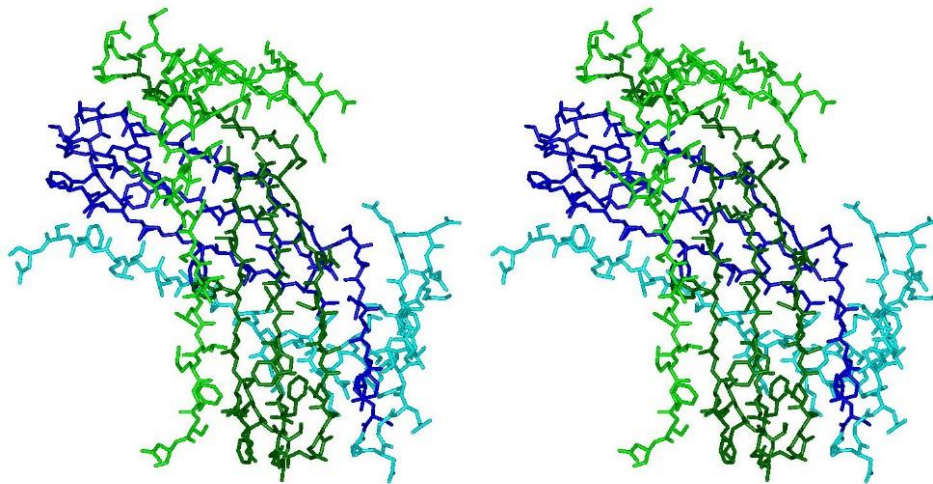
Figure 6.2.10: Stereo representation of the PB3 domain of murine Plk4. The structure (PDB code 1MBY) is a dimer (monomer A in blue, monomer B in green) with extended ß-sheet interactions and domain swapping. Missing parts in the mutant structure are highlighted in light blue and light green (Martin et al. 2014).The figure is drawn for relaxed eyes.

6.2.1.5.2 Protein quantification at the cellular level

To determine the effects of the p.Arg936Serfs*1 mutation on the protein level, total protein was quantified in cell lysates derived from the patient MCP68-7 and two wild type fibroblasts. Immunoblotting was performed and two PLK4-specific bands were observed by using a published antibody to PLK4 (Cizmecioglu et al. 2010). In total cell extracts from control fibroblast, these bands were more obvious. They corresponded to the FL (full length) and ALT (alternative) isoforms. In the patient cells, the band intensity is markedly reduced as would be expected from the reduced mRNA amounts (Fig 6.2.11). In addition, quantitative immunofluorescence was used to estimate PLK4 accumulation at the centrosome in patient and control fibroblasts.
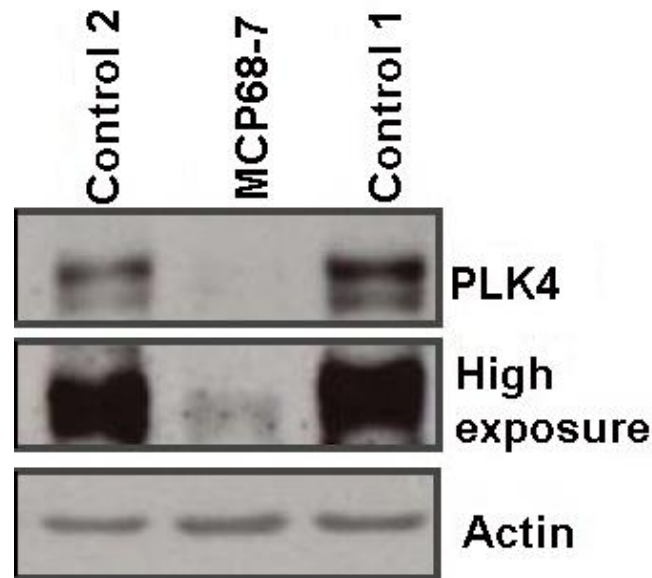
Figure 6.2.11: Immunoblotting demonstrates reduced PLK4 levels in patient-derived fibroblasts. Cell lysates are from asynchronous cells. The two PLK4 bands are evident for control fibroblasts on short exposure. Actin was used as loading control (Martin et al. 2014).

PLK4 is a low abundance protein (Holland et al. 2012a), and to enhance the signal intensity, proteasomal inhibitor MG132 was added to the cells 5 hours prior to fixation. Quantitative analysis of the immunofluorescence signal established that PLK4 levels were significantly reduced ($P < 0.0001$) in mutant fibroblasts with total PLK4 protein levels $15 \pm 0.02\%$ of control levels (Fig. 6.2.12). Therefore it was concluded that the major effect of the p.Arg936Serfs*1 alteration is a reduced PLK4 protein level due to reduced mRNA levels. Reduced protein stability may further reduce the level.

(a)

(b)
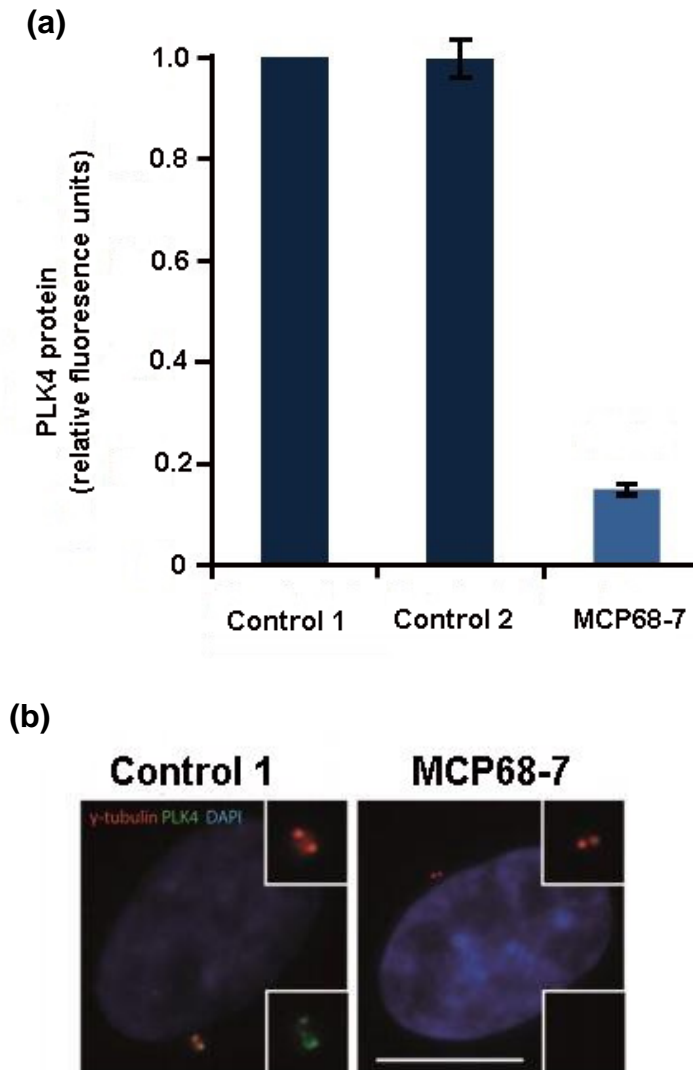
Figure 6.2.12: Immunofluorescence quantification of PLK4 at the centrosome. (a) PLK4 level was quantified at the centrosome by image analysis based quantitative immunofluorescence. *n*=3 experiments, with 50 cells/experiment; error bars, s.e.m for n= 150 cells. *P* values were calculated by two-tailed *t* test: ****$P \leq$ 0.0001. **(b)** The representative image of the patient MCP68-7 and control fibroblasts, treated with proteasome inhibitor (MG132, 10 μM) for 5 hours before fixation. The cells are stained with DAPI (blue), γ-tubulin (red) PLK4 (green). Insets of PLK4 and γ-tubulin staining are shown at 3× magnification. Scale bar = 10 μm (Martin et al. 2014).

6.2.1.6 Effects of reduced PLK4 protein level on the cell cycle

PLK4 is regarded as the master regulator of centriole biogenesis (Bettencourt-Dias et al. 2005). After revealing the consequences on the transcript and protein level, an impaired centriole duplication would be predicted. The number of centrioles changes during the cell cycle, so a mitotic stage was found the most convenient stage, to quantify centrioles in patient fibroblasts.

6.2.1.6.1 Centriole quantification at mitotic phase

To quantify the centriole number during mitosis, primary fibroblasts derived from patient MCP68-7 and two wild type fibroblasts were synchronized at the G2/M phase border using RO-3306 (Vassilev et al. 2006). The patient fibroblasts were treated for 22 hours and then the compound washed away to release the arrest. The cells were stained with centrin 3 specific antibodies to visualize the centriole. A significant reduction of centriole number in patient cells at prometaphase and metaphase was observed. The statistical analysis revealed that 12% of cells had a reduced number of centrioles compared to the two controls (Fig 6.2.13). Further characterization of centriolar phenotypes was carried out in mitotic cells and it was found that 29% of cells had a single centriole and 18% of cells were acentriolar (Fig 6.2.14-c). As the centriole is the key component of the centrosome, one might expect an impaired mitotic spindle formation. In fact, we observed that spindle formation was affected in patient mitotic cells. Furthermore, 9% of the cells were not capable to establish spindle poles and showed spindle disorganization, and 18% of the cells were without centrioles at both poles (balanced). Monopolar spindles were most frequently noticed in (63%) with single or two centrioles (Fig 6.2.14-b). Multipolar configurations were not observed in patient cells. Genomic instability was rarely seen in patient (2.8%) and control cells (2.5%).
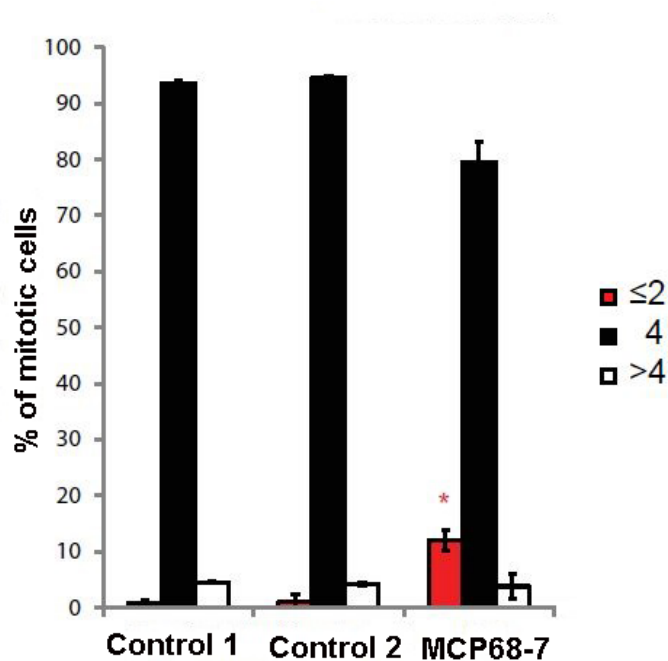
Figure 6.2.13: Centrin foci (centrin 3) quantified in prometaphase and metaphase fibroblasts (n=3 experiments, with 500 cells/experiment). Error bars, s.e.m, P values were calculated by two-tailed t test: *P ≤ 0.05 (Martin et al. 2014).
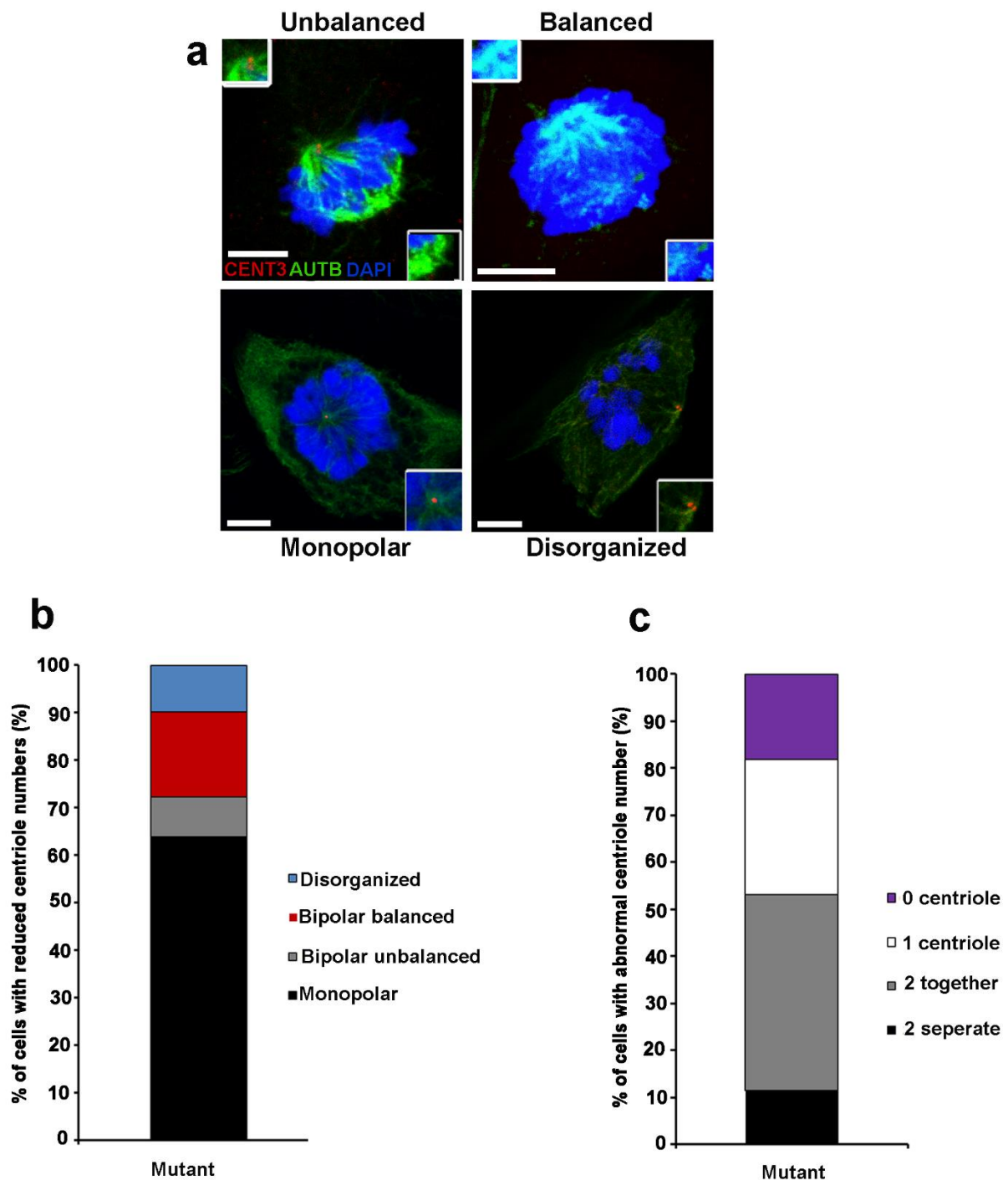
Figure 6.2.14: (**a**) Human fibroblast (MCP68-7, passage 5) carrying a *PLK4* mutation were fixed and stained for immunofluorescence microscopy, using antibodies specific for centrin 3 (red) to stain the centriole and YL1/2 for alpha tubulin (green). DNA is shown in blue (DAPI). The low level of PLK4 resulted in the reduced centriole numbers and impaired spindle formation in patient cells. Cells shown were "balanced" cells with a broad-based bipolar spindle, "unbalanced" cells with an unequal bipolar spindle, "monopolar" cells with only one spindle pole, and "disorganized" cells that failed to establish a spindle pole. Insets show enlargement of the centrioles. Scale bar: 5 µm. (**b**) This column indicates quantification of the mitotic spindle phenotypes observed in the 12% of mitotic MCP68-7-derived fibroblasts (n=3 experiments, with 72 cells/experiment). (**c**) Quantification of the centriole phenotypes observed in 12% mitotic cells (fibroblasts with ≤2 centrioles (n=3 experiments, with 175 cells/experiment). "2 together" indicates two centrioles detected at one spindle pole; 2 separate" indicates two centrioles detected, one at each pole.

6.2.6.2 Centrosome quantification at interphase

Centrosome duplication starts at S phase and separation occurs in late mitosis (Azimzadeh and Bornens 2007). After quantification of centrioles during mitosis, anti γ-tubulin antibody was used to evaluate centrosome numbers in interphase cells of patient MCP68-7 (passage 2) and wild type fibroblast (passage 5). The results showed that a significant percentage of patient cells (7%) was without centrosomes (Fig 6.2.15-b).
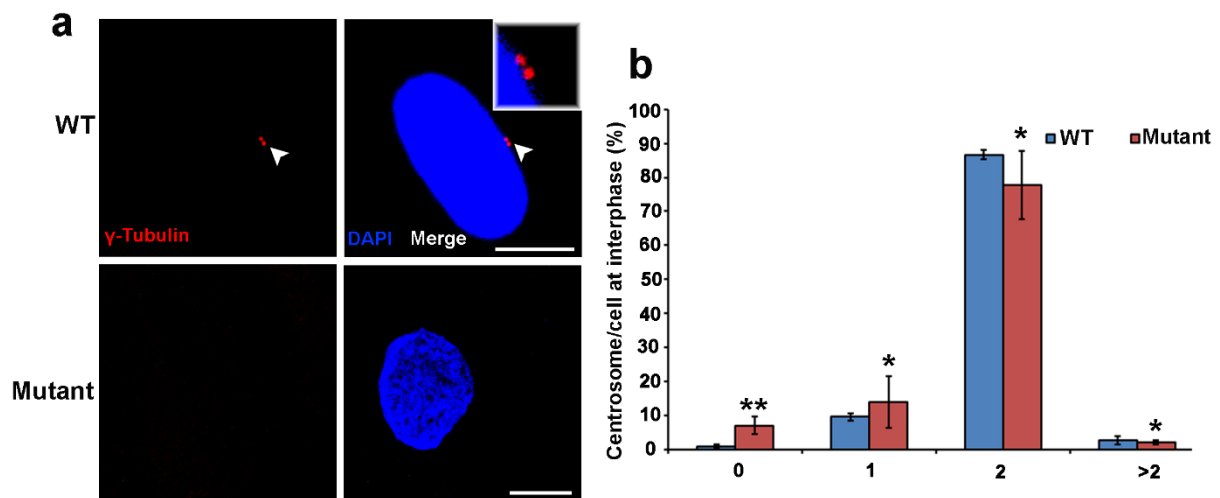
Figure 6.2.15: Centrosome quantification in interphase cells. (a) Images of centrosomes from interphase of unsynchronized fibroblasts derived from MCP68-7 and wild type stained for γ-tubulin and DAPI**.** Insets shows enlargement of the centrosomes**.** Scale bar**:** 10 μm. **(b)** Centrosome number is reduced in interphase of PLK4 deficient cells. Results were obtained from three independent experiments (Exp=3, n=1000). Error bars, s.e.m, P values were calculated by two-tailed t test: **P ≤ 0.0003 for cells with 0 centrosome, *P ≤ 0.0465 for cells with 1 centrosome, *P ≤ 0.0120 for cells with 2 centrosomes, *P ≤ 0.0514 for >2 centrosomes.

6.2.1.7 Centriole overduplication assay

The overexpression of PLK4 causes the mother centriole to assemble more than one procentriole within a cell cycle leading to the characteristic rosette-like arrangement of the procentriole and subsequently to centrosome amplification. The mutation p.Arg936Serfs*1 affects the centriole duplication in 12% of patient mitotic cells. To further confirm this finding, overexpression assays were performed, in which the enzyme activity can be assessed by overduplication of the centriole and increase in number (Kleylein-Sohn et al. 2007). This assay demonstrated that the mutant protein was significantly impaired in its ability to steer centriole biogenesis. In HeLa cells 48 hours after transfection with a plasmid allowing expression of GFP-tagged mutant PLK4, 22.7% of all transfected cells showed evidence of centriole amplification, whereas GFP-tagged wild-type PLK4 triggered centriole amplification in 76.5% of cells (Fig 6.2.16). Importantly, Rosette-like structures were formed in transfected cells. Centrin 2, required for centriole duplication (Salisbury et al. 2002) was used to visualize rosette

structure and localization. In addition, Cyclin A antibody was used to separate G1 from S-G2 phase cells. PLK4 was observed in the lumen of the rosette structure or to say on surface of the centrioles which may imply to the recruitment of centrin 2 to the PLK4 induced precursor structure (Fig 6.2.16). The results revealed that the mutant protein was able to localize effectively to the centrosome but not able to drive centriole amplification which shows that the residual mutant protein was also functionally impaired.
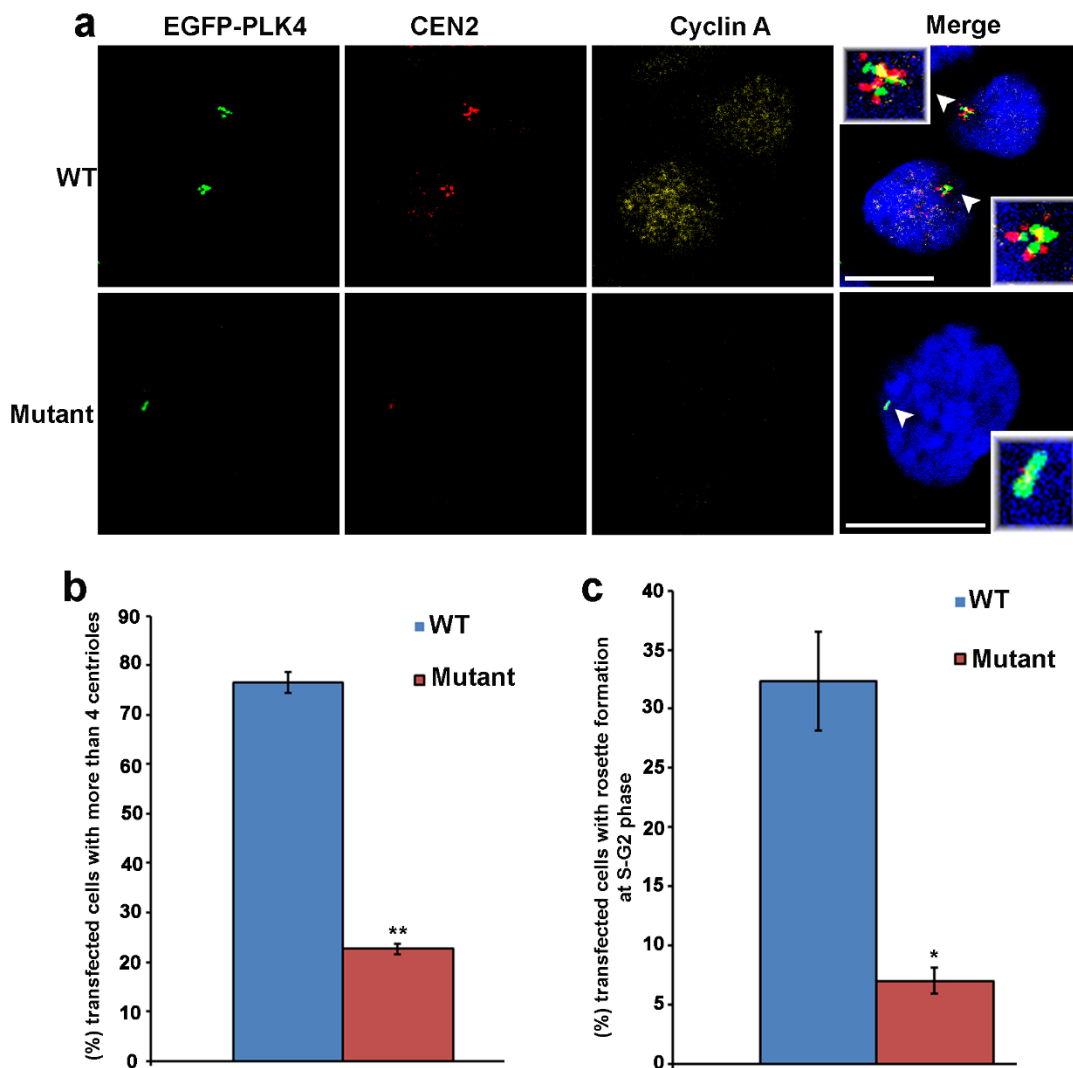
Figure 6.2.16: Overexpression of PLK4 causes centriole amplification. (a) HeLa cells were transfected with GFP-tagged wild type and mutant (p.Arg936Serfs*) PLK4 for 48 hours and prepared for immunofluorescence microscopy. Cells were costained with centrin 2, Cyclin A and DNA with DAP1. Scale bar: 10 μm. **(b-c)** Data taken from three independent experiments, counting 100 cells/ experiment, with error bar indicating the SEM.

6.2.1.8 Mitotic index

To test the hypothesis that abnormal spindle formation delays mitotic progression, the mitotic index (MI) was measured by using Phospho-histone H3 (PH3) antibody. It is known that

Histone-H3 Ser-10 phosphorylation is linked to chromosome condensation (Goto et al. 1999). Therefore, the presence of PH3 could indicate mitotic entry of cells. Fibroblasts were synchronized at the G2/M phase border using RO-3306 and then released for 45 minutes. The mitosis index was determined in patient MCP68-7 (passage 2) and control fibroblasts (passage 5). The result showed that for mutant cells a significantly higher percentage of cells (10%) was in mitosis as compared to control fibroblasts (7%) (Fig 6.2.17).
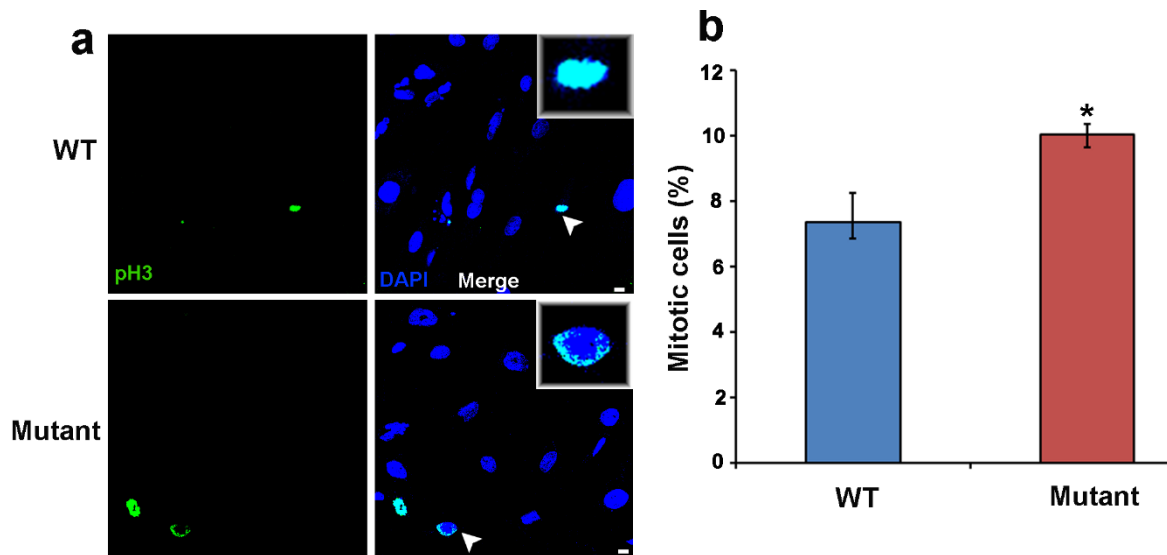
Figure 6.2.17: Increased mitotic index in patient fibroblasts MCP68-7. (a) Images of fibroblasts stained or PH3 and DNA (DAPI). Scale bar: 10 μm **(b)** Quantification of mitotic index in patient and control fibroblasts. Results are from three independent experiments. Error bars, s.e.m, P values were calculated by two-tailed t test: *P ≤ 0.0389 for cells.

6.2.1.9 Apoptosis assay

Fluorescence-activated cell sorting (FACS) analysis was performed in patient and control fibroblasts. The analysis of apoptosis demonstrated that limitated apoptosis occurs in pre-apoptotic patient cells (Fig 6.2.18). This might explain the fate of rare genomic instability seen both in patients and control cells or the cells with balanced configuration (no centriole at both poles).
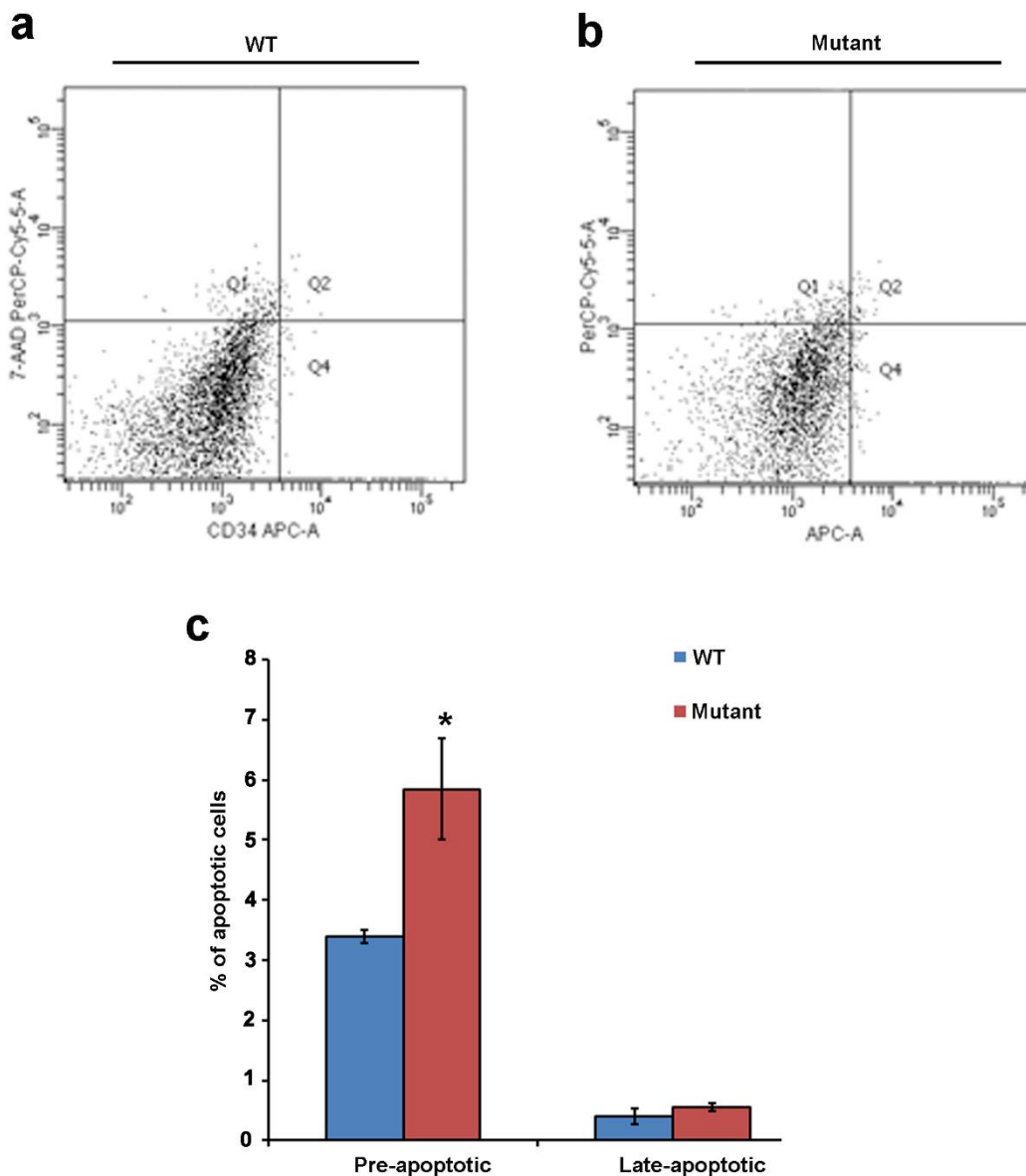
Figure 6.2.18: (a-b) Graphical view of pre- and late-apoptotic primary fibroblasts derived from a wild-type and patient MCP68-7 carrying mutations in PLK4 as analyzed by FACS. Uniform viable cells were observed in the third quarter (lower left). **(c)** Statistical data for pre- and late-apoptotic cells, significant difference was obtained for pre-apoptotic cells between wild type and patient MCP68-7 (passage 2). P value for pre-apoptotic cells is 0.0289 (Student's *t*-test). Four independent experiments were performed. Error bars show SEM.

6.2.1.10 Growth curve

To examine further the effects of spindle abnormalities and delayed mitotic progression, the growth rate was determined in patient and control cells. Patient derived fibroblasts (Fig 6.2.19) showed reduced proliferation rate as compared to the control cells. It appears that the cell cycle is impaired in patient cells due to abnormal mitosis.
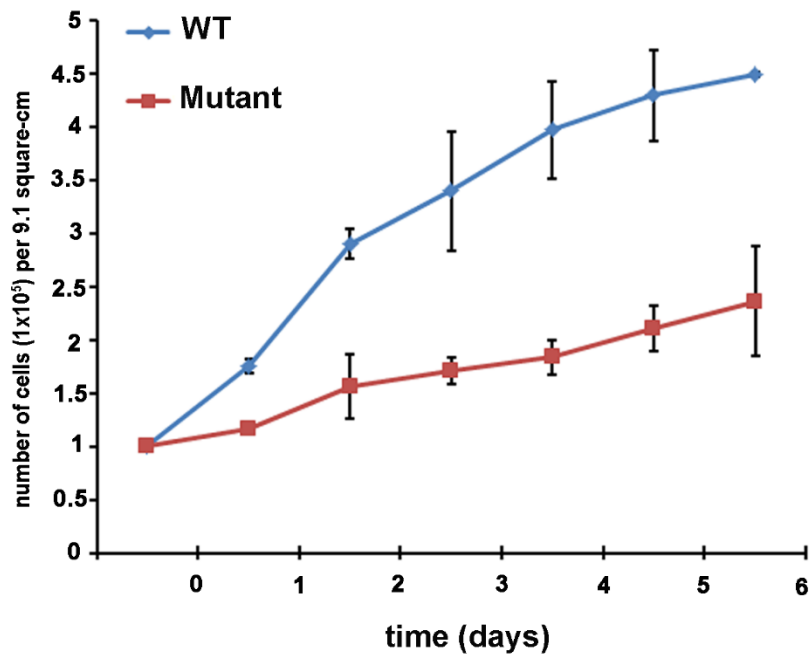
Figure 6.2.19: Growth curves of primary fibroblasts derived from patient MCP68-7 and control. It shows a reduced growth rate in patient cells. The data is taken from three independent experiments. Error bars represent standard deviativation.

7. Discussion

Mammals vary in their body size. For instance, the blue whale (150,000 kg) is 75-million fold bigger than the bumble bee bat (2g). This variation is not only observed in body size but also seen in the individual organ size (Klingseisen and Jackson 2011; Oldham et al. 2000). Most strikingly, the human brain evolved with a dramatic expansion of the cerebral cortex over the course of human evolution (Rakic 2009). The mechanisms that determine the body and brain size are poorly understood (Barkovich et al. 2012; Klingseisen and Jackson 2011). To unlock the blackbox processes of brain development and cellular growth, primary microcephaly (MCPH) and microcephalic primordial dwarfism (MPD) have been identified as the model disoders on the basis of Mendelian genetics (Hu et al. 2014; Klingseisen and Jackson 2011). In the last few years a tremendous progress has been made in the field of genomics, such as the development of genetic tools and sequencing technologies. These rapid advances made possible targeted capture of the genome, high read depth for low level mutation and exome sequencing. The latter is becoming a popular approach for the identification of genes underlying Mendelian disorders (Hu et al. 2014; Ng et al. 2009). The list of candidate genes is growing and 13 genes have been identified in MCPH and more than 20 genes in MPD (Table 3.1; Table 3.2).

In this study, both genetic conditions were studied in a total of 30 consanguineous families from Pakistan. Consanguineous unions increase homozygosity of the genome and expose hidden recessive alleles to express in their offspring. Homozygosity mapping is a powerful tool to spot segments of genomes in individuals that have been passed on to them through their parents (Woods et al. 2006). Hence, all these recessive consanguineous families were investigated through homozygosity mapping and positional candidate gene approach. Exome sequencing strategy complemented with homozygosity mapping was used for the identification of mutations. In addition, candidate-gene linkage analysis using highly polymorphic STR markers combined by Sanger sequencing were also used. The first part of this thesis is about the mutation spectrum and genetic heterogeneity of primary microcephaly while the second part addresses the identification of a new gene in MPD and its functional characterization at the cellular level.

7.1 Mutation screening of MCPH genes

Autosomal Recessive Primary Microcephaly (MCPH) is a rare heterogeneous genetic disorder characterized by reduced head circumference, low cognitive prowess and generally architectonically normal brains (Woods et al. 2005). Molecular genetic analysis of twenty nine consanguineous families revealed linkage to four out of thirteen MCPH loci and their

subsequent sequencing of the candidate gene resulted in different mutations (Table 6.1.2). In total, 23 (79.3%) families linked to the MCPH5 locus, 2 (6.8%) families to the MCPH1 locus, 1 (3.4%) family to the MCPH2 locus, 2 (6.8%) families with MCPH3, and one (3.4%) family was excluded to be linked to any of the known MCPH loci (Fig 6.1.1).

So far heterogeneity studies have been performed in only three populations like Pakistani (Gul et al. 2006; Roberts et al. 2002; Sajid et al. 2013), Indian (Kumar et al. 2004) and Iranian populations (Darvish et al. 2010). The relative frequencies of *ASPM* in Pakistani population is 45% in 42 (Roberts et al. 2002), 54.5% in 33 (Gul et al. 2006), 85.7% in 21 (Kousar et al. 2010), and 29.8% in 57 families with MCPH (Sajid et al. 2013). *ASPM* has also been found in MCPH families from Iran and India with frequency of 14% and 33 % respectively (Darvish et al. 2010; Kumar et al. 2004). *WDR62* (MCPH2) is the second common mutated gene that showed 10% (Roberts et al. 2002), 6% (Gul et al. 2006) and 8.7% (Sajid et al. 2013) prevalence in three independent cohorts of MCPH families from Pakistan. Furthermore, 2 out of 92 Iranian families and, 1 of 9 Indian families showed evidence of linkage to MCPH2. These studies and our findings suggest that mutations in *ASPM* are the most common cause of MCPH and have a high prevalence in the Pakistani populations (Bond et al. 2003; Gul et al. 2006; Hussain et al. 2013). Mutations identified in the known MCPH associated genes are described in the following.

7.1.1 Mutations in *MCPH1*

In family MCP118, a ~164 kb region of *MCPH1* (Fig 6.1.7) is deleted that covered ~160 kb of upstream sequence including exons 1 and 2. On manual inspection, it was found that ~164 kb deletion results in the loss of the first 38 amino acids, but leaves the open reading frame intact. Microcephalin contains three BRCA1-carboxy terminal domains (BRCT domains). The N terminus contains one BRCT domain (1-93 aa) whereas the C terminus has two tandem BRCT domains (672-833 aa) (Jackson et al. 2002; Richards et al. 2010). It was proposed that the mutant protein could start with the methionine that appears at position 96 aa of microcephalin. This microdeletion could affect the N-terminal BRCT domain of microcephalin that is known to interact with the centrosome and is responsible for preventing premature chromosome condensation (Richards et al. 2010). In the second linked family a homozygous deletion of ~577 kb (Fig 6.1.13) covered the first eleven exons and ~493 kb upstream promoter region. It caused an in-frame deletion of 657 amino acids which may lead to the loss of the N-terminal BRCT domain. Interestingly, this deletion did not affect the inversely arranged Angiopoietin 2 encoding gene (ANGPT2), which is situated within introns 12-14 of MCPH1. These overlapping microdeletions, which are not the first of its kind, suggest that the truncated gene is no longer functional (Garshasbi et al. 2006; Pfau et al. 2013). Mutations in *MCPH1* have been reported

to result in premature chromosomal condensation at G2 and delayed decondensation at G1, cause premature chromosomal condensation (PCC) with microcephaly (Garshasbi et al. 2006; Neitzel et al. 2002; Pfau et al. 2013; Trimborn et al. 2004). In both of the current families, patients had normal height and weight and no phenotypic resemblance was observed to PCC syndrome. Cytogenetic analyses were not performed due to geographical remoteness, therefore, PCC syndrome cannot excluded be absolutely. In total 13 mutations have been reported in the *MCPH1* gene, only two from Pakistan (Sajid et al. 2013; Venkatesh and Suresh 2014). The current work has increased the mutation spectrum of *MCPH1* in the Pakistani population. Additionally, it is the first to determine the precise break points in microdeletions >22 kb associated with MCPH genes.

7.1.2 Mutations in *WDR62* and *CDK5RAP2*

In family MCP129 the missense mutation p.Arg111Thr (Fig 6.1.16) was detected in *WDR62* which was previously reported in a family from northern Pakistan (Sajid et al. 2013). This mutation affects the first WD repeat of the WDR62 protein which may interfere with the functions of WDR62. In families MCP105 and MCP121, two novel nonsense mutations were identified in *CDK5RAP2* (Fig 6.1.23). One mutation p.Arg1372*, may result in a truncated protein having lost important functional domains such as the second SMC domain and Pericentrin and Golgi binding sites (Issa et al. 2013). The Second mutation p.Arg427* is predicted to results in a truncated product of 426 amino acids. It would lack most of the 215 kDa protein except γTuRC-binding domain and a part of the N-terminal SMC-domain. Mutations in *CDK5RAP2* are rare and only 8 different mutations have been described so far (Lancaster et al. 2013; Tan et al. 2014b). The *CDK5RAP2* mutations described here increased the total number from eight to ten.

7.1.3 Mutations in *ASPM*

Sequence analysis of the *ASPM* gene in 23 families revealed 12 different mutations including 3 previously published and nine novel mutations. The three known mutations p.Trp1326*, p.Lys1862Glu and p.Tyr3164 (Fig 6.1.24) were reported by (Kumar et al. 2004), (Darvish et al. 2010) and (Kousar et al. 2010) respectively. To date, mutation p.Trp1326 has been described in 32 Pakistani families (Sajid et al. 2013) and proposed as a founder mutation in Northern Pakistani MCPH families (Gul et al. 2006; Gul et al. 2007; Kousar et al. 2010; Muhammad et al. 2009; Sajid et al. 2013). Most of these reported families with the p.Trp1326* mutation have Pashtun ethnicity while some reports did not mention their ethnicity information (Mahmood et al. 2011). Pashtuns reside mainly in Southern and Eastern Afghanistan and in North-Western Pakistan (Haber et al. 2012; Mahmood et al. 2011; Mehdi et al. 1999). In this study, the same

mutation was found in 13 further families having Pashtun ethnicity including two from Afghanistan who are currently living as refugees in Khyber Pakhtunkhwa, Pakistan (Table 6.1.2), which increase the total number from 32 to 45. The nine novel mutations (Fig 6.1.24) identified include p.Gly807Glufs*7, p.Arg2700*, p.Trp1404*, p.Lys412Thrfs*5, p.K412Tfs*5, (p.K2837Mfs*34, p.Q1987*) and (p.Leu2285Argfs*6 p.Gln2377*). The last two are compound heterozygous mutations. These non-sense and frameshift mutations in the *ASPM* most probably induce non-sense-mediated mRNA decay and thus lead to a loss of *ASPM* function. The current study brings the total number of different *ASPM* mutations from 133 to 142 (Stenson et al. 2009; Tan et al. 2014a). Of the 142 mutations, 61 were nonsense mutations, 3 missense, 67 were small deletions or insertions leading to a change in the reading frame, eight were splice site mutations, two microdeletions and one with complex rearrangement (Stenson et al. 2009).

7.2. *PLK4* as a new gene in MPD

In the second part of this thesis, I report the discovery of the genetic basis of microcephalic primordial dwarfism in a consanguineous family from Northern Pakistan. Using homozygosity mapping (Fig 6.2.3-a) in combination with whole-exome sequencing, a biallelic loss-of- function mutation p.Arg936Serfs*1 (Fig 6.2.5), was found in the Polo-like kinase family member (*PLK4*) gene on chromosome 4 (Fig 6.2.3-b). The genetic data together with the functional analyses at the cellular level and its involvement in the regulation of centriole number established that the causal mutation is related to the phenotypes seen in the patients of family MCP68 (Fig 6.2.1).

Microcephalic primordial dwarfism (MPD) is the umbrella term for a group of genetically heterogeneous human syndromes. The hallmark features of MPD is pre and postnatal growth restrictions and smaller brain size (Klingseisen and Jackson 2011). Defective mitosis is one of the molecular mechanics that reflects the effects of MPD mutated genes products (Shaheen et al. 2014b). Mutations have been reported in genes involved in centrosome maturation and centriole biogenesis, and caused primary microcephaly and short stature (Bond et al. 2005; Hussain et al. 2012; Kalay et al. 2011; Kumar et al. 2009; Sir et al. 2011). In the present study, a consanguineous family with seven affected individuals was ascertained from Northern Pakistan. Affected individuals displayed profound microcephaly, severe intellectual disability and short stature. In addition, ocular anomalies were observed as well (Table 6.2.1). The family was subjected to homozygosity mapping and whole-exome sequencing which revealed a homozygous intronic mutation, c.2811–5C>G, in *PLK4.* PLK4 is the most divergent member of all PLKs which are localized to the centrosome and have important roles in centrosome duplication, cells viability, mitotic progression and embryonic development (Brito et al. 2012;

Ko et al. 2005). Plk4[-/-] is reported to be embryonic lethal and arrests the embryo at E7.5 with a large number of mitotic and apoptotic cells (Hudson et al. 2001). Plk4[+/-] mice demonstrated genomic instability with high incidence of lung and liver cancer. These findings were thought to be the consequences of centrosome amplification and multipolar configurations (Ko et al. 2005). In another independent study, failure of cytokinesis and increased levels of tetraploidy were shown in Plk4[-/+] derived mouse embryonic fibroblasts (MEF) (Rosario et al. 2010). On the contrary, a recent study demonstrated that Plk4 did not regulate cytokinesis directly and Plk4[+/-] MEF cells exhibited no centrosome amplification and tetraploidy (Holland et al. 2012a). Moreover, Plk4[+/-] cells showed about 90% Plk4 protein level at centrosome, although there was a 50% transcript reduction which explains the competence of Plk4 to auto regulate its own stability through phosphorylation within a 24-amino acid phosphodegron (Holland et al. 2012a; Holland et al. 2010). In the present study, cells derived from patient MCP68-7 demonstrated a 76% reduction at the transcript level and 15% of PLK4 protein accumulation at the centrosome which impaired the protein function, resulting in reduced centriole duplication (Fig 6.2.13). Meanwhile, an independent group also showed impaired centriole biogenesis and reduced transcript level in their mutant PLK4 cells (Martin et al. 2014). To determine that failure of centriole duplication in the patient is indeed due to reduction of the protein level, a centriole duplication assay was performed and it was found that the mutant protein was functionally impaired for centriole biogenesis (Fig 6.2.16). It is consistent with a recent study which demonstrated that Plk4-ΔPB3 expression increased significantly the percentage of cells with fewer than two centrioles, and related it to the reduced kinase activity (Klebba et al. 2015a). Hence, human PLK4 mutations may result in protein levels where PLK4 autoregulation fails to indemnify full kinase function but the protein still possesses enough enzyme activity to preclude lethality.

From the patient cells with abnormal centrioles number, the most frequent defects (around 63%) are the presence of unipolar spindle with a pair of centrioles and balanced spindle with no centriole at the poles (around 18%) (Fig 6.2.14-b). Two centrioles behave like one centrosome and it remains a highly stable structure in case of centriole duplication failure. Consequently, it will act as an MTOC (Bettencourt-Dias et al. 2005) and will not interfere with the cilia formation in patient cells that inherit this centrosome but would affect spindle formation as having one centrosome is worse than having none, which was previously described in Sak/Plk4 mutants in flies (Bettencourt-Dias et al. 2005). Such an observation has been further confirmed by an independent study which found no cilia formation in patient fibroblasts lacking centrioles (Martin et al. 2014). It indicates that monopolar spindles may be more challenging to mitotic progression than even an acentriolar configuration, however, cells were seen at interphase in patient cells without any centrosome, indicating successful clear up of mitosis (Fig 6.2.15). The most apt explanation for these results will be the formation of a bipolar spindle

to give license for chromosome segregation (Holland et al. 2012a). This phenomenon was further explained by a high mitotic index displayed in patient cells (Fig 6.2.17) which highlights that patient cells are spending more time in mitosis prior to bipolar segregation. In continuation of this experiment, a growth curve was generated which demonstrated that the cell cycle in patient cells is significantly impaired (Fig 6.2.19). Moreover, multipolar spindles were not noticed, suggesting abscence of aneuploidy in patient cells. In general, multipolar spindles predispose to substantial aneuploidy (Ganem et al. 2009). No previous literature and the recent study on PLK4 deficient cells demonstrated that monopolar spindles with reduced centriole number caused aneuploidy (Martin et al. 2014). It is similar to the study of plk4 depletion with small interfering RNA (siRNA) in time-lapse microscopy in RPE1 cells where it was found that Plk4 knock-down cells have a delay in mitosis with increased number of lagging chromosomes, but it was not investigated whether this leads to aneuploidy (Holland et al. 2012a). To further address the consequences of mitotic errors, flow cytometry (Annexin V staining) was performed on patient cells. No increase in apoptotic cells were observed, although early apoptotic cells displayed low level apoptosis in patient cells (Fig 6.2.18). This could be the ultimate fate of cells that have no centrioles at both poles or the rare genomic instability seen in both patient and control.

The findings at the cellular level were further substantiated in zebrafish studies by our collaborator on this project. Plk4 depletion in zebrafish demonstrated that Plk4 protein is required for generating normal cell numbers during development. They confirmed an overall size reduction, and suggest that this results from the lack of cell death because p53 is not involved in the growth arrest. They also detected increased mitotic index and reduction in centrioles number in embryos depleted of PLK4. Phenotypes indicative of ciliary defects such as ventral curvature, abnormal heat looping and retinal abnormalities were also described in morphant embryos (Martin et al. 2014).

Why does the *PLK4* mutation cause reduction of body and head size? Cellular experiments demonstrated that mutation in PLK4 only affects a proportion of mitosis (Fig 6.2.13). It may happen, that most patient cells form monospindle poles at some points in subsequent cell divisions during development. In such circumstances, cells will spend more time in mitosis and have reduced cell cycle efficiency which is enough to cause considerable dwarfism over the ~46 mitotic cell divisions required to reach the 1×10^{14} cells of the human body. To address the second question, a number of possibilities can be proposed in the light of previous literature. Firstly, most microcephalic and Seckel syndrome proteins have been found ubiquitously required for cell division, and they disproportionately affect the brain size. It may reflect a particular requirement of neural progenitor cells to coordinate symmetric and asymmetric cell divisions (Lancaster and Knoblich 2012), which will be disrupted in PLK4

mutated cells with abnormal spindle assembly. Secondly, mitotic progression may be impaired disproportionately in neuroepithelial tissues (Novorol et al. 2013), where high proliferation rates are required during neurogenesis, a process that has a restricted time frame during development. Thirdly, while substantial levels of aneuploidy were not detected in *PLK4* fibroblasts, monopolar spindle formation can predispose to chromosome segregation failure in transformed cells (Thompson and Compton 2008), and aneuploidy itself can cause microcephaly (Marthiens et al. 2013).

7.3 Conclusion and outlook

In this dissertation, I present the identification of the genetic basis of two inherited neurological conditions which share the most prominent phenotype of microcephaly. Morever, this study revealed 13 novel mutations in known MCPH associated genes and discovered a novel genetic cause of microcephalic primordial dwarfism. The first part adds to the mutational spectra of known MCPH-associated genes. It is in agreement with the previous findings that *ASPM* is the most common mutated gene in MCPH. The observed high frequency of the *ASPM* mutation p.Trp1326* underscores its prominent role as a founder mutation in the Pashtun population within Pakistan and Afghanistan. On the basis of these findings, it is recommended to sequence particularly the p.Trp1326* mutation as the first tier when testing MCPH families with Pashtun ethnicity and *ASPM* in general for families from northern Pakistan. The second part uncovers a distinct primordial dwarfism-causing mutation in the Polo-like kinase family member, master regulator of centriole biogenesis. This work highlights the importance of centrosome function for normal brain and body size and extends the phenotype spectrum associated with centriole biogenesis genes. It provides new insights into a gene in which early embryonic lethality has prevented extensive study.

In future, the identification of new genes will reveal additional critical regulators of human embryonic development. It will not only help to uncover the molecular pathology of both heterogeneous diseases but also pave the way to understand some longstanding questions in the field of developmental biology. The mechanism as to how various MCPH-associated proteins cause the same human phenotype and how the same centrosomal genes are implicated in primary microcephaly and microcephalic primordial dwarfism, is not precisely known. In short, the expected outcome of this project will result in improved diagnostics and counseling for primary microcephaly and related syndromes which are relatively more prevalent in Pakistan. The prognosis can be predicted and, the follow-up/treatment may be optimized to each individual/diagnosis.

8. References

Abuelo, D. (2007). "Microcephaly syndromes." *Semin Pediatr Neurol*, 14(3), 118-27.

Afonso, P. V., Zamborlini, A., Saib, A., and Mahieux, R. (2007). "Centrosome and retroviruses: the dangerous liaisons." *Retrovirology*, 4, 27.

Akram, D. S., Fehmina. Arif., and Jabeen, F. (2008). "HOW FREQUENT ARE THE CONSANGUINEOUS MARRIAGES"*JDUHS*. City, pp. 76-79.

Alderton, G. K., Joenje, H., Varon, R., Borglum, A. D., Jeggo, P. A., and O'Driscoll, M. (2004). "Seckel syndrome exhibits cellular features demonstrating defects in the ATR-signalling pathway." *Hum Mol Genet*, 13(24), 3127-38.

Ali, G. (2010). "Genetic deafness in Pakistani population." *J Pak Med Assoc*, 60(6), 418-9.

Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., and Zondervan, K. T. (2010). "Data quality control in genetic case-control association studies." *Nat Protoc*, 5(9), 1564-73.

Antonarakis, S. E., and Beckmann, J. S. (2006). "Mendelian disorders deserve more attention." *Nat Rev Genet*, 7(4), 277-82.

Avidor-Reiss, T., and Gopalakrishnan, J. (2013). "Building a centriole." *Curr Opin Cell Biol*, 25(1), 72-7.

Azimzadeh, J., and Bornens, M. (2007). "Structure and duplication of the centrosome." *J Cell Sci*, 120(Pt 13), 2139-42.

Baig, M. S., Sabih, D., Rahim, M. K., Azhar, A., Tariq, M., Sajid Hussain, M., Saqlan Naqvi, S. M., Raja, G. K., Khan, T. N., Jameel, M., Iram, Z., Noor, S., Baig, U. R., Qureshi, J. A., Baig, S. A., and Bakhtiar, S. M. (2012). "beta-Thalassemia in Pakistan: a pilot program on prenatal diagnosis in Multan." *J Pediatr Hematol Oncol*, 34(2), 90-2.

Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J., Nickerson, D. A., and Shendure, J. (2011). "Exome sequencing as a tool for Mendelian disease gene discovery." *Nature Reviews Genetics*, 12(11), 745-755.

Barbelanne, M., and Tsang, W. Y. (2014). "Molecular and cellular basis of autosomal recessive primary microcephaly." *Biomed Res Int*, 2014, 547986.

Barkovich, A. J., Guerrini, R., Kuzniecky, R. I., Jackson, G. D., and Dobyns, W. B. (2012). "A developmental and genetic classification for malformations of cortical development: update 2012." *Brain*, 135(Pt 5), 1348-69.

Barr, A. R., Kilmartin, J. V., and Gergely, F. (2010). "CDK5RAP2 functions in centrosome to spindle pole attachment and DNA damage response." *J Cell Biol*, 189(1), 23-39.

Bettencourt-Dias, M. (2013). "Q&A: Who needs a centrosome?" *BMC Biol*, 11, 28.

Bettencourt-Dias, M., Hildebrandt, F., Pellman, D., Woods, G., and Godinho, S. A. (2011). "Centrosomes and cilia in human disease." *Trends Genet*, 27(8), 307-15.

Bettencourt-Dias, M., Rodrigues-Martins, A., Carpenter, L., Riparbelli, M., Lehmann, L., Gatt, M. K., Carmo, N., Balloux, F., Callaini, G., and Glover, D. M. (2005). "SAK/PLK4 is required for centriole duplication and flagella development." *Curr Biol*, 15(24), 2199-207.

Bicknell, L. S., Walker, S., Klingseisen, A., Stiff, T., Leitch, A., Kerzendorfer, C., Martin, C. A., Yeyati, P., Al Sanna, N., Bober, M., Johnson, D., Wise, C., Jackson, A. P., O'Driscoll, M., and Jeggo, P. A. (2011). "Mutations in ORC1, encoding the largest subunit of the origin recognition complex, cause microcephalic primordial dwarfism resembling Meier-Gorlin syndrome." *Nat Genet*, 43(4), 350-5.

Bilguvar, K., Ozturk, A. K., Louvi, A., Kwan, K. Y., Choi, M., Tatli, B., Yalnizoglu, D., Tuysuz, B., Caglayan, A. O., Gokben, S., Kaymakcalan, H., Barak, T., Bakircioglu, M., Yasuno, K., Ho, W., Sanders, S., Zhu, Y., Yilmaz, S., Dincer, A., Johnson, M. H., Bronen, R. A., Kocer, N., Per, H., Mane, S., Pamir, M. N., Yalcinkaya, C., Kumandas, S., Topcu, M., Ozmen, M., Sestan, N., Lifton, R. P., State, M. W., and Gunel, M. (2010). "Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations." *Nature*, 467(7312), 207-10.

Bleichert, F., Balasov, M., Chesnokov, I., Nogales, E., Botchan, M. R., and Berger, J. M. (2013). "A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc6 impedes origin recognition complex formation." *Elife*, 2, e00882.

Bond, J., Roberts, E., Mochida, G. H., Hampshire, D. J., Scott, S., Askham, J. M., Springell, K., Mahadevan, M., Crow, Y. J., Markham, A. F., Walsh, C. A., and Woods, C. G. (2002). "ASPM is a major determinant of cerebral cortical size." *Nat Genet*, 32(2), 316-20.

Bond, J., Roberts, E., Springell, K., Lizarraga, S. B., Scott, S., Higgins, J., Hampshire, D. J., Morrison, E. E., Leal, G. F., Silva, E. O., Costa, S. M., Baralle, D., Raponi, M., Karbani, G., Rashid, Y., Jafri, H., Bennett, C., Corry, P., Walsh, C. A., and Woods, C. G. (2005). "A centrosomal mechanism involving CDK5RAP2 and CENPJ controls brain size." *Nat Genet*, 37(4), 353-5.

Bond, J., Scott, S., Hampshire, D. J., Springell, K., Corry, P., Abramowicz, M. J., Mochida, G. H., Hennekam, R. C., Maher, E. R., Fryns, J. P., Alswaid, A., Jafri, H., Rashid, Y., Mubaidin, A., Walsh, C. A., Roberts, E., and Woods, C. G. (2003). "Protein-truncating mutations in ASPM cause variable reduction in brain size." *Am J Hum Genet*, 73(5), 1170-7.

Bornens, M. (2012). "The centrosome in cells and organisms." *Science*, 335(6067), 422-6.

Botstein, D., and Risch, N. (2003). "Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease." *Nat Genet*, 33 Suppl, 228-37.

Boussif, O., Lezoualch, F., Zanta, M. A., Mergny, M. D., Scherman, D., Demeneix, B., and Behr, J. P. (1995). "A Versatile Vector for Gene and Oligonucleotide Transfer into Cells in Culture and in-Vivo - Polyethylenimine." *Proceedings of the National Academy of Sciences of the United States of America*, 92(16), 7297-7301.

Boycott, K. M., Vanstone, M. R., Bulman, D. E., and MacKenzie, A. E. (2013). "Rare-disease genetics in the era of next-generation sequencing: discovery to translation." *Nat Rev Genet*, 14(10), 681-91.

Brito, D. A., Gouveia, S. M., and Bettencourt-Dias, M. (2012). "Deconstructing the centriole: structure and number control." *Curr Opin Cell Biol*, 24(1), 4-13.

Broman, K. W., Murray, J. C., Sheffield, V. C., White, R. L., and Weber, J. L. (1998). "Comprehensive human genetic maps: individual and sex-specific variation in recombination." *Am J Hum Genet*, 63(3), 861-9.

Buchman, J. J., Tseng, H. C., Zhou, Y., Frank, C. L., Xie, Z., and Tsai, L. H. (2010). "Cdk5rap2 interacts with pericentrin to maintain the neural progenitor pool in the developing neocortex." *Neuron*, 66(3), 386-402.

Carvalho-Santos, Z., Azimzadeh, J., Pereira-Leal, J. B., and Bettencourt-Dias, M. (2011). "Evolution: Tracing the origins of centrioles, cilia, and flagella." *J Cell Biol*, 194(2), 165-75.

Chavali, P. L., Putz, M., and Gergely, F. (2014). "Small organelle, big responsibility: the role of centrosomes in development and disease." *Philos Trans R Soc Lond B Biol Sci*, 369(1650).

Chen, J. F., Zhang, Y., Wilde, J., Hansen, K. C., Lai, F., and Niswander, L. (2014). "Microcephaly disease gene Wdr62 regulates mitotic progression of embryonic neural stem cells and brain size." *Nat Commun*, 5, 3885.

Chenn, A., and Walsh, C. A. (2002). "Regulation of cerebral cortical size by control of cell cycle exit in neural precursors." *Science*, 297(5580), 365-9.

Cizmecioglu, O., Arnold, M., Bahtz, R., Settele, F., Ehret, L., Haselmann-Weiss, U., Antony, C., and Hoffmann, I. (2010). "Cep152 acts as a scaffold for recruitment of Plk4 and CPAP to the centrosome." *J Cell Biol*, 191(4), 731-9.

Clancy, B., Darlington, R. B., and Finlay, B. L. (2001). "Translating developmental time across mammalian species." *Neuroscience*, 105(1), 7-17.

Cowie, V. (1960). "The genetics and sub-classification of microcephaly." *J Ment Defic Res*, 4, 42-7.

Cox, J., Jackson, A. P., Bond, J., and Woods, C. G. (2006). "What primary microcephaly can tell us about brain growth." *Trends Mol Med*, 12(8), 358-66.

Darvish, H., Esmaeeli-Nieh, S., Monajemi, G. B., Mohseni, M., Ghasemi-Firouzabadi, S., Abedini, S. S., Bahman, I., Jamali, P., Azimi, S., Mojahedi, F., Dehghan, A., Shafeghati, Y., Jankhah, A., Falah, M., Soltani Banavandi, M. J., Ghani, M., Garshasbi, M., Rakhshani, F., Naghavi, A., Tzschach, A., Neitzel, H., Ropers, H. H., Kuss, A. W., Behjati, F., Kahrizi, K., and Najmabadi, H. (2010). "A clinical and molecular genetic study of 112 Iranian families with primary microcephaly." *J Med Genet*, 47(12), 823-8.

de Carcer, G., Escobar, B., Higuero, A. M., Garcia, L., Anson, A., Perez, G., Mollejo, M., Manning, G., Melendez, B., Abad-Rodriguez, J., and Malumbres, M. (2011). "Plk5, a polo box domain-only protein with specific roles in neuron differentiation and glioblastoma suppression." *Mol Cell Biol*, 31(6), 1225-39.

Devi, A. R., Rao, N. A., and Bittles, A. H. (1987). "Inbreeding and the incidence of childhood genetic disorders in Karnataka, South India." *J Med Genet*, 24(6), 362-5.

Dobbelaere, J., Josue, F., Suijkerbuijk, S., Baum, B., Tapon, N., and Raff, J. (2008). "A genome-wide RNAi screen to dissect centriole duplication and centrosome maturation in Drosophila." *Plos Biology*, 6(9), 1975-1990.

Doxsey, S. (2001). "Re-evaluating centrosome function." *Nat Rev Mol Cell Biol*, 2(9), 688-98.

Elahi, M. M., Elahi, F., Elahi, A., and Elahi, S. B. (1998). "Paediatric hearing loss in rural Pakistan." *J Otolaryngol*, 27(6), 348-53.

Elliott, P. J., Zollner, T. M., and Boehncke, W. H. (2003). "Proteasome inhibition: a new anti-inflammatory strategy." *J Mol Med (Berl)*, 81(4), 235-45.

Fish, J. L., Kosodo, Y., Enard, W., Paabo, S., and Huttner, W. B. (2006). "Aspm specifically maintains symmetric proliferative divisions of neuroepithelial cells." *Proc Natl Acad Sci U S A*, 103(27), 10438-43.

Fong, K. W., Choi, Y. K., Rattner, J. B., and Qi, R. Z. (2008). "CDK5RAP2 is a pericentriolar protein that functions in centrosomal attachment of the gamma-tubulin ring complex." *Molecular Biology of the Cell*, 19(1), 115-125.

Ganem, N. J., Godinho, S. A., and Pellman, D. (2009). "A mechanism linking extra centrosomes to chromosomal instability." *Nature*, 460(7252), 278-82.

Garshasbi, M., Motazacker, M. M., Kahrizi, K., Behjati, F., Abedini, S. S., Nieh, S. E., Firouzabadi, S. G., Becker, C., Ruschendorf, F., Nurnberg, P., Tzschach, A., Vazifehmand, R., Erdogan, F., Ullmann, R., Lenzner, S., Kuss, A. W., Ropers, H. H., and Najmabadi, H. (2006). "SNP array-based homozygosity mapping reveals MCPH1 deletion in family with autosomal recessive mental retardation and mild microcephaly." *Hum Genet*, 118(6), 708-15.

Genin, A., Desir, J., Lambert, N., Biervliet, M., Van Der Aa, N., Pierquin, G., Killian, A., Tosi, M., Urbina, M., Lefort, A., Libert, F., Pirson, I., and Abramowicz, M. (2012). "Kinetochore KMN network gene CASC5 mutated in primary microcephaly." *Hum Mol Genet*, 21(24), 5306-17.

Gibbs, J. R., and Singleton, A. (2006). "Application of genome-wide single nucleotide polymorphism typing: simple association and beyond." *PLoS Genet*, 2(10), e150.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E. M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S., Russ, C., Gabriel, S., Jaffe, D. B., Lander, E. S., and Nusbaum, C. (2009). "Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing." *Nat Biotechnol*, 27(2), 182-9.

Gonzalez, C., Saunders, R. D., Casal, J., Molina, I., Carmena, M., Ripoll, P., and Glover, D. M. (1990). "Mutations at the asp locus of Drosophila lead to multiple free centrosomes in syncytial embryos, but restrict centrosome duplication in larval neuroblasts." *J Cell Sci*, 96 (Pt 4), 605-16.

Gorlin, R. J., Cervenka, J., Moller, K., Horrobin, M., and Witkop, C. J., Jr. (1975). "Malformation syndromes. A selected miscellany." *Birth Defects Orig Artic Ser*, 11(2), 39-50.

Goto, H., Tomono, Y., Ajiro, K., Kosako, H., Fujita, M., Sakurai, M., Okawa, K., Iwamatsu, A., Okigaki, T., Takahashi, T., and Inagaki, M. (1999). "Identification of a novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation." *J Biol Chem*, 274(36), 25543-9.

Gotz, M., and Huttner, W. B. (2005). "The cell biology of neurogenesis." *Nat Rev Mol Cell Biol*, 6(10), 777-88.

Graser, S., Stierhof, Y. D., Lavoie, S. B., Gassner, O. S., Lamla, S., Le Clech, M., and Nigg, E. A. (2007). "Cep164, a novel centriole appendage protein required for primary cilium formation." *J Cell Biol*, 179(2), 321-30.

Grimes, J. E., and Grimes, B. F. (1993). *Ethnologue : language family index*, Dalles, Tex.: Summer Institute of Linguistics.

Gruber, R., Zhou, Z., Sukchev, M., Joerss, T., Frappart, P. O., and Wang, Z. Q. (2011). "MCPH1 regulates the neuroprogenitor division mode by coupling the centrosomal cycle with mitotic entry through the Chk1-Cdc25 pathway." *Nat Cell Biol*, 13(11), 1325-34.

Gudbjartsson, D. F., Jonasson, K., Frigge, M. L., and Kong, A. (2000). "Allegro, a new computer program for multipoint linkage analysis." *Nat Genet*, 25(1), 12-3.

Gul, A., Hassan, M. J., Mahmood, S., Chen, W., Rahmani, S., Naseer, M. I., Dellefave, L., Muhammad, N., Rafiq, M. A., Ansar, M., Chishti, M. S., Ali, G., Siddique, T., and Ahmad, W. (2006). "Genetic studies of autosomal recessive primary microcephaly in 33 Pakistani families: Novel sequence variants in ASPM gene." *Neurogenetics*, 7(2), 105-10.

Gul, A., Tariq, M., Khan, M. N., Hassan, M. J., Ali, G., and Ahmad, W. (2007). "Novel protein-truncating mutations in the ASPM gene in families with autosomal recessive primary microcephaly." *J Neurogenet*, 21(3), 153-63.

Habedanck, R., Stierhof, Y. D., Wilkinson, C. J., and Nigg, E. A. (2005). "The Polo kinase Plk4 functions in centriole duplication." *Nat Cell Biol*, 7(11), 1140-6.

Haber, M., Platt, D. E., Ashrafian Bonab, M., Youhanna, S. C., Soria-Hernanz, D. F., Martinez-Cruz, B., Douaihy, B., Ghassibe-Sabbagh, M., Rafatpanah, H., Ghanbari, M., Whale, J., Balanovsky, O., Wells, R. S., Comas, D., Tyler-Smith, C., Zalloua, P. A., and Genographic, C. (2012). "Afghanistan's ethnic groups share a Y-chromosomal heritage structured by historical events." *PLoS One*, 7(3), e34288.

Hafeez, M., Aslam, M., Ali, A., Rashid, Y., and Jafri, H. (2007). "Regional and ethnic distribution of beta thalassemia mutations and effect of consanguinity in patients referred for prenatal diagnosis." *J Coll Physicians Surg Pak*, 17(3), 144-7.

Hall, J. G. (2013). "The smallest of the small." *Gene*, 528(1), 55-7.

Hall, J. G., Flora, C., Scott, C. I., Jr., Pauli, R. M., and Tanaka, K. I. (2004). "Majewski osteodysplastic primordial dwarfism type II (MOPD II): natural history and clinical findings." *Am J Med Genet A*, 130A(1), 55-72.

Harper, P. S. (2005). "William Bateson, human genetics and medicine." *Hum Genet*, 118(1), 141-51.

Holland, A. J., Fachinetti, D., Da Cruz, S., Zhu, Q., Vitre, B., Lince-Faria, M., Chen, D., Parish, N., Verma, I. M., Bettencourt-Dias, M., and Cleveland, D. W. (2012a). "Polo-like kinase 4 controls centriole duplication but does not directly regulate cytokinesis." *Mol Biol Cell*, 23(10), 1838-45.

Holland, A. J., Fachinetti, D., Zhu, Q., Bauer, M., Verma, I. M., Nigg, E. A., and Cleveland, D. W. (2012b). "The autoregulated instability of Polo-like kinase 4 limits centrosome duplication to once per cell cycle." *Genes Dev*, 26(24), 2684-9.

Holland, A. J., Lan, W., and Cleveland, D. W. (2010). "Centriole duplication: A lesson in self-control." *Cell Cycle*, 9(14), 2731-6.

Hossain, M., and Stillman, B. (2012). "Meier-Gorlin syndrome mutations disrupt an Orc1 CDK inhibitory domain and cause centrosome reduplication." *Genes Dev*, 26(16), 1797-810.

Hu, W. F., Chahrour, M. H., and Walsh, C. A. (2014). "The diverse genetic landscape of neurodevelopmental disorders." *Annu Rev Genomics Hum Genet*, 15, 195-213.

Hudson, J. W., Kozarova, A., Cheung, P., Macmillan, J. C., Swallow, C. J., Cross, J. C., and Dennis, J. W. (2001). "Late mitotic failure in mice lacking Sak, a polo-like kinase." *Curr Biol*, 11(6), 441-6.

Hussain, M. S., Baig, S. M., Neumann, S., Nurnberg, G., Farooq, M., Ahmad, I., Alef, T., Hennies, H. C., Technau, M., Altmuller, J., Frommolt, P., Thiele, H., Noegel, A. A., and Nurnberg, P. (2012). "A truncating mutation of CEP135 causes primary microcephaly and disturbed centrosomal function." *Am J Hum Genet*, 90(5), 871-8.

Hussain, M. S., Baig, S. M., Neumann, S., Peche, V. S., Szczepanski, S., Nurnberg, G., Tariq, M., Jameel, M., Khan, T. N., Fatima, A., Malik, N. A., Ahmad, I., Altmuller, J., Frommolt, P., Thiele, H., Hohne, W., Yigit, G., Wollnik, B., Neubauer, B. A., Nurnberg, P., and Noegel, A. A. (2013). "CDK6 associates with the centrosome during mitosis and is mutated in a large Pakistani family with primary microcephaly." *Hum Mol Genet*, 22(25), 5199-214.

Hussain, R. (1999). "Community perceptions of reasons for preference for consanguineous marriages in Pakistan." *J Biosoc Sci*, 31(4), 449-61.

Ibbetson, D. (1982). Panjab castes : being a reprint of the chapter on "The races, castes, and tribes of the people" in the report on the census of the Panjab published in 1883 by the late Sir Denzil Ibbetson*, K.C.S.I*, Lahore: Sh. Mubarak Ali.

Issa, L., Mueller, K., Seufert, K., Kraemer, N., Rosenkotter, H., Ninnemann, O., Buob, M., Kaindl, A. M., and Morris-Rosendahl, D. J. (2013). "Clinical and cellular features in patients with primary autosomal recessive microcephaly and a novel CDK5RAP2 mutation." *Orphanet J Rare Dis*, 8, 59.

Jackson, A. P., Eastwood, H., Bell, S. M., Adu, J., Toomes, C., Carr, I. M., Roberts, E., Hampshire, D. J., Crow, Y. J., Mighell, A. J., Karbani, G., Jafri, H., Rashid, Y., Mueller, R. F., Markham, A. F., and Woods, C. G. (2002). "Identification of microcephalin, a protein implicated in determining the size of the human brain." *Am J Hum Genet*, 71(1), 136-42.

Jackson, A. P., McHale, D. P., Campbell, D. A., Jafri, H., Rashid, Y., Mannan, J., Karbani, G., Corry, P., Levene, M. I., Mueller, R. F., Markham, A. F., Lench, N. J., and Woods, C. G. (1998). "Primary autosomal recessive microcephaly (MCPH1) maps to chromosome 8p22-pter." *American Journal of Human Genetics*, 63(2), 541-546.

Jamieson, C. R., Fryns, J. P., Jacobs, J., Matthijs, G., and Abramowicz, M. J. (2000). "Primary autosomal recessive microcephaly: MCPH5 maps to 1q25-q32." *Am J Hum Genet*, 67(6), 1575-7.

Jana, S. C., Bazan, J. F., and Bettencourt-Dias, M. (2012). "Polo boxes come out of the crypt: a new view of PLK function and evolution." *Structure*, 20(11), 1801-4.

Kalay, E., Yigit, G., Aslan, Y., Brown, K. E., Pohl, E., Bicknell, L. S., Kayserili, H., Li, Y., Tuysuz, B., Nurnberg, G., Kiess, W., Koegl, M., Baessmann, I., Buruk, K., Toraman, B., Kayipmaz, S., Kul, S., Ikbal, M., Turner, D. J., Taylor, M. S., Aerts, J., Scott, C., Milstein, K., Dollfus, H., Wieczorek, D., Brunner, H. G., Hurles, M., Jackson, A. P., Rauch, A., Nurnberg, P., Karaguzel, A., and Wollnik, B. (2011). "CEP152 is a genome maintenance protein disrupted in Seckel syndrome." *Nature Genetics*, 43(1), 23-26.

Kennedy, M. A. (2001). "Mendelian Genetic Disorders." *ENCYCLOPEDIA OF LIFE SCIENCES*.

Khan, M. A., Rupp, V. M., Orpinell, M., Hussain, M. S., Altmuller, J., Steinmetz, M. O., Enzinger, C., Thiele, H., Hohne, W., Nurnberg, G., Baig, S. M., Ansar, M., Nurnberg, P., Vincent, J. B., Speicher, M. R., Gonczy, P., and Windpassinger, C. (2014). "A missense mutation in the PISA domain of HsSAS-6 causes autosomal recessive primary microcephaly in a large consanguineous Pakistani family." *Hum Mol Genet*, 23(22), 5940-9.

Kilmartin, J. V., Wright, B., and Milstein, C. (1982). "Rat monoclonal antitubulin antibodies derived by using a new nonsecreting rat cell line." *J Cell Biol*, 93(3), 576-82.

Klebba, J. E., Buster, D. W., McLamarrah, T. A., Rusan, N. M., and Rogers, G. C. (2015a). "Autoinhibition and relief mechanism for Polo-like kinase 4." *Proc Natl Acad Sci U S A*, 112(7), E657-66.

Klebba, J. E., Buster, D. W., Nguyen, A. L., Swatkoski, S., Gucek, M., Rusan, N. M., and Rogers, G. C. (2013). "Polo-like kinase 4 autodestructs by generating its Slimb-binding phosphodegron." *Curr Biol*, 23(22), 2255-61.

Klebba, J. E., Galletta, B. J., Nye, J., Plevock, K. M., Buster, D. W., Hollingsworth, N. A., Slep, K. C., Rusan, N. M., and Rogers, G. C. (2015b). "Two Polo-like kinase 4 binding domains in Asterless perform distinct roles in regulating kinase stability." *J Cell Biol*, 208(4), 401-14.

Kleylein-Sohn, J., Westendorf, J., Le Clech, M., Habedanck, R., Stierhof, Y. D., and Nigg, E. A. (2007). "Plk4-induced centriole biogenesis in human cells." *Dev Cell*, 13(2), 190-202.

Klingseisen, A., and Jackson, A. P. (2011). "Mechanisms and pathways of growth failure in primordial dwarfism." *Genes Dev*, 25(19), 2011-24.

Ko, M. A., Rosario, C. O., Hudson, J. W., Kulkarni, S., Pollett, A., Dennis, J. W., and Swallow, C. J. (2005). "Plk4 haploinsufficiency causes mitotic infidelity and carcinogenesis." *Nat Genet*, 37(8), 883-8.

Komai, T., Kishimoto, K., and Ozaki, Y. (1955). "Genetic study of microcephaly based on Japanese material." *Am J Hum Genet*, 7(1), 51-65.

Kousar, R., Nawaz, H., Khurshid, M., Ali, G., Khan, S. U., Mir, H., Ayub, M., Wali, A., Ali, N., Jelani, M., Basit, S., Ahmad, W., and Ansar, M. (2010). "Mutation analysis of the ASPM gene in 18 Pakistani families with autosomal recessive primary microcephaly." *J Child Neurol*, 25(6), 715-20.

Kumar, A., Blanton, S. H., Babu, M., Markandaya, M., and Girimaji, S. C. (2004). "Genetic analysis of primary microcephaly in Indian families: novel ASPM mutations." *Clin Genet*, 66(4), 341-8.

Kumar, A., Girimaji, S. C., Duvvari, M. R., and Blanton, S. H. (2009). "Mutations in STIL, encoding a pericentriolar and centrosomal protein, cause primary microcephaly." *Am J Hum Genet*, 84(2), 286-90.

Lancaster, M. A., and Knoblich, J. A. (2012). "Spindle orientation in mammalian cerebral cortical development." *Curr Opin Neurobiol*, 22(5), 737-46.

Lancaster, M. A., Renner, M., Martin, C. A., Wenzel, D., Bicknell, L. S., Hurles, M. E., Homfray, T., Penninger, J. M., Jackson, A. P., and Knoblich, J. A. (2013). "Cerebral organoids model human brain development and microcephaly." *Nature*, 501(7467), 373-9.

Leung, G. C., Hudson, J. W., Kozarova, A., Davidson, A., Dennis, J. W., and Sicheri, F. (2002). "The Sak polo-box comprises a structural domain sufficient for mitotic subcellular localization." *Nature Structural Biology*, 9(10), 719-724.

Lin, S. Y., Rai, R., Li, K. Y., Xu, Z. X., and Elledge, S. J. (2005). "BRIT1/MCPH1 is a DNA damage responsive protein that regulates the Brca1-Chk1 pathway, implicating checkpoint dysfunction in microcephaly." *Proc Natl Acad Sci U S A*, 102(42), 15105-15109.

Lizarraga, S. B., Margossian, S. P., Harris, M. H., Campagna, D. R., Han, A. P., Blevins, S., Mudbhary, R., Barker, J. E., Walsh, C. A., and Fleming, M. D. (2010a). "Cdk5rap2 regulates centrosome function and chromosome segregation in neuronal progenitors." *Development*, 137(11), 1907-1917.

Lizarraga, S. B., Margossian, S. P., Harris, M. H., Campagna, D. R., Han, A. P., Blevins, S., Mudbhary, R., Barker, J. E., Walsh, C. A., and Fleming, M. D. (2010b). "Cdk5rap2 regulates centrosome function and chromosome segregation in neuronal progenitors." *Development*, 137(11), 1907-17.

Mahmood, S., Ahmad, W., and Hassan, M. J. (2011). "Autosomal Recessive Primary Microcephaly (MCPH): clinical manifestations, genetic heterogeneity and mutation continuum." *Orphanet J Rare Dis*, 6, 39.

Majewski, F., Stoeckenius, M., and Kemperdick, H. (1982). "Studies of microcephalic primordial dwarfism III: an intrauterine dwarf with platyspondyly and anomalies of pelvis and clavicles--osteodysplastic primordial dwarfism type III." *Am J Med Genet*, 12(1), 37-42.

Mardis, E. R. (2008). "Next-generation DNA sequencing methods." *Annu Rev Genomics Hum Genet*, 9, 387-402.

Maridor, G., Gallant, P., Golsteyn, R., and Nigg, E. A. (1993). "Nuclear localization of vertebrate cyclin A correlates with its ability to form complexes with cdk catalytic subunits." *J Cell Sci*, 106 (Pt 2), 535-44.

Marthiens, V., Rujano, M. A., Pennetier, C., Tessier, S., Paul-Gilloteaux, P., and Basto, R. (2013). "Centrosome amplification causes microcephaly." *Nat Cell Biol*, 15(7), 731-40.

Martin, C. A., Ahmad, I., Klingseisen, A., Hussain, M. S., Bicknell, L. S., Leitch, A., Nurnberg, G., Toliat, M. R., Murray, J. E., Hunt, D., Khan, F., Ali, Z., Tinschert, S., Ding, J., Keith, C., Harley, M. E., Heyn, P., Muller, R., Hoffmann, I., Daire, V. C., Dollfus, H., Dupuis, L., Bashamboo, A., McElreavey, K., Kariminejad, A., Mendoza-Londono, R., Moore, A. T., Saggar, A., Schlechter, C., Weleber, R., Thiele, H., Altmuller, J., Hohne, W., Hurles, M. E., Noegel, A. A., Baig, S. M., Nurnberg, P., and Jackson, A. P. (2014). "Mutations in PLK4, encoding a master regulator of centriole biogenesis, cause microcephaly, growth failure and retinopathy." *Nat Genet*, 46(12), 1283-92.

Matise, T. C., Sachidanandam, R., Clark, A. G., Kruglyak, L., Wijsman, E., Kakol, J., Buyske, S., Chui, B., Cohen, P., de Toma, C., Ehm, M., Glanowski, S., He, C., Heil, J., Markianos, K., McMullen, I., Pericak-Vance, M. A., Silbergleit, A., Stein, L., Wagner, M., Wilson, A. F., Winick, J. D., Winn-Deen, E. S., Yamashiro, C. T., Cann, H. M., Lai, E., and Holden, A. L. (2003). "A 3.9-centimorgan-resolution human single-nucleotide polymorphism linkage map and screening set." *Am J Hum Genet*, 73(2), 271-84.

References

McIntyre, R. E., Lakshminarasimhan Chavali, P., Ismail, O., Carragher, D. M., Sanchez-Andrade, G., Forment, J. V., Fu, B., Del Castillo Velasco-Herrera, M., Edwards, A., van der Weyden, L., Yang, F., Sanger Mouse Genetics, P., Ramirez-Solis, R., Estabel, J., Gallagher, F. A., Logan, D. W., Arends, M. J., Tsang, S. H., Mahajan, V. B., Scudamore, C. L., White, J. K., Jackson, S. P., Gergely, F., and Adams, D. J. (2012). "Disruption of mouse Cenpj, a regulator of centriole biogenesis, phenocopies Seckel syndrome." *PLoS Genet*, 8(11), e1003022.

McKusick, V. (1955). "Primordial dwarfism and ectopia lentis." *Am J Hum Genet*, 7(2), 189-98.

Mehdi, S. Q., Qamar, R., Ayub, Q., Khaliq, S., Mansoor, A., Ismail., M., Hammer, M. F., Underhill, P. A., and Cavalli-Sforza, L. L. (1999). "The origin of Pakistani populations: evidence from Y chromosome markers", in S. S. Herausgeber: Papiha, Deka, Ranjan, Chakraborty, Ranajit (Eds.), (ed.), *Genomic Diversity*. New York: Springer US, pp. 83-90.

Moynihan, L., Jackson, A. P., Roberts, E., Karbani, G., Lewis, I., Corry, P., Turner, G., Mueller, R. F., Lench, N. J., and Woods, C. G. (2000). "A third novel locus for primary autosomal recessive microcephaly maps to chromosome 9q34." *American Journal of Human Genetics*, 66(2), 724-727.

Muhammad, F., Mahmood Baig, S., Hansen, L., Sajid Hussain, M., Anjum Inayat, I., Aslam, M., Anver Qureshi, J., Toilat, M., Kirst, E., Wajid, M., Nurnberg, P., Eiberg, H., Tommerup, N., and Kjaer, K. W. (2009). "Compound heterozygous ASPM mutations in Pakistani MCPH families." *Am J Med Genet A*, 149A(5), 926-30.

Naidoo, N., Pawitan, Y., Soong, R., Cooper, D. N., and Ku, C. S. (2011). "Human genetics and genomics a decade after the release of the draft sequence of the human genome." *Hum Genomics*, 5(6), 577-622.

Neitzel, H., Neumann, L. M., Schindler, D., Wirges, A., Tonnies, H., Trimborn, M., Krebsova, A., Richter, R., and Sperling, K. (2002). "Premature chromosome condensation in humans associated with microcephaly and mental retardation: a novel autosomal recessive condition." *Am J Hum Genet*, 70(4), 1015-22.

Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E. E., Bamshad, M., Nickerson, D. A., and Shendure, J. (2009). "Targeted capture and massively parallel sequencing of 12 human exomes." *Nature*, 461(7261), 272-6.

Nicholas, A. K., Khurshid, M., Desir, J., Carvalho, O. P., Cox, J. J., Thornton, G., Kausar, R., Ansar, M., Ahmad, W., Verloes, A., Passemard, S., Misson, J. P., Lindsay, S., Gergely, F., Dobyns, W. B., Roberts, E., Abramowicz, M., and Woods, C. G. (2010). "WDR62 is associated with the spindle pole and is mutated in human microcephaly." *Nat Genet*, 42(11), 1010-4.

Nicholas, A. K., Swanson, E. A., Cox, J. J., Karbani, G., Malik, S., Springell, K., Hampshire, D., Ahmed, M., Bond, J., Di Benedetto, D., Fichera, M., Romano, C., Dobyns, W. B., and Woods, C. G. (2009). "The molecular landscape of ASPM mutations in primary microcephaly." *J Med Genet*, 46(4), 249-53.

Noctor, S. C., Martinez-Cerdeno, V., Ivic, L., and Kriegstein, A. R. (2004). "Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases." *Nat Neurosci*, 7(2), 136-44.

Novorol, C., Burkhardt, J., Wood, K. J., Iqbal, A., Roque, C., Coutts, N., Almeida, A. D., He, J., Wilkinson, C. J., and Harris, W. A. (2013). "Microcephaly models in the developing

zebrafish retinal neuroepithelium point to an underlying defect in metaphase progression." *Open Biol*, 3(10), 130065.

O'Connell, K. F., Caron, C., Kopish, K. R., Hurd, D. D., Kemphues, K. J., Li, Y., and White, J. G. (2001). "The C. elegans zyg-1 gene encodes a regulator of centrosome duplication with distinct maternal and paternal roles in the embryo." *Cell*, 105(4), 547-58.

Ogi, T., Walker, S., Stiff, T., Hobson, E., Limsirichaikul, S., Carpenter, G., Prescott, K., Suri, M., Byrd, P. J., Matsuse, M., Mitsutake, N., Nakazawa, Y., Vasudevan, P., Barrow, M., Stewart, G. S., Taylor, A. M., O'Driscoll, M., and Jeggo, P. A. (2012). "Identification of the first ATRIP-deficient patient and novel mutations in ATR define a clinical spectrum for ATR-ATRIP Seckel Syndrome." *PLoS Genet*, 8(11), e1002945.

Oldham, S., Bohni, R., Stocker, H., Brogiolo, W., and Hafen, E. (2000). "Genetic control of size in Drosophila." *Philos Trans R Soc Lond B Biol Sci*, 355(1399), 945-52.

Opitz, J. M., and Holt, M. C. (1990). "Microcephaly: general considerations and aids to nosology." *J Craniofac Genet Dev Biol*, 10(2), 175-204.

Passemard, S., Kaindl, A. M., and Verloes, A. (2013). "Microcephaly." *Handb Clin Neurol*, 111, 129-41.

Petrosino, J. F., Highlander, S., Luna, R. A., Gibbs, R. A., and Versalovic, J. (2009). "Metagenomic pyrosequencing and microbial identification." *Clin Chem*, 55(5), 856-66.

Pfau, R. B., Thrush, D. L., Hamelberg, E., Bartholomew, D., Botes, S., Pastore, M., Tan, C., del Gaudio, D., Gastier-Foster, J. M., and Astbury, C. (2013). "MCPH1 deletion in a newborn with severe microcephaly and premature chromosome condensation." *Eur J Med Genet*, 56(11), 609-13.

Pulvers, J. N., Bryk, J., Fish, J. L., Wilsch-Brauninger, M., Arai, Y., Schreier, D., Naumann, R., Helppi, J., Habermann, B., Vogt, J., Nitsch, R., Toth, A., Enard, W., Paabo, S., and Huttner, W. B. (2010). "Mutations in mouse Aspm (abnormal spindle-like microcephaly associated) cause not only microcephaly but also major defects in the germline." *Proc Natl Acad Sci U S A*, 107(38), 16595-600.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., and Sham, P. C. (2007). "PLINK: a tool set for whole-genome association and population-based linkage analyses." *Am J Hum Genet*, 81(3), 559-75.

Quintyne, N. J., Reing, J. E., Hoffelder, D. R., Gollin, S. M., and Saunders, W. S. (2005). "Spindle multipolarity is prevented by centrosomal clustering." *Science*, 307(5706), 127-9.

Qvist, P., Huertas, P., Jimeno, S., Nyegaard, M., Hassan, M. J., Jackson, S. P., and Borglum, A. D. (2011). "CtIP Mutations Cause Seckel and Jawad Syndromes." *PLoS Genet*, 7(10), e1002310.

Rakic, P. (2009). "Evolution of the neocortex: a perspective from developmental biology." *Nat Rev Neurosci*, 10(10), 724-35.

Richards, M. W., Leung, J. W., Roe, S. M., Li, K., Chen, J., and Bayliss, R. (2010). "A pocket on the surface of the N-terminal BRCT domain of Mcph1 is required to prevent abnormal chromosome condensation." *J Mol Biol*, 395(5), 908-15.

Riparbelli, M. G., Callaini, G., Glover, D. M., and Avides Mdo, C. (2002). "A requirement for the Abnormal Spindle protein to organise microtubules of the central spindle for cytokinesis in Drosophila." *J Cell Sci*, 115(Pt 5), 913-22.

Ripoll, P., Pimpinelli, S., Valdivia, M. M., and Avila, J. (1985). "A cell division mutant of Drosophila with a functionally abnormal spindle." *Cell*, 41(3), 907-12.

Roberts, E., Hampshire, D. J., Pattison, L., Springell, K., Jafri, H., Corry, P., Mannon, J., Rashid, Y., Crow, Y., Bond, J., and Woods, C. G. (2002). "Autosomal recessive primary microcephaly: an analysis of locus heterogeneity and phenotypic variation." *J Med Genet*, 39(10), 718-21.

Roberts, E., Jackson, A. P., Carradice, A. C., Deeble, V. J., Mannan, J., Rashid, Y., Jafri, H., McHale, D. P., Markham, A. F., Lench, N. J., and Woods, C. G. (1999). "The second locus for autosomal recessive primary microcephaly (MCPH2) maps to chromosome 19q13.1-13.2." *Eur J Hum Genet*, 7(7), 815-20.

Ronaghi, M. (2001). "Pyrosequencing sheds light on DNA sequencing." *Genome Res*, 11(1), 3-11.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M., and Nyren, P. (1996). "Real-time DNA sequencing using detection of pyrophosphate release." *Analytical Biochemistry*, 242(1), 84-89.

Rosario, C. O., Ko, M. A., Haffani, Y. Z., Gladdy, R. A., Paderova, J., Pollett, A., Squire, J. A., Dennis, J. W., and Swallow, C. J. (2010). "Plk4 is required for cytokinesis and maintenance of chromosomal stability." *Proc Natl Acad Sci U S A*, 107(15), 6888-93.

Sajid, H. M., Marriam Bakhtiar, S., Farooq, M., Anjum, I., Janzen, E., Reza Toliat, M., Eiberg, H., Kjaer, K. W., Tommerup, N., Noegel, A. A., Nurnberg, P., Baig, S. M., and Hansen, L. (2013). "Genetic heterogeneity in Pakistani microcephaly families." *Clin Genet*, 83(5), 446-51.

Salisbury, J. L., Suino, K. M., Busby, R., and Springett, M. (2002). "Centrin-2 is required for centriole duplication in mammalian cells." *Curr Biol*, 12(15), 1287-92.

Scherer, W. F., Syverton, J. T., and Gey, G. O. (1953). "Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix." *J Exp Med*, 97(5), 695-710.

Seckel, H. P. G. (1960). *Bird-headed dwarfs*, Springfield, Ill.,: C. C. Thomas.

Shaheen, R., Al Tala, S., Almoisheer, A., and Alkuraya, F. S. (2014a). "Mutation in PLK4, encoding a master regulator of centriole formation, defines a novel locus for primordial dwarfism." *J Med Genet*, 51(12), 814-6.

Shaheen, R., Faqeih, E., Ansari, S., Abdel-Salam, G., Al-Hassnan, Z. N., Al-Shidi, T., Alomar, R., Sogaty, S., and Alkuraya, F. S. (2014b). "Genomic analysis of primordial dwarfism reveals novel disease genes." *Genome Res*, 24(2), 291-9.

Sillibourne, J. E., and Bornens, M. (2010). "Polo-like kinase 4: the odd one out of the family." *Cell Div*, 5, 25.

Sir, J. H., Barr, A. R., Nicholas, A. K., Carvalho, O. P., Khurshid, M., Sossick, A., Reichelt, S., D'Santos, C., Woods, C. G., and Gergely, F. (2011). "A primary microcephaly protein complex forms a ring around parental centrioles." *Nat Genet*, 43(11), 1147-53.

Sonawane, N. D., Szoka, F. C., Jr., and Verkman, A. S. (2003). "Chloride accumulation and swelling in endosomes enhances DNA transfer by polyamine-DNA polyplexes." *J Biol Chem*, 278(45), 44826-31.

Stenson, P. D., Mort, M., Ball, E. V., Howells, K., Phillips, A. D., Thomas, N. S., and Cooper, D. N. (2009). "The Human Gene Mutation Database: 2008 update." *Genome Med*, 1(1), 13.

Stiff, T., Alagoz, M., Alcantara, D., Outwin, E., Brunner, H. G., Bongers, E. M., O'Driscoll, M., and Jeggo, P. A. (2013). "Deficiency in origin licensing proteins impairs cilia formation: implications for the aetiology of Meier-Gorlin syndrome." *PLoS Genet*, 9(3), e1003360.

Stiles, J., and Jernigan, T. L. (2010). "The basics of brain development." *Neuropsychol Rev*, 20(4), 327-48.

Sunkel, C. E., and Glover, D. M. (1988). "polo, a mitotic mutant of Drosophila displaying abnormal spindle poles." *J Cell Sci*, 89 (Pt 1), 25-38.

Syed, A. M. S., and Shehla, A. (2011). "Cultural Diversity in Pakistan: National vs Provincial " *Mediterranean Journal of Social Sciences*, Vol.2.

Tan, C. A., del Gaudio, D., Dempsey, M. A., Arndt, K., Botes, S., Reeder, A., and Das, S. (2014a). "Analysis of ASPM in an ethnically diverse cohort of 400 patient samples: perspectives of the molecular diagnostic laboratory." *Clin Genet*, 85(4), 353-8.

Tan, C. A., Topper, S., Ward Melver, C., Stein, J., Reeder, A., Arndt, K., and Das, S. (2014b). "The first case of CDK5RAP2-related primary microcephaly in a non-consanguineous patient identified by next generation sequencing." *Brain Dev*, 36(4), 351-5.

Thiele, H., and Nürnberg, P. (2005). "HaploPainter: a tool for drawing pedigrees with complex haplotypes." *Bioinformatics*, 21(8), 1730-2.

Thompson, S. L., and Compton, D. A. (2008). "Examining the link between chromosomal instability and aneuploidy in human cells." *J Cell Biol*, 180(4), 665-72.

Thornton, G. K., and Woods, C. G. (2009). "Primary microcephaly: do all roads lead to Rome?" *Trends Genet*, 25(11), 501-10.

Tibelius, A., Marhold, J., Zentgraf, H., Heilig, C. E., Neitzel, H., Ducommun, B., Rauch, A., Ho, A. D., Bartek, J., and Kramer, A. (2009). "Microcephalin and pericentrin regulate mitotic entry via centrosome-associated Chk1." *J Cell Biol*, 185(7), 1149-57.

Tolmie, J. L., McNay, M., Stephenson, J. B., Doyle, D., and Connor, J. M. (1987). "Microcephaly: genetic counselling and antenatal diagnosis after the birth of an affected child." *Am J Med Genet*, 27(3), 583-94.

Trimborn, M., Bell, S. M., Felix, C., Rashid, Y., Jafri, H., Griffiths, P. D., Neumann, L. M., Krebs, A., Reis, A., Sperling, K., Neitzel, H., and Jackson, A. P. (2004). "Mutations in microcephalin cause aberrant regulation of chromosome condensation." *Am J Hum Genet*, 75(2), 261-6.

Van Den Bosch, J. (1959). "Microcephaly in the Netherlands: a clinical and genetical study." *Ann Hum Genet*, 23(2), 91-116.

Vassilev, L. T., Tovar, C., Chen, S., Knezevic, D., Zhao, X., Sun, H., Heimbrook, D. C., and Chen, L. (2006). "Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1." *Proc Natl Acad Sci U S A*, 103(28), 10660-5.

Venkatesh, T., and Suresh, P. S. (2014). "Emerging roles of MCPH1: Expedition from primary microcephaly to cancer." *European Journal of Cell Biology*, 93(3), 98-105.

Verloes, A., Drunat, S., Gressens, P., and Passemard, S. (1993). "Primary Autosomal Recessive Microcephalies and Seckel Syndrome Spectrum Disorders", in R. A. Pagon, M. P. Adam, H. H. Ardinger, S. E. Wallace, A. Amemiya, L. J. H. Bean, T. D. Bird, C. R. Dolan, C. T. Fong, R. J. H. Smith, and K. Stephens, (eds.), *GeneReviews(R)*. Seattle (WA).

Wang, Z., Wu, T., Shi, L., Zhang, L., Zheng, W., Qu, J. Y., Niu, R., and Qi, R. Z. (2010). "Conserved motif of CDK5RAP2 mediates its localization to centrosomes and the Golgi complex." *J Biol Chem*, 285(29), 22658-65.

Wodarz, A., and Huttner, W. B. (2003). "Asymmetric cell division during neurogenesis in Drosophila and vertebrates." *Mech Dev*, 120(11), 1297-309.

Woods, C. G. (2004). "Human microcephaly." *Curr Opin Neurobiol*, 14(1), 112-7.

Woods, C. G., Bond, J., and Enard, W. (2005). "Autosomal recessive primary microcephaly (MCPH): a review of clinical, molecular, and evolutionary findings." *Am J Hum Genet*, 76(5), 717-28.

Woods, C. G., Cox, J., Springell, K., Hampshire, D. J., Mohamed, M. D., McKibbin, M., Stern, R., Raymond, F. L., Sandford, R., Malik Sharif, S., Karbani, G., Ahmed, M., Bond, J., Clayton, D., and Inglehearn, C. F. (2006). "Quantification of homozygosity in consanguineous individuals with autosomal recessive disease." *Am J Hum Genet*, 78(5), 889-96.

Xu, X. Z., Lee, J., and Stern, D. F. (2004). "Microcephalin is a DNA damage response protein involved in regulation of CHK1 and BRCA1." *Journal of Biological Chemistry*, 279(33), 34091-34094.

Yang, Y. J., Baltus, A. E., Mathew, R. S., Murphy, E. A., Evrony, G. D., Gonzalez, D. M., Wang, E. P., Marshall-Walker, C. A., Barry, B. J., Murn, J., Tatarakis, A., Mahajan, M. A., Samuels, H. H., Shi, Y., Golden, J. A., Mahajnah, M., Shenhav, R., and Walsh, C. A. (2012). "Microcephaly gene links trithorax and REST/NRSF to control neural stem cell proliferation and differentiation." *Cell*, 151(5), 1097-112.

Yu, T. W., Mochida, G. H., Tischfield, D. J., Sgaier, S. K., Flores-Sarnat, L., Sergi, C. M., Topcu, M., McDonald, M. T., Barry, B. J., Felie, J. M., Sunu, C., Dobyns, W. B., Folkerth, R. D., Barkovich, A. J., and Walsh, C. A. (2010). "Mutations in WDR62, encoding a centrosome-associated protein, cause microcephaly with simplified gyri and abnormal cortical architecture." *Nat Genet*, 42(11), 1015-20.

9. Appendix

Appendices

Appendix 1

List of microsatellite markers used for exclusion mapping. These markers spanning the first seven MCPH loci. Of note, the *CEP152* (MCPH4) locus name is now designated as MCPH9 (OMIM# 614852) according to the new nomenclature. See 5.5 for details.

Locus	Marker	PCR Prog	Dye	Repe-at unit	PCR product size (bp)	Position (cM)	Physical position (hg19)
MCPH1	D8S518	MS	Fam	Di	229-253	5.6	4487605-4487855
	D8S1798	MS	Tet	Di	145-165	6.7	5103286-5103608
	D8S1742	MS	Fam	Di	130-150	7.7	6214007-6214142
	D8S277	MS	Fam	Di	148-180	8.34	6516725-6516870
	D8S1819	TDM	Fam	Di	207-223	10.0	6749967-6750187
	D8S1825	MS	Hex	Di	127-143	15.4	8924866-8924998
MCPH2	D19S213	MS	Hex	Di	174-184	59.4	34111259-34111441
	D19S425	MS	Fam	Di	252-280	59.4	35493932-35494196
	M19SH14	MS	Hex	Di	293	61.0	36474261-36474287
	D19S224	MS	Tet	Di	240-262	61.5	36528072-36528329
	D19S896	MS	Fam	Di	194-220	62.0	37478642-37478839
	D19S570	MS	Fam	Di	186-210	62.0	37727743-37727942
MCPH3	D9S302	TDM	Fam	Tetra	232-316	123.3	117081115-117081372
	D9S934	MS	Hex	Tetra	206-230	128	121095754-121095977
	D9S1850	MS	Ned	Di	218-230	132.1	123450263-123450486
	D9S1682	MS	Hex	Di	200-208	132.1	124993185-124993386
	D9S1881	MS	Fam	Di	220-236	135.9	126979479-126979708
	D9S1825	MS	Ned	Di	127-145	136.5	127888132-127888262
MCPH4	D15S659	MS	Hex	Tetra	166-210	43.5	46374008-46374191
	D15S1006	MS	Fam	Di	208-228	44.9	47546963-47547176
	D15S992	MS	Fam	D1	256-274	45.6	48839742-48840011
	D15S126	TDM	Fam	Di	188-218	45.6	49302271-49302458
	D15S1016	MS	Hex	Di	239-267	47.3	53532829-53533077
	D15S962	MS	Fam	Di	282-294	47.9	56575068-56575351
MCPH5	D1S2816	MS	Fam	Di	248-252	211.1	196650545-196650792
	D1S1660	MS	Fam	Tetra	218-250	212.4	198611347-198611574

	D1S373	MS	Hex	Tetra	283-330	214.0	200254486-200254796
	D1S1723	MS	Hex	Di	167-181	215.1	201390888-201391058
	D1S2655	MS	Fam	Di	224-260	216.8	202565270-202565517
MCPH6	D13S787	MS	Hex	Tetra	231-271	8.87	24380738-24380989
	D13S742	MS	Hex	Tetra	364	10.7	25282814-25283171
	D13S283	MS	Fam	Di	128-153	11.5	25600985-25601113
	D13S1294	MS	Hex	Di	247-273	12.9	26476909-26477159
	D13S221	MS	Ned	Di	318-366	12.91	26576866-26577100
	D13S1304	MS	Hex	Di	149-165	13.45	27370064-27370222
MCPH7	D1S2797	MS	Fam	Di	148-180	75.7	46933592-46933758
	D1S2874	MS	Hex	Di	230-240	75.7	47908120- 47908484
	D1S2134	MS	Ned	Tetra	249-301	75.7	48281268-48281554
	D1S197	MS	Hex	Di	115-129	76.3	50750498-50750635
	D1S427	MS	Fam	Di	275-297	76.3	69213345-69213632

Appendix 2

Primers used for PCR and Sanger sequencing of MCPH1

MCPH1-Ex1F	CTCCCGCCTCACCTACAGAG
MCPH1-Ex1R	AGCAGTACGGGGGAGGAAG
MCPH1-Ex2F	GTCTCAAACCCCTGACTTCG
MCPH1-Ex2R	TCCTCCTCCTCACTCAATGC
MCPH1-Ex3F	TTCACCGTTGTAACTGGAACAG
MCPH1-Ex3R	TCCCCCTAGGCAGAGTTAGG
MCPH1-Ex4F	GGGAAGTTTGATTTATACTGACTTTTG
MCPH1-Ex4R	ACCACAGGCTTTTCCATTTC
MCPH1-Ex5F	TTGCCAGTTCACATACAGTGC
MCPH1-Ex5R	TATTATGGCTCCCAGCCAAG
MCPH1-Ex6F	CAACATCACTGCCTGTGGAG
MCPH1-Ex6R	CAAAGCCAGCCATGAAATAAG
MCPH1-Ex7F	CAGGCAAGTTGACTTTAAGATCC
MCPH1-Ex7R	GAGCAATGGATTCTGTAGCAAG
MCPH1-Ex8-1F	CCCTTAAGTGGAAATGAGAAGAAC
MCPH1-Ex8-1R	CAAACGATACTTCTCTTCAAACG
MCPH1-Ex8-2F	TCATCACCATCTTTCACTCACC
MCPH1-Ex8-2R	AACTGAGCAGGGCTTGATG
MCPH1-Ex8-3F	CTGGAGGCTCTTAGCTGTGG
MCPH1-Ex8-3R	CTTTTCAGACCAACCGCTTC
MCPH1-Ex8-4F	TTCGAGTTGCGTGACTTCTG

MCPH1-Ex8-4R	TCATTGGAATGATCAAATGCAC
MCPH1-Ex9F	TTTATTTAAAAGGCCTATAACTTCCTG
MCPH1-Ex9R	GCAAACAACGCTTTCAGTTTC
MCPH1-Ex10F	GCTTTGGGGACAGTATCTGAG
MCPH1-Ex10R	CTGCCTAAACAACCCCAGAG
MCPH1-Ex11F	TTTATTTCCCCAGGTTTCAAAG
MCPH1-Ex11R	GACCAACAGGAGGAAAGACG
MCPH1-Ex12F	TTCATAACATTTATGCAACATGAAG
MCPH1-Ex12R	AAAAGGTGTTAAGTTCTGTGAATGTC
MCPH1-Ex13F	TGTCATCATCTTCTCTGGATTCTC
MCPH1-Ex13R	CGTCTGCTAACAGCAAGGAG
MCPH1-Ex14F	TTGTATTGAATTTCGTTTCACG
MCPH1-Ex14R	TCTTAAGAACCACAACACATGG

Appendix 3

List of primers used for PCR amplification and sequencing of *ASPM*

ASPM-Ex1-1-F	GATTCCGCTACCGTTCTCAC
ASPM-Ex1-1-R	AAGGAGACCTGCAGAAGTGG
ASPM-Ex1-2-F	GAATTCACTCCCACGACCTC
ASPM-Ex1-2-R	AAGAGCCACCCACAGTTATTG
ASPM-Ex2-F	TTCTTATAATTAAGCAGATAGGGTAGG
ASPM-Ex2-R	GCAATATTACTTCAAGCCTGTTATC
ASPM-Ex3-1-F	ATGTTTTAACCATTCTGTGATTTACTC
ASPM-Ex3-1-R	ACGTACAGAGAGTGGCAAGC
ASPM-Ex3-2-F	CCCCCAACAGAAAACAATTC
ASPM-Ex3-2-R	GGGGATAAAATAGGATTAACTGACTC
ASPM-Ex3-3-F	AGCCTGTGCATTTGGAATC
ASPM-Ex3-3-R	ACCCACTGCACTGTTGAGAC
ASPM-Ex3-4-F	GAAGGCCACCTGTACCAGAG
ASPM-Ex3-4-R	AGGAAATGTACCCAGCAAATAAG
ASPM-Ex4-F	TCATAGCATTAGTGCGTGGAG
ASPM-Ex4-R	TTCTTCCAGGCTGTTATTCAAC
ASPM-Ex5-F	TTCAGTGTTTTAAAGATGGTATTGC
ASPM-Ex5-R	GCTAATGAACAGGGAATTATGC
ASPM-Ex6-F	TGAAATTGCATTTTATTGCTG
ASPM-Ex6-R	AAAAACACAAATCTCTTGAATG
ASPM-Ex7-F	TTTCCCACTGATATACTCTCCTTG
ASPM-Ex7-R	AAAAGATGAATCAAGCTAACCTAATG

ASPM-Ex8-F	TTATGGTCTGCCATTCCTCAG
ASPM-Ex8-R	GGAATGAGGGTGGAGGAAG
ASPM-Ex9-F	TTGCTACCCTACACTTTTTGTTTTAC
ASPM-Ex9-R	GATGGGACTCACCAGACAGG
ASPM-Ex10-F	CAGAATGATTTGGAGGATTTG
ASPM-Ex10-R	GACATAACATTGATGTACCACTTCC
ASPM-Ex11-F	ATGGAGCAACAGAGTTTTGAG
ASPM-Ex11-R	AAAAGAAAAGGTTGTCCATTAGC
ASPM-Ex12-F	AGAGGTAAAACTTATTCTCCTCACTG
ASPM-Ex12-R	GTTACTGGGGCAAAATAAACC
ASPM-Ex13-F	TCCATTTCAGGCACTTTATTTTC
ASPM-Ex13-R	AAGAAATTCAAGAAAAGACTTCAGG
ASPM-Ex14-F	CTGCCTTTATTATTCACAGGTATTC
ASPM-Ex14-R	CTATTATGCAGGAGGAAAGGAG
ASPM-Ex15-F	TCCCTTTTGGTTACTTTTCC
ASPM-Ex15-R	AAAAACCCACAAAAGATAAAAAC
ASPM-Ex16-F	AATTCAATAATAAAGTCTCAGAAGATG
ASPM-Ex16-R	ATACTCATACCTCCCCAACC
ASPM-Ex17-F	TTCCAAAGATGAACACAAGAAG
ASPM-Ex17-R	ACCATAACGAGCTTTTCAGG
ASPM-Ex18-1-F	TGAATTGGCTACAGGTATATCAATG
ASPM-Ex18-1-R	CTTGTGATTGCATTTTACGTTG
ASPM-Ex18-2-F	GTTACAATTCAGAGGCATTGG
ASPM-Ex18-2-R	CGCTCAATCTTTCCTTTCAG
ASPM-Ex18-3-F	AGAAAAGATTTCGGTGCTTTC
ASPM-Ex18-3-R	AGACTGCAGCACAATGACAG
ASPM-Ex18-4-F	TCATTTCCGAGCTTATATTTTG
ASPM-Ex18-4-R	AATACTGCCGAGCCTTTCTC
ASPM-Ex18-5-F	AGGATACCTTGTCCGAAAGC
ASPM-Ex18-5-R	AAATCACAGCTGCCTTTGTC
ASPM-Ex18-6-F	ACGCCAGCTAATCAAACAAC
ASPM-Ex18-6-R	CTTTTGTTGCACATGCATTC
ASPM-Ex18-7-F	TGTGGAAGGGAAAAACACTGAG
ASPM-Ex18-7-R	ATAAAGTGGCTGCCCTGTG
ASPM-Ex18-8-F	AAGACAGCAATTAAAATCCAATCTG
ASPM-Ex18-8-R	GAATGAGAGTTGCGGCTATATG
ASPM-Ex18-9-F	CAGCAAAGATACTGGGCAATG
ASPM-Ex18-9-R	ATGCTGCCTCTGCAGTTTTG

ASPM-Ex18-10-F	GGGCTGCTACTTTCATCCAG
ASPM-Ex18-10-R	GCATGTTTCCAAGTCTGAAATG
ASPM-Ex18-11-F	AAGGGCTGCAGTTCTCATTC
ASPM-Ex18-11-R	AATAATAATGGCAGCCTGGTG
ASPM-Ex18-12-F	CTTGTGTTCAGGCAGGTTTTC
ASPM-Ex18-12-R	GGCTGCCATCTTTTCCTCTG
ASPM-Ex18-13-F	GCTGCTTTTAGAGGCATGAAAG
ASPM-Ex18-13-R	TTGGTCAAAAGAAAGACTCACAG
ASPM-Ex19-F	ATGCAATTTATTTTTGTCTGCAC
ASPM-Ex19-R	CAGTACGAGAGATGTATTTTGTTTCC
ASPM-Ex20-F	ACTTCTTTCGTGTGCGTGTG
ASPM-Ex20-R	TGTGTGAAATAAATGCATACTTAGGTC
ASPM-Ex21-F	TGAAACATGGCATTCTTAGAC
ASPM-Ex21-R	TCAGTGCTCTTGTCACTTACC
ASPM-Ex22-F	TGATTAACTTAGGAGGCAGATACTTG
ASPM-Ex22-R	AACCCTCAAATCATTGTAACAGC
ASPM-Ex23-F	TGGACTGGAAAGGTTTTTGC
ASPM-Ex23-R	TGAGTTATTCTACCGGCTAATGC
ASPM-Ex24-25-F	GATGTAGATATAAATAGAAAACATTGG
ASPM-Ex24-25-R	CAGGGGCATATTTGTTGAC
ASPM-Ex26-F	AAAGTCCTTTGCACTTGCTG
ASPM-Ex26-R	TTTATCCGTGCAAAAAGCAG
ASPM-Ex27-F	GCGACAGAGCAAGAGAGACC
ASPM-Ex27-R	TGTTGTTTCTCCACTGAAAAGC
ASPM-Ex28-F	AAGGTGAAGGTAACTCTAAAGAGAAG
ASPM-Ex28-R	TTTTATGAAAATTCCTCAACTTACC

Appendix 4

Primers used for junction amplification and sequencing in *MCPH1*

|	
MCP118-F	ACTCATTCTGCTTGCTGCAC
MCP118-R	GGCAAATTGTTTGCTATCTCC
MCP125-F	CGATTTTGAAGGTATCTTTCTCC
MCP125-R	CTTTCAGAAACAAGAACCCAAG

Appendix 5

List of primers used for *WDR62* sequencing.

WDR62-AMP.1-F	GCCCCCATTGGTTCTAAGC
WDR62-AMP.2-F	GGGTGTTGAATGTAGCAGGAC
WDR62-AMP.3-F	ATTTGTGGGCTTTTCTGGTG
WDR62-AMP.4-F	AGGCAGAAGAGGGCAGAGTC
WDR62-AMP.5-F	GACTTGGAGTGGGGACGAG
WDR62-AMP.6-F	CCATTTGGAAGAGCTTCTTAGAG
WDR62-AMP.7-F	TGGGTGTCTTAGACCCAAGG
WDR62-AMP.8-F	GATGGATGGCCACAAATACC
WDR62-AMP.9-F	TTGCACGATACCTGTTTATCTTTAAC
WDR62-AMP.10-F	CTCTGCATCACTGCCTCTTG
WDR62-AMP.11-F	ATTGGTGAAACTCCCAGCTC
WDR62-AMP.12-F	TTCAACACGGCTCCCTTTAC
WDR62-AMP.13-F	GAACCAAAGACCCCTTCTCC
WDR62-AMP.14-F	TTGATTGATGGCTCTTGCAG
WDR62-AMP.15-F	TCAGCTTGAGATGGGGTTTC
WDR62-AMP.16-F	AGAGTCCCATGTGCTGTGC
WDR62-AMP.17-F	CAGGAGCATCTGGAGAATGG
WDR62-AMP.18-F	CACATGGAGCTTAGCACAGG
WDR62-AMP.19-F	CTTCTGACTTCTGGGTTGTGG
WDR62-AMP.20-F	CAGATGGGCTGTGTGAAATG
WDR62-AMP.21-F	CTGCTGGCATGGTTCCTG
WDR62-AMP.22-F	AGGGGGCATTTGGAGGAG
WDR62-AMP.23-F	TACCATCCCTCCTCCAGATG
WDR62-AMP.24-F	CCCACTGAGCCTGGAAGAC
WDR62-AMP.25-F	CTGCTCGTACCCTGTGTCTG
WDR62-AMP.26-F	CCGGATGTCCAGCCTAGC
WDR62-AMP.27-F	GGGACCCAGAGAAGTGTTCC
WDR62-AMP.1-R	GAACAACCTGAGGGGACTGG
WDR62-AMP.2-R	GAGCTGGGACACAAATCCAG
WDR62-AMP.3-R	CGTCTCCTCTGGTCACTGC
WDR62-AMP.4-R	TGCGTCCTCTCATGTAGCTC
WDR62-AMP.5-R	GAGAGTAGGCCCAGAAGCAG
WDR62-AMP.6-R	AACCTGTGTGGTCAATGCTG
WDR62-AMP.7-R	TGGGTATTTGTGGGATGGAC
WDR62-AMP.8-R	AGAGGGTTACTGCTGCTTGG

| WDR62-AMP.9-R	TTGCCCTCATGGAACCTAAC
WDR62-AMP.10-R	CTCAGCTCCCCAGAGACG
WDR62-AMP.11-R	TATTCACAGCCCCTCATTCC
WDR62-AMP.12-R	AACTCCCAGCTGTCATCCAC
WDR62-AMP.13-R	CGGCCCCAGTCATTATTCTG
WDR62-AMP.14-R	ATTCCTTCCCAGCTTCTTGG
WDR62-AMP.15-R	GGCACAGGTCTCAGGTGTG
WDR62-AMP.16-R	ACTCCCACAATCCAAACTGG
WDR62-AMP.17-R	TCAGCTGTGAGTGTGCTAAGG
WDR62-AMP.18-R	GCTGGGATTACAGGTATCAGC
WDR62-AMP.19-R	ACTGAGCCTGGAAGGATGAC
WDR62-AMP.20-R	GGAGTGGGAGTCAGCCTAAG
WDR62-AMP.21-R	CCACTGGGAAGGCAGAGAG
WDR62-AMP.22-R	CAGAATCCTCAGGCAGCAG
WDR62-AMP.23-R	CGTCTTCCAGGCTCAGTGG
WDR62-AMP.24-R	CTCTCCACTCACCCTCCTG
WDR62-AMP.25-R	CACAGAACTGATGGCTAGGG
WDR62-AMP.26-R	CGAATGAATGAATGGCACAG
WDR62-AMP.27-R	TCTGGGCATCACCTTCTACC

Appendix 6

Pedigrees of 14 consanguineous families in which *ASPM* is Sanger sequenced without performing linkage analysis. These families were mostly ascertained from the North West part of Pakistan
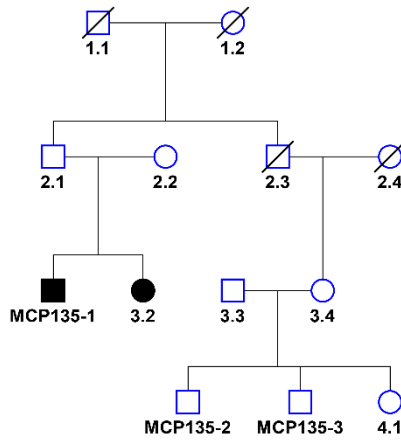

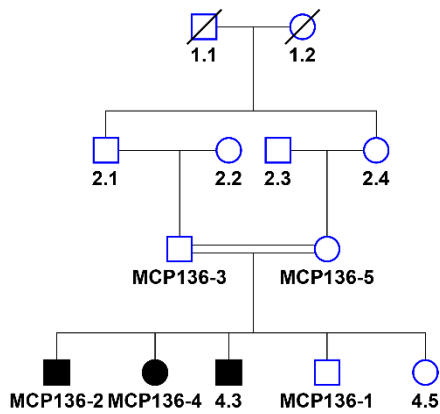**Family MCP103**
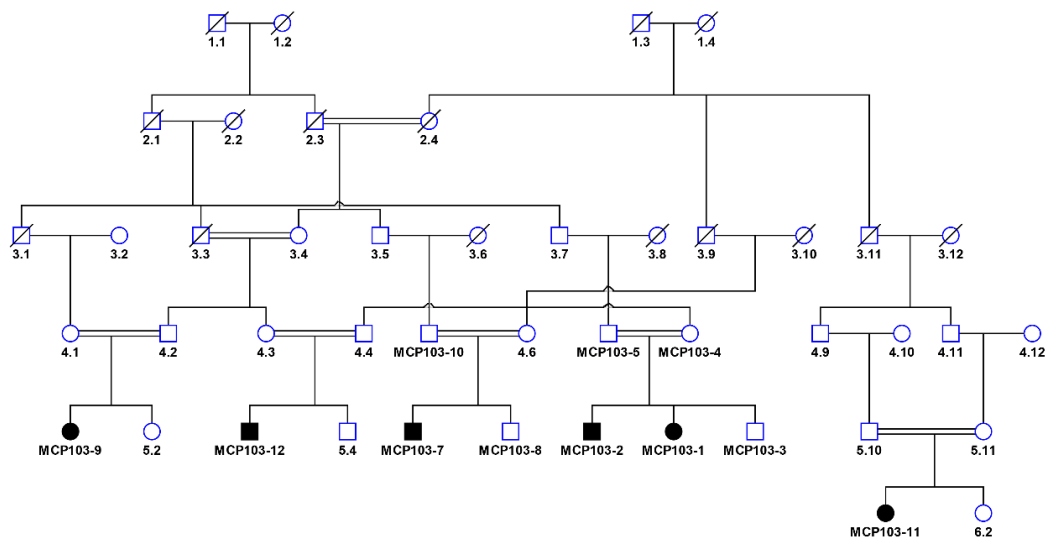
1. Pedigree MCP103
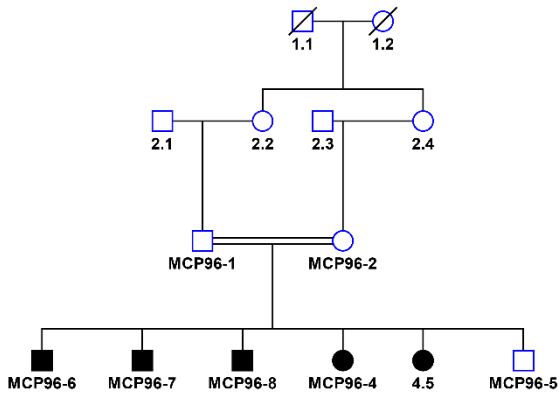
Family MCP96

Family MCP135

2. Pedigree MCP96
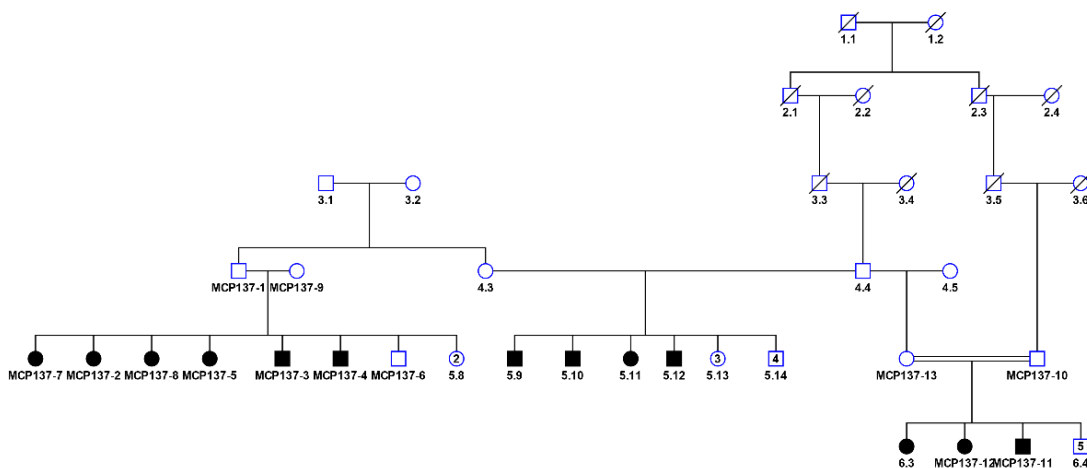
3. Pedigree MCP135

Family MCP136

Family MCP138
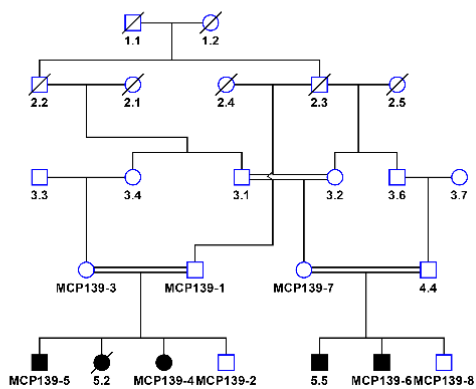
4. Pedigree MCP136

5. Pedigree MCP138
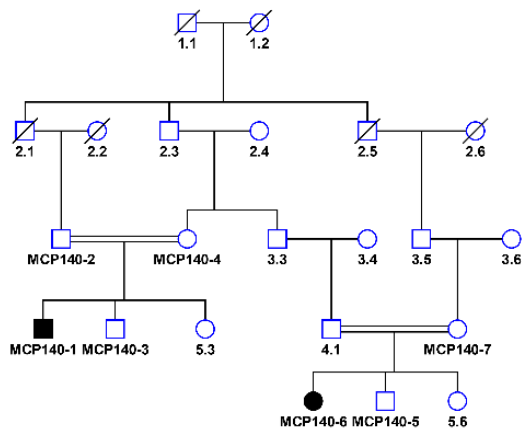
Family MCP137

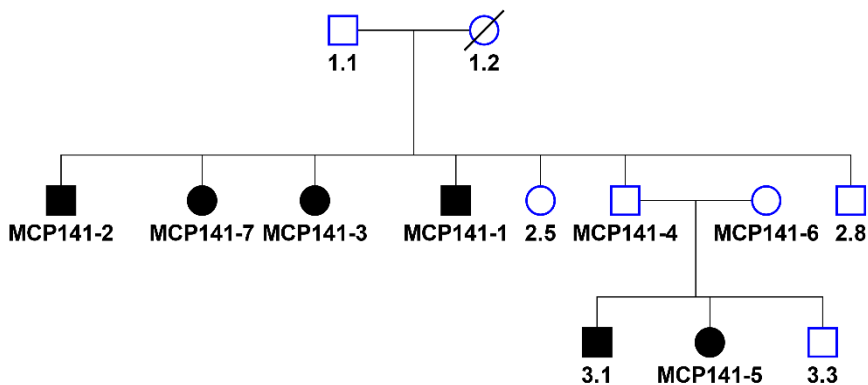## 6. Pedigree MCP137

Family MCP139

7. Pedigree MCP139

Family MCP140

8. Pedigree MCP140

Family MCP141

9. Pedigree MCP141

Family MCP143
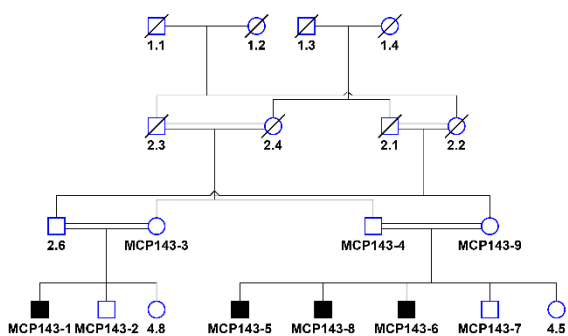
Family MCP174
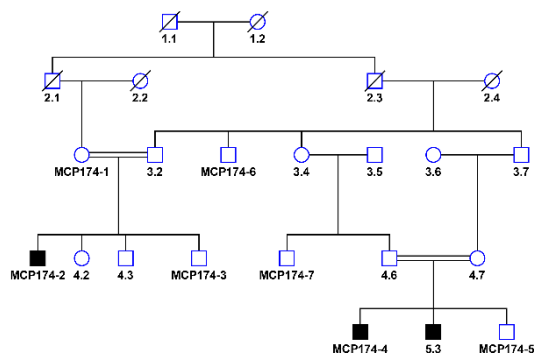
10. Pedigree MCP143   **11. Pedigree MCP174**

Family MCP175
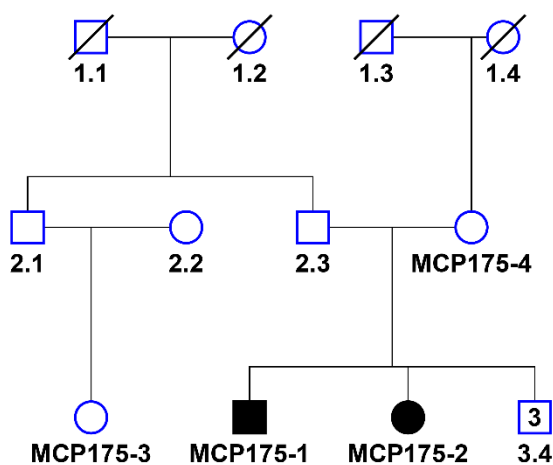
12. Pedigree MCP175

Family MCP176
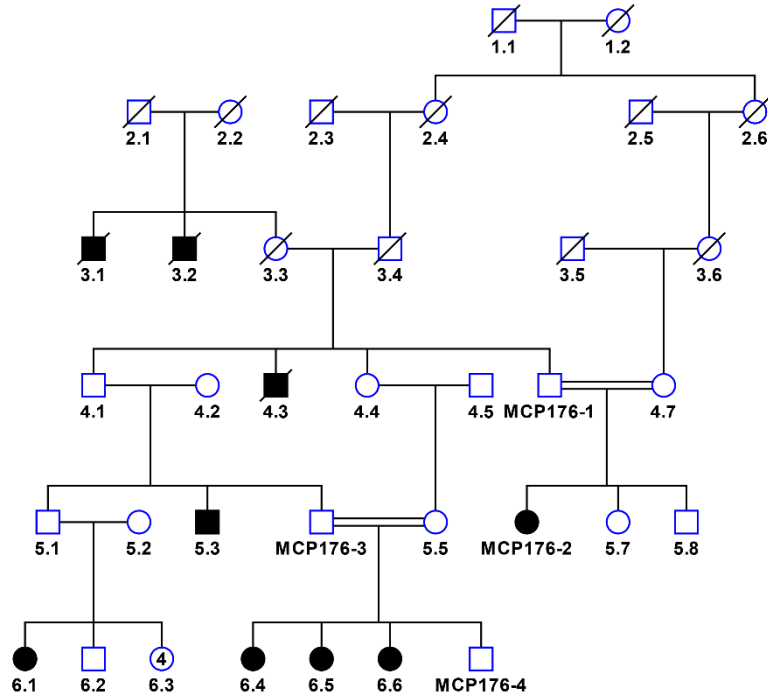
13. Pedigree MCP176

Family MCP177
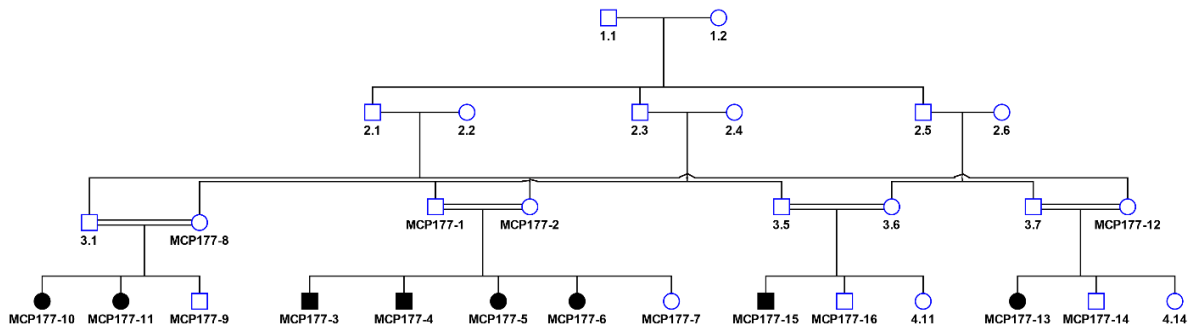
14. Pedigree MCP177

Appendix 7

Primer sequences related to pyrosequncing, RT-PCR, quantitative PCR and capillary electrophoresis in *PLK4*.

Pyrosequencing

PLK4-Ex16CT-F   BiotiN-TAATGTGCTTTTTGGGACTCTCA

PLK4-Ex16CT-R   AAGGATGGAAGACAGACACTGTAA

PLK4-Ex16CT-S   TTTCATTTTCTCCATACCT

RT-PCR and capillary electrophoresis

PLK4-CFr-F   FAM-AATGTTGGTTGGGCTACACAG

PLK4-CFr-R   CATCAAAAGGATGGAAGACAGAC

Sanger sequencing

PLK4-Ex16-F   ttttctgaggtgcctcttg
PLK4-EX16-R   CATCAAAAGGATGGAAGACAGAC

Quantitative PCR

GAPDH-F   CCAGGTGGTCTCCTCTGACTTC

GAPDH-R   GGTGTCGCTGTTGAAGTCAGA

PLK4-QT-F1   GGAGCTGTGTGGGTTCAGTT

PLK4-QT-R1   CATCAAAAGGATGGAAGACAGA

.

10. Abbreviations

Asp	Abnormal spindle gene
ASPM	Abnormal Spindle-like Microcephaly-associated
Asl	Asterless
APS	Adenosine 5' Phosphosulfate
ATR	Ataxia Telangiectasia Mutated
BRCT	Breast cancer 1(BRCA1) C-terminal
CDK5RAP2	Cyclin Dependent Kinase 5 Regulatory Associated Protein 2
CDK6	Cyclin-dependent Kinase 6
CEP135	Centrosomal protein 135 kDa
CENPJ	Centromere Protein J
CEP152	Centrosomal Protein of 152 kDa
CNV	Copy Number Variation
CPAP	Centrosomal protein 4.1- associated protein
CT	Computerized tomography
CMV	Cytomegalovirus
CPB	Cryptic Polo Box
DNA	Deoxyribonucleic acid
DAPI	4', 6-Diamidino-2'-phenylindole
dNTP	Deoxyribonucleotide Triphosphate
DDC1	DNA damage checkpoint 1
DMEM	Dulbecco´s Modified Eagle´s Medium
DMSO	Dimethyl Sulfoxide
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
FBS	Fetal bovine serum
γTuRC	γ-tubulin ring complex
HC	Head circumference
Kb	Kilobases
kDa	kilodaltons
LMW	Low Molecular Weight
LD	Linkage Disequilibrium
LOD	Logarithm of the odds
MCPH	Autosomal recessive primary microcephaly
MEF	Mouse Embryonic Fibroblasts
MGS	Meier Gorlin Syndrome

mRNA	Messenger Ribonucleic Acid
MRI	Magnetic resonance imaging
MCPH1	Autosomal recessive primary microcephaly locus 1
MCPH2	Autosomal recessive primary microcephaly locus 2
MCPH3	Autosomal recessive primary microcephaly locus 3
MCPH4	Autosomal recessive primary microcephaly locus 4
MCPH5	Autosomal recessive primary microcephaly locus 5
MCPH6	Autosomal recessive primary microcephaly locus 6
MCPH7	Autosomal recessive primary microcephaly locus 7
MCPH8	Autosomal recessive primary microcephaly locus 8
MCPH9	Autosomal recessive primary microcephaly locus 9
MCPH10	Autosomal recessive primary microcephaly locus 10
MCPH11	Autosomal recessive primary microcephaly locus 11
MCPH12	Autosomal recessive primary microcephaly locus 12
MCPH13	Autosomal recessive primary microcephaly locus 13
MOPD	microcephalic osteodysplastic primordial dwarfism
MPD	Microcephalic Primordial Dwarfism
MRI	Magnetic Resonance Imaging
MTOC	Microtubule Organising Centre
NCBI	National Center for Biotechnology Information
NE	Neuroepithelial
NGS	Next Generation Sequencing
NMR	Nuclear magnetic resonance
ORF	Open Reading Frame
OFC	Occipital Frontal Circumference
PB3	Polo Box Domain 3
PBS	Phosphate buffered saline
PCC	Premature Chromosome Condensation
PCM	Pericentriolar matrix
PCNT	Pericentrin
PCR	Polymerase chain reaction
PEI	polyethylenimine
PFA	paraformaldehyde
PIC	Proteinase Inhibitor Cocktail
Plk4	Polo-like-kinase 4
PKU	Phenylketonuria
PSQ	Pyromark Assay Design

RT-PCR	Reverse transcriptase PCR
SAP	Shrimp Alkaline Phosphatase
SD	Standard deviation
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
STR	Short tandem repeat
SNP	Single Nucleotide Polymorphism
SVZ	Subventricular zone
STIL/SIL	SCL/TAL1 interrupting locus
TBS	Tris-buffered saline
UV	Ultra-Violet
V	Volt
VZ	Ventricular zone
WDR62	WD repeat domain 62
WHO	World Health Organization

11. Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit- einschließlich Tabellen und Abbildungen-, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht ist, sowie, dass ich eine Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Herrn Prof. Dr. Peter Nürnberg betreut worden.

Köln, den June,2016

_____
Ilyas Ahmad

Nachfolgend genannte Teilpublikationen liegen vor:

* These authors contributed equally to this work

1.Martin CA*, **<u>Ahmad I</u>***, Klingseisen A*, Hussain MS*, Bicknell LS, Leitch A, Nürnberg G, Toliat MR, Murray JE, Hunt D, Khan F, Ali Z, Tinschert S, Ding J, Keith C, Harley ME, Heyn P, Müller R, Hoffmann I, Daire VC, Dollfus H, Dupuis L, Bashamboo A, McElreavey K, Kariminejad A, Mendoza-Londono R, Moore AT, Saggar A, Schlechter C, Weleber R, Thiele H, Altmüller J, Höhne W, Hurles ME, Noegel AA, Baig SM, Nürnberg P, Jackson AP: Mutations in *PLK4*, encoding a master regulator of centriole biogenesis, cause microcephaly, growth failure and retinopathy. *Nat Gene*t 2014. 3122, 1283–1292.

2. <u>Ahmad I</u>, Baig S.M, Abdulkareem A.R, Hussain M.S, Sur I, Toliat M.R, Nürnberg G, Dalibor N, Moawia A, Waseem S.S, Asif M, Hafsa, Sher M, Khan M.M, Hassan I, Rehman S, Thiele H, Altmüller J, Noegel A.A and Nürnberg P. Genetic heterogeneity in Pakistani microcephaly families revisited. *Clinical Genetics* 2016 (Submitted:CGE-00372-2016).

3.Salil K. Sukumaran, Maria Stumpf, Sarah Salamon, **Ilyas Ahmad**, Kurchi Bhattacharya, Sarah Fischer, Rolf Müller, Priyanka Kohli, Bernhard Schermer, Janine Altmüller, Birgit Budde, Holger Thiele, Peter Nürnberg, Shahid Mahmood Baig, Muhammad Sajid Hussain, Angelika A. Noegel. A role for CDK5RAP2 in the Hippo signaling pathway and its potential impact on primary microcephaly. *Human Genetics* 2016 (Submitted:HUGE-D-16-00190).