CHAPTER 1

General introduction
Fanconi anemia (FA) is a genetic disorder that is associated with mutations in one of the 16 known FA-related genes. The encoding FA proteins play a role in a DNA repair pathway for intrastrand and interstrand DNA cross-links. Groups of genes involved in maintaining the genetic information in the cell are called caretaker or DNA repair genes (Vogelstein and Kinzler 2004). Fanconi anemia is one of the several recessively inherited chromosomal instability syndromes that are associated with an increased cancer risk due to defects in DNA repair genes (Knoch et al. 2012). Fanconi anemia is named after Guido Fanconi who first characterized the disease as anemia combined with developmental abnormalities such as short stature and small head (microcephaly). In line with Fanconi's initial description, the great majority of patients with mutations in an FA-related gene exhibit hematological abnormalities. The variable clinical (and cellular) appearance of Fanconi anemia patients (and derived cells) cannot be explained by defects in DNA repair alone. This thesis will focus on exploring other biological networks besides DNA repair, in which the proteins encoded by FA-related genes are involved.

1. Fanconi anemia
The Swiss pediatrician Guido Fanconi described Fanconi anemia in 1927 (Fanconi 1927). It later became clear that the disease is a rare, often recessive genetically heterogeneous chromosomal instability disorder. Worldwide, the disease occurs in 1-5 individuals per million with a heterozygous mutation carrier frequency between 0.3 and 1% (Auerbach et al. 2001). Current preclinical research has so far identified 16 FA disease-causing genes, but since there are still, in rare cases, FA patients with an unidentified etiologic agent, this list is likely to grow. Nonetheless, the 16 known FA-related genes have led to the development of clinical protocols for detecting the disease-causing gene in the vast majority of FA patients.

In vitro research on the proteins expressed by the known FA-related genes has led to the unraveling of a DNA repair pathway, which is now intensively studied. However, recent research is having a wider scope since it has become increasingly clear that defects in DNA repair alone may not explain all the clinical features observed in FA patients or explain the specific characteristics of patient derived cells.

1.1 Clinical characteristics
FA patients exhibit a variable appearance. Developmental abnormalities range from skin pigmentation abnormalities (café-au-lait spots), skeletal abnormalities (absence or hypoplasia of radius and/or thumb), short stature, to microcephaly (Auerbach 2009; Shimamura and Alter 2010). Patients with FA show a high incidence of aplastic anemia,
myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML). Furthermore they have a higher risk for solid tumors, such as head-and-neck squamous cell carcinoma, gynecological squamous cell carcinoma, esophageal carcinoma, liver tumors, brain tumors, skin tumors, and renal tumors (Alter 2003; Rosenberg et al. 2003; Kutler et al. 2003). As a result the average life expectancy of FA patients is strongly reduced to about 20 years (range 0-50).

1.2 Cellular characteristics
The disorder is characterized at the cellular level by hypersensitivity to DNA cross-linking agents, such as mitomycin C (MMC), diepoxybutane, and cisplatin (Ishida and Buchwald 1982). These agents mainly generate intrastrand cross-links, connecting bases in the same strand, and a small fraction of interstrand cross-links, in which the antiparallel DNA strands in a double helix are covalently bound. Interstrand cross-links are considered the most toxic lesion, because they inhibit DNA strand separation and thereby block DNA replication, transcription, and segregation (Dronkert and Kanaar 2001). These cross-linking 'bifunctional alkylating' agents exacerbate the excessive chromosomal breakage levels typically present in cultured FA lymphocytes. The exclusive hypersensitivity to these cross-linking agents is a hallmark for FA and used to confirm the suspicion of FA in a diagnostic test (Oostra et al. 2012).

1.3 Complementation groups: the FA/BRCA pathway
The hypersensitivity of FA cells to for example, MMC as described above has been implemented in a ‘complementation analysis’ assay in preclinical research in order to identify the number of different FA (complementation) groups. In this assay two FA cell lines are fused and cultured as a tetraploid hybrid cell line, if the hybrid cell line appears still sensitive to MMC, the original two ‘parental’ FA cell lines belong to the same complementation group; if, however the hybrid cell line has become resistant to MMC (is complemented or corrected for the defect) they apparently belong to different complementation groups (Joenje and Patel 2001).

At present sixteen complementation groups have been found: FA-A, -B, -C, -D1, -D2, -E, -F, -G, -I, -J, -L, -M, -N, -O, -P, and -Q (see for references Table 1). Each individual complementation group corresponds to a distinct gene that if defective causes FA. The individual disease genes have been identified using various approaches, including genetic linkage analysis, candidate-gene-approaches, and massively parallel sequencing. The sixteen FA genes are spread over the genome. Most of them encode orphan proteins that appeared unrelated to any other protein and had no known functional domains (Blom et al. 2002) that could provide a clue to actual function. Mutations in BRCA1
<table>
<thead>
<tr>
<th>Subtype</th>
<th>FA gene</th>
<th>Chromosomal location</th>
<th>Exons</th>
<th>Protein size (aa)</th>
<th>Percentage Identity (aa) Human-Mouse</th>
<th>Domains</th>
<th>Protein structure</th>
<th>Cellular localization</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>FANCB</td>
<td>Xp22.2</td>
<td>9</td>
<td>859</td>
<td>49</td>
<td>-</td>
<td>-</td>
<td>N</td>
<td>(Meetei et al. 2003b; Meetei et al. 2004a)</td>
</tr>
<tr>
<td>C</td>
<td>FANCC</td>
<td>9q22.3</td>
<td>15</td>
<td>558</td>
<td>67</td>
<td>-</td>
<td>-</td>
<td>N/C</td>
<td>(Strathdee et al. 1992)</td>
</tr>
<tr>
<td>D1</td>
<td>FANCD1/BRCA2</td>
<td>13q12.3</td>
<td>28</td>
<td>3418</td>
<td>56</td>
<td>BRCA2 repeats</td>
<td>-</td>
<td>N</td>
<td>(Timmers et al. 2001; Howlett et al. 2002)</td>
</tr>
<tr>
<td>D2</td>
<td>FANCD2</td>
<td>3p26</td>
<td>44</td>
<td>1471</td>
<td>73</td>
<td>-</td>
<td>yes</td>
<td>N</td>
<td>(Whitney et al. 1995; Timmers et al. 2001; Joo et al. 2011)</td>
</tr>
<tr>
<td>E</td>
<td>FANCE</td>
<td>6p22-p21</td>
<td>10</td>
<td>536</td>
<td>65</td>
<td>-</td>
<td>-</td>
<td>N</td>
<td>(Waisfisz et al. 1999b; de Winter et al. 2000a)</td>
</tr>
<tr>
<td>F</td>
<td>FANCF</td>
<td>11p15</td>
<td>1</td>
<td>374</td>
<td>47</td>
<td>-</td>
<td>yes</td>
<td>N</td>
<td>(de Winter et al. 2000b; Kowal et al. 2007)</td>
</tr>
<tr>
<td>G</td>
<td>FANCG/XRCC9</td>
<td>9p13</td>
<td>14</td>
<td>622</td>
<td>72</td>
<td>TPR repeats</td>
<td>-</td>
<td>N/C</td>
<td>(Saar et al. 1998; de Winter et al. 1998; Waisfisz et al. 1999a)</td>
</tr>
<tr>
<td>I</td>
<td>FANCI/KIAA1794</td>
<td>15q25-26</td>
<td>38*</td>
<td>1328*</td>
<td>80</td>
<td>-</td>
<td>yes</td>
<td>N</td>
<td>(Dorsman et al. 2007; Smogorzewska et al. 2007; Sims et al. 2007; Joo et al. 2011)</td>
</tr>
<tr>
<td>J</td>
<td>FANCI/BRIP1</td>
<td>17q22-q24</td>
<td>20</td>
<td>1249</td>
<td>69</td>
<td>Helicase</td>
<td>-</td>
<td>N</td>
<td>(Levitus et al. 2005; Levran et al. 2005)</td>
</tr>
<tr>
<td>L</td>
<td>Gene</td>
<td>Chromosome</td>
<td>Gene ID</td>
<td>Protein Domains</td>
<td>Function</td>
<td>Localization</td>
<td>References</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----------------</td>
<td>------------</td>
<td>---------</td>
<td>-----------------</td>
<td>----------</td>
<td>--------------</td>
<td>------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>FANCM</td>
<td>14q21.3</td>
<td>2048</td>
<td>63</td>
<td>Helicase</td>
<td>N</td>
<td>(Meetei et al. 2005; Singh et al. 2009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>FANCN/PALB2</td>
<td>16p12.1</td>
<td>1186</td>
<td>58</td>
<td>WD repeats</td>
<td>N</td>
<td>(Xia et al. 2006; Xia et al. 2007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>FANCO/RAD51C</td>
<td>17q25.1</td>
<td>376</td>
<td>76</td>
<td>-</td>
<td>N/C</td>
<td>(Vaz et al. 2010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>FANCP/SLX4</td>
<td>16p13.3</td>
<td>1834</td>
<td>51</td>
<td>BTB (POZ) domain</td>
<td>N</td>
<td>(Kim et al. 2011; Stoepker et al. 2011; Wan et al. 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>FANCQ/XPF/ERCC4</td>
<td>16p13.12</td>
<td>916</td>
<td>86</td>
<td>Helicase, Catalytic, and HhH domains</td>
<td>N</td>
<td>(Tripsianes et al. 2005; Bogliolo et al. 2013)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Derived from http://www.genecards.org
2 Derived from http://www.ensembl.org
* Exception: data obtained from lab
(Domchek et al. 2012; D’Andrea 2013), XRCC2 (Shamseldin et al. 2012), and deletion of the RUNX1 gene (encoding a transcription factor (Click et al. 2011)) have been linked to FA or a FA-like syndrome.

The FA proteins function together in the FA/BRCA pathway. In the last decade, an important part of FA research has focused on the role of the FA/BRCA pathway in the repair of DNA damage. The canonical FA/BRCA pathway is therefore well established now and the proteins can be divided into 3 groups: a core complex, central players, and downstream players. The core complex is assembled in different stages. Three subcomplexes are formed (Medhurst et al. 2006), FANCA/FANCG (Waisfisz et al. 1999a; Garcia-Higuera et al. 2000), FANCB/FANCL (Meetei et al. 2004a), and FANCE/FANCC/FANCF (de Winter et al. 2000c; Pace et al. 2002; Taniguchi and D’Andrea 2002; Léveillé et al. 2004; Léveillé et al. 2006), that together with FANCM form the final core complex. The fully assembled core complex facilitates via FANCL’s E3 ubiquitin ligase activity the monoubiquitination of the central players FANCD2 and FANCI, which together form the ID-complex (Meetei et al. 2003a; Meetei et al. 2004b; Smogorzewska et al. 2007; Sims et al. 2007). The monoubiquitinated ID-complex promotes the recruitment of the downstream players, FANCD1/BRCA2, FANCN/PALB2, FANCJ/BRIP1, FANCO/RAD51C, FANCP/SLX4, FANCQ/XPF/ERCC4, and BRCA1 to repair the DNA.

Besides the sixteen known FA proteins, there are several other proteins associated with the FA core complex, but no mutations in the corresponding genes have thus far been found in FA patients. These proteins are called the Fanconi Anemia Associated Proteins (FAAPs): FAAP100, FAAP24, FAAP20, FAAP16/MHF1, and FAAP10/MHF2 (Ciccia et al. 2007; Ling et al. 2007; Singh et al. 2010; Ali et al. 2012). Further an FANCD2/FANCI associated nuclease, FAN1 (Kratz et al. 2010; MacKay et al. 2010; Smogorzewska et al. 2010; Liu et al. 2010), and a FANCD2/FANCI deubiquitinated enzyme complex consisting of USP1 and UAF1 are associated with the FA/BRCA pathway (Figure 1) (Nijman et al. 2005; Cohn et al. 2007).

Figure 1. Schematic representation of the FA/BRCA pathway.
For detailed information see main text.
1.4 FA/BRCA pathway and cancer

The increased risk for cancer in FA patients is associated with a defect in the FA/BRCA (DNA repair) pathway. A defect in the DNA repair pathway results in genomic instability which creates a cancer prone environment (Knoch et al. 2012; Hosoya and Miyagawa 2014). While genomic instability afflicts all of the cells of FA patients, some cancers in non-FA patients have a tumor-specific (non-somatic) defect in one of the components of the FA/BRCA pathway. This also means that such tumors are hypersensitive to DNA cross-linking agents (see § 1.2 Cellular characteristics) – information that could be of value during treatment. In contrast, resistance to therapies based on DNA cross-linking agents can be due to the hyperactivation of the FA/BRCA pathway (Stecklein and Jensen 2012; Dietlein et al. 2014). The dual role of the FA/BRCA pathway, either enhancing a cancer-prone environment by being permissive to genomic instability when defective or enhancing the survival of cancer cells upon exposure to DNA damage when hyperactivated, is used in studies to refine cancer therapies (Balacescu et al. 2014; Hosoya and Miyagawa 2014).

1.5 Diagnosis and treatment

Before the cellular properties were known, diagnosis of FA patients was based on the phenotypic descriptions of Guido Fanconi (Fanconi 1927), which are still relevant when a clinician suspects Fanconi anemia. The first standard test to confirm or reject the clinical suspicion of FA is based on the occurrence of excessive chromosomal breaks induced by diepoxybutane (DEB) or other DNA cross-linking agents (Oostra et al. 2012). To further classify the breakage-positive patients the candidate-FA genes are investigated in order to find the disease-causing mutation(s).

For molecular diagnosis in the pre massive parallel sequencing era, a stepwise (cumbersome, time-consuming, and expensive) approach was taken for the molecular classification of FA. First, the mode of inheritance is determined, since in one case, FANCB, FA is X-linked (Meetei et al. 2004a). When FANCB can be excluded as a candidate-gene, possible mutations are investigated FANCA, FANCC, and FANCG, which will in 80-85% of the DNA breakage-positive cases identify the disease gene. If none of these genes are mutated, FANCD2 monoubiquitination in patient-derived cell lines is determined by Western blot. A defect in FANCD2 can distinguish between mutations up- or downstream of the FANCD2 monoubiquitination step (see the previous section on the FA-BRCA pathway). Monoubiquitination of FANCD2 is only accomplished with an intact core complex. If the Western blot shows loss of FANCD2 monoubiquitination this points to a defect upstream of FANCD2. In case of intact FANCD2 monoubiquitination a defect in the downstream pathway should be investigated. If the defect is upstream, the FANCL,
FANCE, FANCF, FANCM, and FANCI genes can be scanned for mutations. If the FANCD2 Western blot points out a downstream effect, RAD51 foci formation can be examined. If RAD51 foci are not formed, a mutation screening in FANCD1/BRCA2, FANCN/PALB2, and FANCO/RAD51C will be carried out. With normal RAD51 foci formation, FANCl/BRIP1, FANCP/SLX4 and FANCO/XPF/ERCC4 are screened. In addition, BRCA1 and XRCC2 may be screened. If no mutations are found, the disease genes for Nijmegen breakage syndrome (NBS1) and Warsaw breakage syndrome (ESCO2) need to be sequenced before one may decide that a novel FA complementation group has been discovered (Ameziane et al. 2008; Gille et al. 2012). With the emerging implementation of Whole Exome Sequencing (WES) and Whole Genome Sequencing (WGS) the molecular diagnostics of FA will become more efficient, quicker, and less expensive (Knies et al. 2012).

Treatment of FA is firstly focused on curing the anemia. At present, FA treatment consists of administration of either androgen or hematopoietic growth factors to increase blood cell production. The only beneficial treatment, so far, remains hematopoietic stem cell transplantation which, when successful, produces new blood cells in FA patients. Nonetheless, after successful hematopoietic stem cell transplantation FA patients stay at risk for developing cancer and are in need of regular checkups (Dufour and Svaen 2008; Smith and Wagner 2012; Peffault de Latour et al. 2013).

1.6 Uncovering new FA genes

Traditional approaches aimed at identifying candidate FA genes, such as linkage analysis, have been successful in relatively large patient cohorts or in kindreds with consanguinity. With the continuous development of high-throughput sequencing methods, the speed of such analysis is constantly improving. Nonetheless, both old and new molecular techniques, used to identify FA genes, often give rise to a large list of candidate genes, and therefore gene prioritization is highly desirable to minimize costs and time spent performing laboratory work (Yu et al. 2013).

Candidate-gene prioritization for a particular disease is based on the assumption that the encoded proteins of the disease-causing genes often interact directly or indirectly in a so-called protein-protein-interaction (PPI) network or pathway (Yu et al. 2013). PPI network information criteria can also be obtained from other protein properties such as length, functional domains, nuclear localization signal (NLS), iso-electric point (pI), and ortholog protein conservation as demonstrated for the identification of FANCI (Dorsman et al. 2007). Furthermore, information can be obtained from databases that contain data concerning mRNA expression and involved transcription factors, information about cellular localization via Gene Ontology (GO), and even text-mining tools can give valuable information (Haitjema et al. 2013). Integrating such information
about genes that are known to be involved in FA can be used to create a prioritized candidate gene list from large datasets such as generated from next generation sequencing. Development of gene prioritization methods that are tailored for specific diseases will aid in the timely discovery of new disease genes.

2. Roles of FA proteins in other networks than DNA repair

FA research has resulted in the identification of FA genes, thereby unraveling the FA/BRCA pathway and subsequently the role of the pathway in cancer. However, how the dysfunction of the DNA repair mechanism causes other features of the FA phenotype is still unclear (Kaddar and Carreau 2012), e.g. bone marrow failure (Garaycoechea and Patel 2013), abnormalities in glucose metabolism (Elder et al. 2008), and microcephaly (Shimamura and Alter 2010). This suggests that FA proteins, besides their role in the DNA repair network, have additional functions in other networks which may explain some of the observed phenotypes. Several studies have suggested roles for FA proteins in oxidative metabolism, cell cycle progression, apoptosis, and transcriptional regulation (Kaddar and Carreau 2012). In addition, the descriptions of FA-like patients (BRCA1 (Domchek et al. 2012; D’Andrea 2013); XRCC2 (Shamseldin et al. 2012); RUNX1 (Click et al. 2011)) or cellular phenotypes (mutant EGFR (Pfaffle et al. 2013)) suggests a possible role for FA proteins in a centrosome network (BRCA1 (Kais et al. 2012), XRCC2 (Cappelli et al. 2011), RUNX proteins (Chuang et al. 2012), and EGFR (Mardin et al. 2013)). Furthermore, cells from Seckel syndrome (SCKS) patients, a closely related genetic syndrome to FA (Andreassen et al. 2004), have increased number of centrosomes (Alderton et al. 2004). Of note, SCKS together with primary autosomal recessive microcephaly’s (MCPH) represent a spectrum of disorders characterized by microcephaly of which the involved gene products have roles, among others, in the centrosome cycle. Investigating the role of FA protein in transcriptional regulation and in centrosome protein networks (MCPH-SCKS spectrum) may uncover other networks than DNA repair in which FA proteins play a role.

2.1 Transcriptional regulation

Transcriptional regulation is a process that regulates levels of gene expression. Gene expression is the process in which a gene is transcribed (mRNA) from the DNA, and the mRNA template is subsequently translated into a protein which functions in a particular biological process (Figure 2A). Transcription regulation is initiated by a signal that gives rise to a cascade to activate (or repress) transcription factors, cofactors, chromatin remodelers (modifying condensed DNA), and recruitment of the transcription machinery (Figure 2B). Multiple transcription factors bind to the DNA in the promoter region of the
target gene, and together with the other factors initiate transcription of the DNA into the mRNA. This is a tightly orchestrated and controlled process. Dysfunctions in this process can cause a spectrum of diseases, such as cancer and neurological disorders and other developmental deficiencies (Lee and Young 2013).

**Figure 2. Transcription, translation, and transcriptional regulation.**
Schematic representation of transcription, translation, and transcriptional regulation. (A) DNA is transcribed into mRNA which is translated into a protein. (B) Prior to transcription of a gene several transcription factors (TF), cofactors (CO), and the general transcription machinery (Gen. TF machine) have to recruit to the DNA, a process called transcription regulation.

Only a few studies have investigated the transcriptional regulation of FA genes (Meier and Schindler 2011). The transcription factors that have been reported to regulate the transcription of FA genes indicate that there is an important role for the E2F/Rb pathway in the transcriptional regulation of FA genes (Tategu et al. 2007; Hoskins et al. 2008; Hoskins et al. 2012). Strikingly, FA proteins themselves also play an active role in regulating gene expression (Tremblay et al. 2008; Tremblay et al. 2009; Kaddar and Carreau 2012) or interact with proteins involved in transcriptional regulation (Kaddar and Carreau 2012). Consequently, the dysfunction of FA proteins, that regulate expression of genes, may have vital effects on certain biological networks.

Interestingly, FA proteins interact with key transcription factors involved in the oxidative stress response such as KEAP1 (FANCN/PALB2; (Ma et al. 2012)) and FOXO3a (FANCD2; (Li et al. 2010)). Early and recent reports have provided evidence on FA-associated oxidative stress. For example, FA cells have an excess in oxygen toxicity and mitochondria dysfunction (review in (Pagano et al. 2012)). Insight into the role of FA proteins in transcriptional regulation and into how FA genes are regulated can lead to potential therapeutic strategies for FA patients in particular and cancer in general (Pagano et al. 2012).

**2.2 Centrosome protein network**
A centrosome protein network is often affected in MCPH and SCKS, which are part of a spectrum of disorders characterized by microcephaly and the absence of visceral
malformations. Microcephaly (Greek: micro = small and cephaly = head) occurs in FA and other DNA repair disorders, for example, Nijmegen breakage syndrome (NBS), Warsaw breakage syndrome (WBS), and Xeroderma pigmentosum (XP) (Verloes et al. 1993).

Currently, 12 MCPH loci (MCPH1-12) and 8 SCKS loci (SCKL1-8) encoding 17 identified genes and 1 unidentified gene (two genes correspond to both MCPH and SCKL loci; Table 2). The encoding proteins have roles in diverse molecular pathways that are involved in proper spindle orientation, microtubule dynamics, cell cycle regulation, DNA damage response signaling, transcriptional regulation and the centrosome cycle (Mahmood et al. 2011). It is important to note, that MCPH1 is downregulated in breast, prostate, and ovarian cancers and is suggested to function as a tumor suppressor (Venkatesh et al. 2013), whereas MCPH2/WDR62 is found overexpressed in ovarian cancers and is suggested to be associated with the tumorigenesis of ovarian cancer (Zhang et al. 2013). Studying genes/proteins with characteristics that overlap across diseases such as FA and MCPH-SCKS might provide novel insight into the functioning of gene/protein networks. Elucidating these interaction networks can ultimately lead to the design of therapeutic approaches targeting specific cancers and the design for FA-subtype specific treatments (Kee and D’Andrea 2012).

3. Aim and outline of this thesis
The aim of this thesis was to expand our knowledge of the FA/BRCA pathway beyond its apparent role in DNA repair through genomic and proteomic approaches. In order to identify additional genes/proteins that participate in or connect to the FA/BRCA pathway we report in Chapter 2 how one could rank candidate-genes/proteins in terms of their relevance for the pathway. Chapter 3 describes a novel, non-standard, FA subtype (“FANCR”) and we demonstrate how the ranking procedure, described in chapter 2, is implemented in FA gene discovery research. In Chapter 4 we determined how FA genes are regulated, especially in relation to the cell cycle. In Chapter 5 we present an inventory of the currently known transcription factors thought to be involved in the transcriptional regulation of the FA/BRCA pathway, which can be used as an important tool for data mining. In Chapter 6 we demonstrate that CEP170 (a protein identified as an endogenous binding partner of MCPH2/WDR62; an important candidate on the basis of ranking criteria in chapter 2) causes microcephaly in an animal model when inactivated via knockdown. In Chapter 7 we present a novel hypothesis about the possible interaction of the FA/BRCA proteins with the MCPH-SCKL proteins. This interaction might explain some of the developmental abnormalities typically observed in patients affected by FA.
<table>
<thead>
<tr>
<th>Subtype</th>
<th>Gene</th>
<th>Chromosomal location</th>
<th>Exons</th>
<th>Protein size (aa)</th>
<th>Percentage Identity (aa) Human-Mouse</th>
<th>Domains</th>
<th>Protein structure</th>
<th>Cellular Localization^{1}</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCPH1</td>
<td>MCPH1</td>
<td>8p23.1</td>
<td>46</td>
<td>835</td>
<td>57</td>
<td>BRCT domain</td>
<td>yes</td>
<td>CP; CyS; MOC; CS</td>
<td>(Tommerup et al. 1993; Jackson et al. 2002; Trimborn et al. 2004; Trimborn et al. 2005; Garshasbi et al. 2006; Farooq et al. 2010; Darvish et al. 2010; Leung et al. 2011; Ghani-Kakhki et al. 2012; Sajid Hussain et al. 2012)</td>
</tr>
<tr>
<td>MCPH2</td>
<td>WDR62</td>
<td>19q13.12</td>
<td>22</td>
<td>1518</td>
<td>77</td>
<td>WD40 repeat</td>
<td>-</td>
<td>N; C; CyS; SP</td>
<td>(Bilgüvar et al. 2010; Nicholas et al. 2010; Yu et al. 2010; Kousar et al. 2011; Murdock et al. 2011; Bhat et al. 2011; Bacino et al. 2012; Sajid Hussain et al. 2012; Memon et al. 2012)</td>
</tr>
<tr>
<td>MCPH3</td>
<td>CDK5RAP2</td>
<td>9q33.2</td>
<td>49</td>
<td>1893</td>
<td>70</td>
<td>-</td>
<td>-</td>
<td>CP; CyS; MOC; CS; GA</td>
<td>(Bond et al. 2005; Hassan et al. 2007; Pagnamenta et al. 2012; Issa et al. 2013; Tan et al. 2013)</td>
</tr>
<tr>
<td>MCPH4*</td>
<td>CASC5</td>
<td>15q14</td>
<td>32</td>
<td>2342</td>
<td>61</td>
<td>-</td>
<td>-</td>
<td>N; Chr; CeM; K</td>
<td>(Genin et al. 2012)</td>
</tr>
<tr>
<td>MCPH5</td>
<td>ASPM</td>
<td>1q31</td>
<td>45</td>
<td>3477</td>
<td>70</td>
<td>CH domain</td>
<td>-</td>
<td>CP; CyS; N</td>
<td>(Bond et al. 2002; Bond et al. 2003; Passemard et al. 2009)</td>
</tr>
<tr>
<td>MCPH6 SCKL4</td>
<td>CENPJ</td>
<td>13q12.12</td>
<td>34</td>
<td>1338</td>
<td>73</td>
<td>TCP10</td>
<td>-</td>
<td>CP; CyS; MOC; CS; CeN</td>
<td>(Bond et al. 2005; Gul et al. 2006; Al-Dosari et al. 2010; Darvish et al. 2010; Sajid Hussain et al. 2012)</td>
</tr>
<tr>
<td>MCPH7</td>
<td>STIL</td>
<td>1p32</td>
<td>25</td>
<td>1287</td>
<td>75</td>
<td>-</td>
<td>-</td>
<td>CP</td>
<td>(Kumar et al. 2009; Darvish et al. 2010; Papari et al. 2013)</td>
</tr>
<tr>
<td>MCPH8</td>
<td>CEP135</td>
<td>4q12</td>
<td>8</td>
<td>1140</td>
<td>83</td>
<td>-</td>
<td>-</td>
<td>CP; CyS; MOC; CeN</td>
<td>(Hussain et al. 2012)</td>
</tr>
</tbody>
</table>

Table 2. Overview MCPH-SCKS spectrum.
### General Introduction

1. Chr = chromosome; CeM = centromere; CeN = centriole; CP = cytoplasm; CS = centrosome; CyS = cytoskeleton; GA = golgi apparatus; K = kinetochore; MOC = microtubule organization center; N = nucleus; SP = spindle pole

*Previous annotated as MCPH4 since the locus was published (Jamieson et al. 1999) and others identified mutations in CEP152 that is located in the linkage region of MCPH4 (Guernsey et al. 2010). However, due to re-analyzes of the MCPH4 locus CASC5 was identified in the initial MCPH4 family (Genin et al. 2012), and therefore CEP152 is now annotated as MCPH9 since mutations in MCPH patients were reported (Guernsey et al. 2010).*

<table>
<thead>
<tr>
<th>Region</th>
<th>Gene</th>
<th>Chromosome</th>
<th>Number</th>
<th>Size</th>
<th>Description</th>
<th>Function</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCPH9</td>
<td>SCKL5</td>
<td>15q21.1</td>
<td>23</td>
<td>1710</td>
<td>-</td>
<td>CP; CyS; MOC; CS</td>
<td>(Guernsey et al. 2010; Kalay et al. 2011)</td>
</tr>
<tr>
<td>MCPH10</td>
<td>ZNF335</td>
<td>20q13.12</td>
<td>52</td>
<td>1342</td>
<td>C2H2 Zinc Finger</td>
<td>-</td>
<td>N</td>
</tr>
<tr>
<td>MCPH11</td>
<td>PHC1</td>
<td>12p13</td>
<td>17</td>
<td>1004</td>
<td>SAM domain</td>
<td>-</td>
<td>N</td>
</tr>
<tr>
<td>MCPH12</td>
<td>CDK6</td>
<td>17q21-22</td>
<td>8</td>
<td>326</td>
<td>Kinase yes CP; N</td>
<td>(Schulze-Gahmen and Kim 2002; Lu and Schulze-Gahmen 2006; Hussain et al. 2013)</td>
<td></td>
</tr>
<tr>
<td>SCKL1</td>
<td>ATR</td>
<td>3q23</td>
<td>28</td>
<td>2644</td>
<td>Kinase - N; Chr</td>
<td>(O’Driscoll et al. 2003; Ogi et al. 2012; Mokrani-Benhelli et al. 2013)</td>
<td></td>
</tr>
<tr>
<td>SCKL2</td>
<td>RBBP8</td>
<td>18q11.2</td>
<td>28</td>
<td>897</td>
<td>PXDLS motif yes</td>
<td>N</td>
<td>(Børglum et al. 2001; Hassan et al. 2008; Qvist et al. 2011; Stokes et al. 2013)</td>
</tr>
<tr>
<td>SCKL3</td>
<td>-</td>
<td>14q22.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(Kiliç et al. 2003)</td>
</tr>
<tr>
<td>SCKL6</td>
<td>CEP63</td>
<td>3q22.2</td>
<td>37</td>
<td>703</td>
<td>-</td>
<td>CP; CyS; MOC; CS</td>
<td>(Sir et al. 2011)</td>
</tr>
<tr>
<td>SCKL7</td>
<td>NIN</td>
<td>14q22.1</td>
<td>40</td>
<td>2090</td>
<td>EF-hand -</td>
<td>CP; CyS; MOC; CS</td>
<td>(Dauber et al. 2012)</td>
</tr>
<tr>
<td>SCKL8</td>
<td>ATRIP</td>
<td>3p21.31</td>
<td>42</td>
<td>791</td>
<td>EEXXXD-DL motif</td>
<td>-</td>
<td>N</td>
</tr>
</tbody>
</table>
References


Faconti anaemia/Breast cancer consortium (1996) Positional cloning of the Fanconi anae-
mia group A gene. Nat Genet 14:324–328. doi: 10.1038/ng1196-324

Fanconi G (1927) Familiäre infantile perniziö-


Hussain MS, Baig SM, Neumann S, et al (2012) A truncating mutation of CEP135 causes primary microcephaly and disturbed centroso-


Kumar A, Gimel’jani SC, Duvvari MR, Blanton SH (2009) Mutations in STIL, encoding a pericentriolar and centrosomal protein, cause prima-


tein Hef is defective in Fanconi anemia complementation group M. Nat Genet 37:958–963. doi: 10.1038/ng1626


General introduction


