RAS-MAPK syndromes - a Clinical and Molecular Investigation

ANNA-MAJA NYSTRÖM
Dissertation presented at Uppsala University to be publicly examined in Rudbecksalen, Rudbecklaboratoriet, Dag Hammarskjölds väg 20, Uppsala, Wednesday, May 20, 2009 at 09:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract

The RAS-MAPK syndromes are a group of clinically and genetically related disorders, characterized by cardiac defects, facial dysmorphism, cutaneous abnormalities and neurocognitive impairment. The pathogenesis is dysregulation of the RAS-MAPK pathway, and several genes within the pathway are involved.

The present thesis aimed at identifying genetic causes in three of the RAS-MAPK syndromes - Noonan syndrome (NS), cardio-facio-cutaneous syndrome (CFC) and Neurofibromatosis-Noonan syndrome (NFNS) - and at correlating genotype with phenotype.

A mutation analysis of six genes associated with the RAS-MAPK syndromes in NS and CFC patients revealed mutations in 10/31 patients. The results suggested more complex genetic overlap and genetic heterogeneity among these syndromes than previously believed. Subsequently, gene dosage imbalances of seven RAS-MAPK-syndrome-related genes were investigated in mutation-negative patients. A multiplex ligation-dependent probe amplification strategy was developed that excluded copy number changes of these genes as a common mechanism in NS.

Genetic causes of clinical variability in NS were investigated where an atypical and severe NS patient was described. In addition, multiple café-au-lait (CAL) spots affected the patient and four otherwise healthy family members. Molecular analysis of four candidate genes revealed a previously described de novo PTPN11 mutation and an inherited NFI variant in the patient. Neither of them explained the CAL spots trait, which consequently represented a distinct entity. The results suggested that the atypical and severe phenotype in the patient could be a consequence of an additive effect.

Finally, a family displaying NFNS was investigated clinically and molecularly revealing a novel mutation in the GAP-domain of NFI. Furthermore, the results suggested that other RAS-MAPK-syndrome-related genes are not involved in NFNS. A proposal of prioritizing the GAP-domain of NFI in NFNS was presented.

Conclusively, these studies contribute to further understanding of the RAS-MAPK syndromes and facilitate the diagnostic process and future prognosis prediction.

Keywords: RAS-MAPK syndromes, Noonan syndrome, Neurofibromatosis type I, cardio-facio-cutaneous syndrome, mutation, multiplex ligation-dependent probe amplification, gene

Anna-Maja Nyström, Department of Genetics and Pathology, Uppsala University, SE-75185 Uppsala, Sweden

© Anna-Maja Nyström 2009

ISSN 1651-6206
ISBN 978-91-554-7508-6
urn:nbn:se:uu:diva-100804 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-100804)
Till hela min underbara familj
Supervisors:  

Marie-Louise Bondeson, Assoc. Prof.  
Dept. of Genetics and Pathology  
Uppsala University  
Uppsala, Sweden  

Göran Annerén, Prof.  
Dept. of Genetics and Pathology  
Uppsala University  
Uppsala, Sweden  

Faculty opponent:  

Anna Wedell, Prof.  
Dept. of Molecular Medicine and Surgery  
Karolinska Institutet  
Stockholm, Sweden  

Review board:  

Göran Andersson, Prof.  
Dept. of Animal Breeding and Genetics  
University of Agricultural Sciences  
Uppsala, Sweden  

Jovanna Dahlgren, Assoc. Prof.  
Dept. of Paediatrics  
The Queen Silvia Children's Hospital  
Göteborg University  
Göteborg, Sweden  

Fredrik Öberg, Assoc. Prof.  
Dept. of Genetics and Pathology  
Uppsala University  
Uppsala, Sweden
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


*These authors contributed equally to the work

Reprints were made with permission from the publishers.
Paper I: BMJ Publishing Group Ltd
Paper III: Wiley-Blackwell Publishing
Other publications from the author


Contents

Introduction ................................................................................................... 13
  Human genetic variation .......................................................................... 13
    Structural variations ............................................................................. 14
    Repeat sequences ................................................................................. 15
    Single nucleotide polymorphism ......................................................... 15
    Pathogenic variants .............................................................................. 15
    Consequences of mutations .................................................................. 16
  Human genetic disease ............................................................................. 17
    Monogenic disorders ............................................................................ 17
  Methods to identify disease-causing genes .............................................. 18
    Genetic mapping .................................................................................. 19
    Physical mapping ................................................................................. 20
    Mutational screening methods ............................................................. 22
    Determining the significance of mutations .......................................... 23

Cell signalling and the RAS-MAPK pathway .............................................. 24
  The RAS-MAPK pathway ....................................................................... 24
    RAS-MAPK substrates ........................................................................ 26
    RAS-MAPK activation and the different cellular responses ............... 27
    Regulation of the RAS-MAPK pathway ............................................. 29
    RAS-MAPK and cancer ...................................................................... 31

RAS-MAPK syndromes ................................................................................ 35
  Clinical description ................................................................................ 35
    Noonan syndrome ................................................................................ 35
    Cardio-facio-cutaneous syndrome ..................................................... 37
    Costello syndrome ............................................................................. 38
    LEOPARD syndrome .......................................................................... 38
    Neurofibromatosis type I and its variants ............................................ 39
    Neurofibromatosis type 1-like syndrome ............................................ 42
  Genetic studies ....................................................................................... 42
    PTPN11 ............................................................................................ 44
    SOS1 ............................................................................................... 48
    RAF1 ............................................................................................... 49
    KRAS ............................................................................................. 51
    BRAF .................................................................
Present investigations.................................................................61
Background ....................................................................................61
Aims of the study .............................................................................62
Paper I and II ..................................................................................63
Subjects............................................................................................63
Results and discussion ..................................................................63
Paper III ..........................................................................................68
Subjects............................................................................................68
Results and discussion ..................................................................68
Paper IV ..........................................................................................70
Subjects............................................................................................70
Results and discussion ..................................................................71
Concluding remarks and future perspectives ..................................73
Populärvetenskaplig svensk sammanfattning ...................................78
Acknowledgements ..........................................................................80
References........................................................................................83
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>array-CGH</td>
<td>Array-Comparative Genomic Hybridisation</td>
</tr>
<tr>
<td>ALL</td>
<td>Acute lymphoblastic leukaemia</td>
</tr>
<tr>
<td>BRAF</td>
<td>V-raf murine sarcoma viral oncogene homolog B1</td>
</tr>
<tr>
<td>CAL</td>
<td>Café-au-lait</td>
</tr>
<tr>
<td>CDC25</td>
<td>CDC-homology domain</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>CFC</td>
<td>Cardio-facio-cutaneous syndrome</td>
</tr>
<tr>
<td>C-FOS</td>
<td>V-fos FBJ murine osteosarcoma viral oncogene</td>
</tr>
<tr>
<td>CNV</td>
<td>Copy number variant</td>
</tr>
<tr>
<td>COSMIC</td>
<td>Catalogue of Somatic Mutations in Cancer</td>
</tr>
<tr>
<td>CR</td>
<td>Conserved region</td>
</tr>
<tr>
<td>CRD</td>
<td>Cysteine-rich domain</td>
</tr>
<tr>
<td>CS</td>
<td>Costello syndrome</td>
</tr>
<tr>
<td>DECIPHER</td>
<td>Database of Chromosomal Imbalance and Phenotype in Humans using Ensemble resources</td>
</tr>
<tr>
<td>DGV</td>
<td>Database of Genomic Variants</td>
</tr>
<tr>
<td>DH</td>
<td>Dbl-homology domain</td>
</tr>
<tr>
<td>dHPLC</td>
<td>Denaturing High-Performance Liquid Chromatography</td>
</tr>
<tr>
<td>DUSP1</td>
<td>Dual specificity phosphatase 1</td>
</tr>
<tr>
<td>DUSP6</td>
<td>Dual specificity phosphatase 6</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal growth factor</td>
</tr>
<tr>
<td>EGFR</td>
<td>Epidermal growth factor receptor</td>
</tr>
<tr>
<td>ELK1</td>
<td>ELK1, member of ETS oncogene family</td>
</tr>
<tr>
<td>ERK</td>
<td>Extracellular signal regulated kinase</td>
</tr>
<tr>
<td>GAB</td>
<td>GRB2-associated binding protein</td>
</tr>
<tr>
<td>GAP</td>
<td>GTPase activating protein</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanine diphosphate</td>
</tr>
<tr>
<td>GEF</td>
<td>Guanine nucleotide exchange factor</td>
</tr>
<tr>
<td>GH</td>
<td>Growth hormone</td>
</tr>
<tr>
<td>GRB2</td>
<td>Growth factor receptor-bound protein 2</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine triphosphate</td>
</tr>
<tr>
<td>HD</td>
<td>Histone-like domain</td>
</tr>
<tr>
<td>HGMD</td>
<td>Human Gene Mutation Database</td>
</tr>
<tr>
<td>HRAS</td>
<td>V-Ha-ras Harvey rat sarcoma viral oncogene homolog</td>
</tr>
<tr>
<td>IEG</td>
<td>Immediate early gene</td>
</tr>
<tr>
<td>IGF-1</td>
<td>Insulin-like growth factor-1</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>IGFBP3</td>
<td>Insulin-like growth factor binding protein 3</td>
</tr>
<tr>
<td>IQGAP1</td>
<td>IQ motif containing GTPase activating protein 1</td>
</tr>
<tr>
<td>JMML</td>
<td>Juvenile myelomonocytic leukaemia</td>
</tr>
<tr>
<td>KRAS</td>
<td>V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog</td>
</tr>
<tr>
<td>KSR</td>
<td>Connector enhancer of kinase suppressor of RAS</td>
</tr>
<tr>
<td>LS</td>
<td>LEOPARD syndrome</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MAP2K</td>
<td>Mitogen-activated protein kinase kinase</td>
</tr>
<tr>
<td>MAP3K</td>
<td>Mitogen-activated protein kinase kinase kinase</td>
</tr>
<tr>
<td>Mb</td>
<td>Mega base</td>
</tr>
<tr>
<td>MEK</td>
<td>MEK1 and MEK2</td>
</tr>
<tr>
<td>MEK1</td>
<td>Mitogen-activated protein kinase kinase 1</td>
</tr>
<tr>
<td>MEK2</td>
<td>Mitogen-activated protein kinase kinase 2</td>
</tr>
<tr>
<td>MLPA</td>
<td>Multiplex ligation-dependent probe amplification</td>
</tr>
<tr>
<td>MP1</td>
<td>MAPK scaffold protein 1</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>NF1</td>
<td>Neurofibromatosis type I</td>
</tr>
<tr>
<td>NF1</td>
<td>Neurofibromin</td>
</tr>
<tr>
<td>NF1S</td>
<td>Neurofibromatosis type 1-like syndrome</td>
</tr>
<tr>
<td>NFNS</td>
<td>Neurofibromatosis-Noonan syndrome</td>
</tr>
<tr>
<td>N-RAS</td>
<td>Neuroblastoma RAS viral (v-ras) oncogene homolog</td>
</tr>
<tr>
<td>NS</td>
<td>Noonan syndrome</td>
</tr>
<tr>
<td>OMIM</td>
<td>Online Mendelian Inheritance in Man database</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PH</td>
<td>Pleckstrin-homology domain</td>
</tr>
<tr>
<td>PTP</td>
<td>Protein tyrosine phosphatase domain</td>
</tr>
<tr>
<td>PTPN11</td>
<td>Protein tyrosine phosphatase, non-receptor type 11 gene</td>
</tr>
<tr>
<td>RAF</td>
<td>RAF proteins, BRAF, RAF1, ARAF</td>
</tr>
<tr>
<td>RAF1 (CRAF)</td>
<td>V-raf-1 murine leukaemia viral oncogene homolog 1</td>
</tr>
<tr>
<td>RAS</td>
<td>Rat sarcoma viral oncogene (small GTP binding protein/GTPase)</td>
</tr>
<tr>
<td>RAS-MAPK</td>
<td>RAS-induced Mitogen-Activated Protein Kinase</td>
</tr>
<tr>
<td>RBD</td>
<td>RAS-binding domain</td>
</tr>
<tr>
<td>REM</td>
<td>RAS exchange domain</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction Fragment Length Polymorphism</td>
</tr>
<tr>
<td>SH2</td>
<td>Src-homology 2 domains</td>
</tr>
<tr>
<td>SHP-2</td>
<td>Protein tyrosine phosphatase, non-receptor type 11</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SOS</td>
<td>SOS1 and SOS2, guanine exchange factors</td>
</tr>
<tr>
<td>SOS1</td>
<td>Son of sevenless homolog 1</td>
</tr>
<tr>
<td>SPRED1</td>
<td>Sprouty-related, EVH1 domain containing 1</td>
</tr>
<tr>
<td>SPRY</td>
<td>Sprouty homolog, antagonist of FGF signalling</td>
</tr>
<tr>
<td>WS</td>
<td>Watson syndrome</td>
</tr>
</tbody>
</table>
Introduction

Human diseases are caused by several factors and one of them lies within the human genome itself. The degree and type of genetic involvement in human genetic disorders vary substantially in different disorders; for instance high blood pressure, cancer and diabetes are caused by a combination of environmental and genetic factors, whereas other disorders are only attributed to defects in the genome. A fundamental milestone in the research field of human genetic disorders was the publication of the first draft of the human genome sequence almost a decade ago [1, 2]. In 2004, a near complete sequence was presented, which revealed that the human genome is approximately 3 billion bases and contains 20,000-25,000 protein-coding genes [3]. The information gained from the human sequence has greatly facilitated the field of human genetic disorders and has paved the path to characterizing genetic variations that influence normal human diversity.

In the present thesis, a group of human genetic disorders will be discussed where variations in the human genome are responsible for the disorders. These disorders are sometimes denoted RAS-MAPK syndromes, as they are all caused by different defects within the intracellular signal transduction pathway RAS-MAPK. These disorders share several clinical features and this has complicated the diagnosis of these patients. However, the RAS-MAPK syndromes also differ with regard to disease severity, in that some disorders are associated with a predisposition to cancer, some with severe mental retardation and some have a higher risk for progressive cardiac complications, whereas in some of the other disorders the affection status can occasionally be almost normal. Therefore, it is of crucial importance to increase our understanding of the clinical and genetic aspects of these disorders in order to facilitate the diagnosis and clinical management of these patients.

Human genetic variation

The fact that the human genome contains variations has been known since the 60’s, but our knowledge of the type and extent of this variation is constantly being updated. It has been estimated that the human genome differs by 0.1% to at least 1% between two randomly selected individuals [4]. These differences in the genome are of various types and size, ranging from one
single nucleotide to very large variations (see below). Regardless of the size or type of the variant, one important aspect is when a variant is disease-causing and when it is normal (polymorphism) and thus only contributes to the diversity seen among individuals. In general, a polymorphism is traditionally defined as a genetic variant that is present in \( \geq 1\% \) of the population. However, as our knowledge about human genetic diversity and disease has increased, it has become clear that variants cannot be strictly categorized into disease causing or polymorphism; instead there is a dynamic range between the two extremes, with several reported polymorphisms being functional and predisposing to disease.

Structural variations

Structural variations are defined as a DNA sequence that is greater than 1 kb in size and several different categories exist. One type of structural variation, denoted copy number variant (CNV), includes duplications, deletions or insertions of DNA compared to a reference sequence, and they can contain genes and regulatory regions. A large amount of data concerning CNVs has been reported since their discovery in 2004, and this has been dependent on the development of a novel technology: array-Comparative Genomic Hybridisation (array-CGH). The CNVs have been collected in the Database of Genomic Variants (DGV; http://projects.tcag.ca/variation/), which currently contains \(~6000\) CNVs. By analysing 270 healthy individuals from four different populations (HapMap; http://www.hapmap.org/index.html.en), it has been estimated that large CNVs cover \(~12\%\) of the human genome [5]. However, a recent report on the same HapMap samples using a higher resolution array-CGH has revealed several important findings regarding CNVs. First, the sizes of the CNVs have been overestimated 5-15 times. Second, large CNVs affect <5% of the human genome instead of 12%. Third, the human genome differs less between two individuals than previously estimated [6]. The frequency of these CNVs in the population varies from more than 1% to presence in only one individual, and in the study by McCarroll et al., it was observed that as many as 80% of their CNVs had an allele frequency exceeding 5% [6].

The CNVs often co-localize with another type of structural variation, the segmental duplications, which constitute about 5% of the human sequence [3]. In fact 25-50% of the CNVs were identified close to a segmental duplication [7]. Segmental duplications are low copy repeats present in more than one location in the genome, with the different copies sharing a sequence identity of \( >90\% \) [7]. There are two types of segmental duplications, inter-chromosomal and intrachromosomal, of which the latter is often involved in structural chromosomal rearrangements associated with disease [7].

Another category of structural variation is the inversions, \( i.e. \) a DNA segment that is reversed compared to the rest of the chromosomes. Finally,
translocations and segmental uniparental disomy also belong to the category structural variation [7].

Repeat sequences

The human genome sequencing revealed that various repetitive sequences account for more than 50% of the genome (which also includes segmental duplications) [1]. There are different types of repetitive sequences, either tandemly repeated or interspersed. The tandem repeats are composed of a repeat unit, e.g. CA that varies in length from one individual to another, e.g. (CA) 6. Different classes exist, satellites, minisatellites, microsatellites and mononucleotide tracts, which range from a total repeat tract of \( \sim 10^2 \) (microsatellites) to \( >10^5 \). Microsatellites have been extensively and successfully used in mapping of disease-causing genes because they are easy to genotype and have a high degree of length polymorphism among individuals.

Single nucleotide polymorphism

The smallest type of variant is the Single Nucleotide Polymorphism (SNP), and it is also one of the most common types of variation. A SNP can be either a deletion, insertion or a substitution of one single base. The number of SNPs in the human genome collected in the NCBI’s SNP database is \( \sim 16 \) million (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi; build 129). On average there is one SNP every 100-300 base (http://www.ncbi.nlm.nih.gov/SNP/), although the density of SNPs in the genome across regions varies [8]. In 2002, the International HapMap project started (http://www.hapmap.org/index.html.en), and today 3.1 million SNPs have been genotyped in the 270 HapMap samples and haplotype maps have been created [8]. These haplotype maps together with the abundance of SNPs in the genome and the development of techniques to study SNPs, such as the array technology, have made SNPs suitable for mapping of disease-causing genes in the human population.

Pathogenic variants

Variants representing all the above-mentioned categories that cause human genetic disease exist, hence these are pathogenic variants or, more commonly denoted, mutations. The most frequent types of mutations are those affecting one or a few bases. This bias towards small disease-causing variants is partly a reflection of the rather late technical development that has taken place during the past years, which now enables characterization of structural variants and their association with human genetic disease.

Regarding structural variations, there are currently 58 different recognizable syndromes listed in the Database of Chromosomal Imbalance and Phe-
notype in Humans using Ensemble resources (DECIPHER; https://decipher.sanger.ac.uk/application/), which has also collected numerous pathogenic structural variants in isolated cases. Well-characterized structural pathogenic variants include e.g. the microduplication syndrome Charcot-Marie-Tooth type 1A (CMT1A) and the reciprocal microdeletion syndrome hereditary neuropathy with liability to pressure palsies (HNPP) that involves a 1.4Mb region on chromosome 17p including the PMP22 gene [9].

Pathogenic repeat sequences include repeats expanding over a certain length and to date about 20 disorders have been associated with this pathogenetic mechanism [10]. Fragile-X syndrome is one example where a trinucleotide repeat, CCG, upstream of the FMR1 gene is pathogenic when the repeat expands over a certain number.

Regarding simple pathogenic variants that affect 20 bases or less, to date there are ~78,000 such mutations described in the Human Gene Mutation Database release 2008.4 (HGMD; http://www.hgmd.cf.ac.uk/ac/index.php). These mutations are either located in coding regions or in non-coding regions affecting promoters, regulatory sequences or splice sites. Coding region mutations include nonsense, missense, frameshift and silent mutations, where silent mutations are generally regarded as polymorphisms, as they do not change the protein sequence, however occasionally these can cause disease through different mechanisms involving aberrant splicing.

Consequences of mutations

The effect of a pathogenic variant on the phenotype is generally categorized as either loss-of-function or gain-of-function mutations.

In loss-of-function mutations, the protein has a decreased function or no function at all. For some genes, a reduced gene dosage to 50% is still enough for a normal phenotype, but when one functioning allele is not sufficient, it is denoted haploinsufficiency. One specific form of loss-of-function is when the mutation has a dominant negative effect, i.e. when a protein product that is formed does not function properly and in addition interferes and inhibits the function of the wildtype protein. There are several different types of mutations that cause loss-of-function, where deletions, translocations or insertions are examples of large variants, and nonsense, frameshift and splicing mutations are commonly involved small-scale mutations. Missense mutations are either loss- or gain-of-function depending on the role of the substituted amino acid. Regarding nonsense mutations, the truncated protein is rarely expressed due to activation of nonsense-mediated mRNA decay.

One specific type of loss-of-function effect is aberrant imprinting, which can occur due to e.g. deletions, balanced translocations or uniparental disomy (the presence of two paternal or two maternal homologous chromosomes in a cell instead of one from each). Aberrant imprinting is attributed to allele-specific expression of some genes, thus only one of the two parental
alleles is expressed and the other allele is inactivated through methylation. Examples of imprinting disorders are Prader-Willi syndrome and Angelman syndrome. These are caused by loss of imprinting in an imprinted domain containing several involved genes at 15q11-13. Loss of paternally expressed genes causes Prader-Willi and loss of maternally expressed genes causes Angelman syndrome.

In gain-of-function mutations, the protein has an enhanced or has acquired a new function, e.g. expressed in the wrong tissue or developmental stage. A gain-of-function effect commonly observed in various cancer types is the chromosomal rearrangement that fuses different exons from different genes into a new functional gene, e.g. the BCR-ABL fusion gene, which is a constitutively active tyrosine kinase causing the majority of cases of chronic myeloid leukaemia. Another mechanism is duplications that increase the dosage of genes. This gain-of-function mechanism is also frequently observed in human cancers and has hitherto been described in about ¼ of the syndromes described in DECIPHER. Finally, missense mutations can result in gain-of-function of the protein, a mechanism further discussed later in the thesis, as it is one of the main pathogenic variants involved in the studied disorders.

Several examples of disorders exist where the loss-of-function phenotype can occur either as an effect of a nonsense mutation in a dosage-sensitive gene or as a deletion of the same gene, e.g. Sotos syndrome, HNPP, or one of the disorders studied here, Neurofibromatosis type I [9, 11, 12]. The same has been shown to be true for gain-of-function missense mutations in dosage-sensitive genes producing a similar phenotype as a duplication of the same gene, e.g. CMT1A [9].

### Human genetic disease

Traditionally, the field of human genetic disease is subdivided into several categories; chromosomal disorders, mendelian or monogenic disorders, multifactorial or complex disorders and mitochondrial disorders, depending on what type of genetic aberration underlies the disorder and the extent of genetic vs. environmental involvement in the disorder. In the present thesis, the studied disorders belong to the so-called monogenic group of disorders, and therefore a few aspects of this category of disorders are explained.

#### Monogenic disorders

Mendelian or monogenic diseases are rare disorders, approximately 6000 are described in the Online Mendelian Inheritance in Man database (OMIM; http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM), and in about 2500 of them the genetic aetiology is known. The monogenic disorders have differ-
ent inheritance patterns. These are autosomal dominant and autosomal recessive, in which the defective gene is located on one of the autosomes, chromosome pair 1-22. Then there are X-linked dominant and X-linked recessive, in which the causative factor is located on the X-chromosome (Y-linked disorders are very rare). The term dominant means that one allele and recessive that two alleles need to be affected in order to express the phenotype.

Despite its simplistic name (“mono”) indicating an uncomplicated situation in monogenic disorders, there are several important factors that must be considered when studying these disorders. One aspect regards heterogeneity, where locus heterogeneity is when mutations in several different loci produce the same phenotype, allelic heterogeneity is when there are different types of mutations in the same gene in patients with the same disorder and finally clinical heterogeneity is when defects in the same gene are associated with different disorders and these disorders are often denoted as allelic disorders. All three types of heterogeneity are observed in the group of disorders studied here.

Other complicating factors when investigating patients/families with especially dominant disorders are reduced penetrance, i.e. the probability that an individual having the mutation will express the disorder, and variable expression, i.e. different features/severity of features are manifested in different individuals with the same disorder. These two factors are connected to each other because a disorder with very high variable expressivity can occasionally mimic a reduced penetrance, a situation that has been discussed in the context of the RAS-MAPK disorders [13, 14]. Related to these factors are also the age of onset of symptoms and the change of symptoms with age, two factors that are also a problem in the RAS-MAPK syndromes [14, 15]. Finally, anticipation is when the severity of a disorder can increase in subsequent generations, a situation that has been seen with the repeat expansion disorders, e.g. Fragile-X syndrome.

It is usually not known what causes the variability in expression or reduced penetrance in a monogenic disorder, however, some general causes are modifier genes, allelic heterogeneity and to a lesser extent environmental factors. One recently suggested alternative hypothesis involves CNVs as modifying factors in particularly loss-of-function mutations, where a dominant inherited mutation could be masked and result in reduced penetrance or reduced severity in the offspring if the child simultaneously inherits a CNV gain of the same gene from the unaffected parent [4].

Methods to identify disease-causing genes

In general, there are two major strategies to identify the defect that causes a monogenic disorder: genetic mapping and physical mapping. In both strate-
gies, different techniques exist depending on what type of patient material is available. In the next sections, some of the techniques and the basic theory of importance in the present thesis will be explained.

Genetic mapping
This type of strategy can be used when neither the disease-causing gene nor its location is known and when large well-characterized families are available. Briefly, two loci on different chromosomes have a 50% possibility of being inherited together. With loci on the same chromosome, the possibility of being inherited together depends on the distance between the loci, where two closely positioned loci are more likely to be inherited together than are two loci that are far apart. Closely positioned loci are thus inherited as a block or a haplotype in a family. This type of segment-like inheritance is an effect of recombination events during meiosis and recombination events are more likely to occur between distantly located loci. The proportion of recombination, or the recombination fraction, occurring between two loci is hence a measurement of the genetic distance between them. When two loci display 1% recombination, the distance between them is defined as 1 centi-Morgan (1 cM), which corresponds roughly to 1 Mb as measured in physical distance.

In linkage analysis, the recombination fraction occurring between two loci is calculated, where unlinked loci show 50% (0.5) recombination. To test the significance of the observation, hence if linkage has occurred merely by chance, the ratio between the likelihood of loci being linked vs. likelihood of loci being unlinked is calculated. This ratio gives the LOD score (Logarithm of the Odds) and for statistical significance in monogenic disorders, a LOD score of >3 is required, which means that there is a 1000:1 odds in favour of linkage. For monogenic disorders, a parametric or model-based procedure is used, requiring mode of inheritance, penetrance and allele frequencies. Requirements for successful linkage analysis are large enough families, a clear phenotype so that affected and unaffected family members can be distinguished and densely mapped informative markers. Traditionally, microsatellite markers have been used, although recent technological development enables the use of SNPs as markers. A successful linkage analysis identifies a candidate region in the genome, where all affected individuals in a pedigree share the same haplotype, which is not present in unaffected family members. This candidate region is delimited by recombination on both sides of the disease locus, and searching for candidate genes within the region can be initiated.

In cases where a gene is known for a disease, the fact that closely located markers co-segregate in a haplotype when transmitted to the next generation is possible to use as an indirect method to investigate whether or not a disease in a family is related to that specific gene. When using this method, it is
of crucial importance to use markers in close proximity to the gene and preferably intragenic markers or markers on either side of the gene, so the likelihood of having a recombination between the marker and the disease gene is reduced as much as possible. Furthermore, the markers must be informative so that the inheritance pattern of the alleles can be determined.

Physical mapping

In physical mapping, the specific location of a disease gene is determined, and one successful strategy has been to investigate different chromosomal rearrangements, particularly translocations, which may be causing the disease directly. The traditional techniques involving chromosome banding, e.g. G-bandning, have a low resolution with a detection limit of >3 Mb. The improvements of techniques in recent years has resulted in the widespread use of array-CGH, which has greatly facilitated the physical mapping of structural disorders, with a detection limit of a few kb today [6, 7].

The array-CGH technique is based on the immobilization of probes onto a slide/chip; thereafter patient DNA and a reference DNA are differentially labelled with fluorescent dyes, mixed and hybridized onto the slide. The fluorescent intensities are measured and quantified for each spot, and the intensities correspond to how much DNA is present in the patient and the reference. A ratio between the hybridized patient DNA vs. reference DNA is calculated for each probe on the array slide. A ratio of above 1.5 is thus a duplication, whereas a ratio of 0.5 corresponds to a deletion in the patient [7]. One alternative procedure to the comparative genomic hybridization in detecting genomic gains or losses is the use of comparative analysis of hybridization signals, which is used in SNP arrays [5, 6]. Today, the recent development of the Affymetrix SNP 6.0, which contains 1.8 million markers, has proven valuable in detecting CNVs with a very high resolution [6].

Similarly to the genome-wide screening of structural rearrangements, the investigation of specific loci has also been improved during recent years. Previously, Fluorescent In Situ Hybridisation (FISH) and Southern blot techniques were widely used. Today, one commonly used technique to detect or confirm gains and losses as well as other types of mutations is multiplex ligation-dependent probe amplification (MLPA) (Fig. 1) [7, 16]. MLPA is suitable for simultaneous detection of losses or gains of several different targets, e.g. genes or exons. For each target, two probes are designed to bind adjacent to each other on the target sequence. Each probe consists of a target-specific hybridization sequence and a universal PCR primer recognition sequence (occasionally a stuffer sequence is added) and the total length of the two probes is unique to each target. An initial hybridization step of the probes to its target is followed by ligation of the probes and denaturation. The universal PCR primer sequences enable simultaneous amplification of all the ligated probes with a fluorescently labelled primer pair, and the
amount of the amplification product is directly related to the amount of the target. Thereafter, the amplification products are separated according to size using capillary electrophoresis. The MLPA experiment also includes probes targeting control loci to enable normalization of the probes of interest vs. the control probes. Furthermore, normalization of each patient vs. healthy control individuals included in the run enables the calculation of a ratio for each target. The ratio is thereafter used to determine the copy number of the targets. Expected ratio for a gain is 1.5 and for a loss 0.5 [16]. Commonly used thresholds for gains and losses are <0.7 or 0.75 for a deletion and >1.3 or 1.25 for a duplication [17, 18]. Initially, the long MLPA probe of a probe pair was designed and prepared using M13-derived vectors [16], however this process is now facilitated by the development of using synthetic probe sets [19].

![Diagram of MLPA technique](http://www.mrc-holland.com/WebForms/WebFormMain.aspx)
Mutational screening methods

Once a candidate gene has been identified, either by physical or genetic mapping, or a disease-causing gene is known, a number of different methods are available to screen the gene for the presence of mutations. The screening of genes is mainly focused on the coding region and intron-exon boundaries and usually starts with these regions. However, because mutations can also occur in promoter regions, regulatory regions, or further away than the exon/intron boundaries, including screening of these regions should be considered upon a negative screening result of the coding region. Here, some methods used to screen the known genes associated with the RAS-MAPK syndromes studied in the thesis will be explained briefly.

When a large patient cohort is screened for unknown variations in a gene, one suitable method is denaturing High-Performance Liquid Chromatography (dHPLC). This method involves the separation of PCR-fragments at specific temperatures. Because a fragment containing a mutation (heteroduplex) will have reduced thermostability due to the mismatch, it will start to denature at a lower temperature than a fragment that is a perfect match between the two strands (homoduplex). The disadvantage of using this method is that it does not identify what type of mutation is present in the analysed fragment; only the presence of a mutation is detected.

Sequencing is another method to screen a gene for mutations, and compared to dHPLC the type of mutation is directly identified. The basis for this method is that during a PCR amplification, fluorescently labelled nucleotide terminators are incorporated into the PCR-product. This will stop the synthesis, so that at the end, different sized fragments will exist in the pool of fragments. Each fragment ends with one of the four differentially labelled nucleotides. These fragments are then separated according to size, and the colour of the fluorophore is detected corresponding to the base in the original sequence. The target in the amplification can either be genomic DNA or cDNA, where the latter enable the detection of aberrant splicing.

When the type of mutation is known and the aim is to screen a cohort of patients, such as relatives in a family, or unrelated controls for this specific mutation, a method based on Restriction Fragment Length Polymorphism (RFLP) is suitable. The basis for this method is that restriction enzymes have different recognition motifs in a sequence. If a mutation creates or deletes a motif for a specific enzyme, cleavage of the mutation-containing fragment will yield a different cleavage pattern than in a fragment that does not contain the mutation. The different cleavage pattern is analysed using gel-electrophoresis.
Determining the significance of mutations

Different mutations require different methods to assess the functional importance of the mutation. Either the mutation is a large structural rearrangement or a small scale mutation, information in databases such as DGV, DECIPHER, NCBI’s SNP database or HGMD is useful in determining whether the mutation has been reported before and in that case as a common polymorphism or as a disease-associated variant. However, caution must be taken when interpreting DGV and DECIPHER, because the hitherto collected data are based on array-CGH that often overestimated the size of the CNVs, as has been recently observed [6].

If the mutation has not been reported previously as a common SNP or CNV, the next step is to carry out segregation analysis, if possible, to determine whether or not the mutation has occurred de novo and to investigate whether the variation segregates with the disease. For small-scale mutations, analysis of healthy control individuals follows to rule out the possibility of having identified a common polymorphism.

Depending on the type of the novel small-scale mutation, the significance of the mutation can be difficult or easy to predict. Loss-of-function mutations are easier, whereas gain-of-function mutations are difficult to predict, especially the outcome of a missense mutation. The type of amino acid replacement and the degree of conservation of the affected amino acid among different species are often used to evaluate the importance of preserving the amino acid. To facilitate the evaluation of a mutation, different mutation prediction programs have been developed such as PolyPhen (http://genetics.bwh.harvard.edu/pph/) and PMUT (http://mmb2.pcb.ub.es:8080/PMut/), although the ideal assessment of the significance of a mutation is the functional characterization of it in vitro or in vivo.
Cell signalling and the RAS-MAPK pathway

A strictly regulated gene expression regarding both where and when is essential to achieving a correct outcome during all cellular processes, such as development, differentiation, proliferation, apoptosis and stress response. Generally, gene expression is regulated on three different levels: transcriptional, post-transcriptional and through epigenetic mechanisms. One mechanism to regulate gene expression on a transcriptional level is through signal transduction pathways. This mechanism is initiated by extracellular ligands, such as growth factors, hormones, neurotransmitters or other types of environmental signals, which bind to cell-surface receptors or diffuse through the plasma membrane and target intracellular receptors. The binding of the ligand activates the receptors, which activates a signal transduction pathway resulting in activation or inhibition of transcription factors. In the case of intracellular receptors, these can act as transcription factors themselves upon ligand binding. The subsequent role of transcription factors is to bind to response elements on the DNA, thereby regulating gene expression, which finally results in a cellular response, e.g. cell differentiation.

One central group of signal transduction pathways are the mitogen-activated protein kinase (MAPK) cascades, which include four major groups of MAPKs in eukaryotic cells, the extracellular signal regulated kinase (ERK), c-Jun N-terminal kinase (JNK), p38 and extracellular signal regulated kinase-5 (ERK5) [20]. The focus here is on the ERK cascade (also denoted RAS-MAPK) due to its involvement in the RAS-MAPK syndromes. However, the MAPK cascades are not isolated from each other, instead complex interactions occur between them. Therefore, one cannot exclude involvement of other MAPKs in the pathogenesis of this group of syndromes.

The RAS-MAPK pathway

The RAS-MAPK pathway as well as the other MAPK cascades consists of several levels, generally three or more, which enable the signal to be amplified and tightly regulated. These levels are commonly denoted as MAP3K, MAP2K and MAPK (Fig. 2).

The RAS-MAPK cascade is activated by several different receptors, such as G-protein-coupled receptors, growth factor receptors, ion channels, etc.
Upon stimulation of receptors, e.g. epidermal growth factor (EGF) binding to the epidermal growth factor receptor (EGFR), the receptors dimerize and autophosphorylate on tyrosine residues. This autophosphorylation recruits and activates enzymes (e.g. the phosphatase SHP2), adaptors (e.g. GRB2) and docking proteins (e.g. GAB) [20, 21]. The adaptor protein GRB2 interacts with SOS, which translocates to the cell membrane. However, SOS is also able to bind directly to the plasma membrane [22]. In any case, the recruitment of SOS to the plasma membrane is important for the next step in the cascade. SOS proteins are members of the guanine nucleotide exchange factors family and two highly homologous SOS proteins exist, SOS1 and SOS2. [22]. SOS catalyzes the conversion of inactive guanosine diphosphate bound RAS (RAS-GDP) to active guanosine triphosphate bound RAS (RAS-GTP). This GDP-GTP exchange activity of SOS is regulated by an autoinhibition of SOS and the regulation of SOS activation is important for the extent, degree and output of the signalling [21, 22].

Figure 2. The RAS-MAPK signalling pathway

The RAS proteins are small GTP-binding proteins, which are also denoted GTPases, and the activation of RAS is central to more than 20 effector cascades, among them the RAS-MAPK cascade. These effector pathways are involved in different cellular processes, and generalized the RAS-MAPK is a
key pathway involved in proliferation and differentiation [23]. Four different RAS proteins exist, NRAS, HRAS and two isoforms of KRAS, KRASA and KRASB, where KRASB is the major transcript of KRAS and therefore often denoted only as KRAS. These RAS proteins are modified post-translationally in different ways, which is responsible for their different subcellular localization. KRAS is localized primarily at the plasma membrane, but can upon different cellular changes redirect to other endomembrane compartments such as the Golgi, endoplasmatic reticulum and mitochondria. NRAS and HRAS, however, constantly move back and forth between the Golgi and the plasma membrane [21].

The active form of RAS, RAS-GTP, interacts with the RAF family kinases, ARAF, BRAF and RAF1 (also denoted CRAF), which constitute the MAP3K level. RAF is located in the cytosol in an autoinhibited inactive conformation, which is stabilized by a protein called 14-3-3. When the cell is stimulated, RAF is recruited to the plasma membrane and binds to RAS-GTP. Several mechanisms disrupt the autoinhibition of RAF and stabilize the active form of RAF. These mechanisms involve membrane interactions, the RAS-binding and phosphorylation events. Active RAF subsequently phosphorylates and activates the next level in the cascade MAP2K, also denoted MEK [21]. Two MEK proteins exist, MEK1 and MEK2, which are dual specificity protein kinases. Active MEK phosphorylates two residues, a threonine and a tyrosine, on ERK (the MAPK level of the cascade) and two ERK proteins exist, ERK1 and ERK2. MEK thus has a central role regarding the specificity of the RAS-MAPK pathway, because its only known substrates are ERK1 and ERK2 [24]. The phosphorylation of the two residues on ERK causes a conformational change, which subsequently results in a substantial increase in catalytic activity. When ERK is activated, it either translocates into the nucleus or remains in the cytoplasm. Totally, in the nucleus and in the cytoplasm, the number of known ERK substrates is approximately 160, and activated ERK influences these substrates either by phosphorylations or by direct protein-protein interactions [25].

Although the above-described components are the most common ones involved in various cell stimulations, the MAP3K level and levels upstream of it contain additional other proteins, which are activated in certain conditions or at certain subcellular localizations [25].

RAS-MAPK substrates
The substrates of ERK vary under different conditions, different cell types and even different cellular compartments and include transcription factors, protein kinases and phosphatases, cytoskeletal and scaffold proteins, receptors and signalling molecules, apoptosis-related proteins and others [25].

The nuclear substrates include transcription factors such as ELK1 and C-FOS, which belong to a group of genes, denoted immediate early genes
(IEGs). These genes are not dependent on a preceding protein synthesis to be expressed, but instead use constitutively present transcription factors for their expression. Therefore, the IEG products represent the first changes regarding transcription that follow upon cell stimulation, and these are present within minutes after cell stimulation. Subsequently, the IEG products are involved in the transcriptional regulation of secondary response genes that are dependent on protein synthesis. Regarding ELK1, it is phosphorylated by ERK at multiple sites, which results in an enhanced DNA binding capacity, enables the recruitment of other coactivators and substantially increases the transcriptional activity of ELK1 on downstream target genes. ELK1 is involved in the rapidly induced expression of C-FOS. C-FOS has an important role as transcription factor, as it regulates the early transcriptional response of several genes after cell stimulation and is involved in determining the cellular response upon stimulation (see below) [25]. Other well-characterized IEGs that are substrates of ERK belong to the JUN and MYC families.

Additional nuclear substrates include e.g. the transcription factor AML1, which is important for the myeloid hematopoietic differentiation, the transcription factor MITF, involved in the development of melanocytes and the transcription factor PAX 6, which is involved in central nervous system development.

Among the protein kinases, there are both cytosolic and nuclear kinases that are targets of ERK (see also section – Regulation of RAS-MAPK pathway). The ribosomal S6 kinase family (RSK) is primarily present in the cytoplasm before stimulation. When RSK is phosphorylated and activated by ERK, they contribute to the distribution of the ERK signal to additional substrates both in the cytoplasm and in the nucleus [25].

Examples of the cytoskeletal proteins are Synapsin I, which upon phosphorylation is involved in modulating synaptic plasticity and in regulation of neurotransmitter release, Connexin-43, which is involved in GAP-junctions and upon phosphorylation is thought to inhibit the coupling between the junctions, and Statmin, which is involved in cell cycle progression and cell differentiation facilitated by its phosphorylation [25].

Note that several of the ERK substrates are not uniquely activated by ERK, but also by the other MAPK cascades [25, 26].

**RAS-MAPK activation and the different cellular responses**

The cellular response to different ligands varies, e.g. fibroblast growth factor treatment of PC12 (pheochromocytoma) cells leads to cell differentiation, whereas if the same cells are stimulated with EGF, proliferation follows. These different effects were correlated with the duration of the ERK signal, where a prolonged signal was associated with translocation of ERK to the nucleus and subsequently with differentiation, whereas a short-lived signal
did not lead to nuclear translocation. Based on these findings, a model was proposed in which the signal duration determines the cell fate and different cell types use transient and prolonged signalling in different ways. For instance, it has been observed that in fibroblasts, the sustained signal results in proliferation [27].

Complicating the model suggested by Marshall et al., was the subsequent observation that different ligands can induce similar IEG response regardless of whether the signalling was sustained or transient. To explain how similar induction of IEGs correlated with different cell fates, the IEG product C-FOS was further investigated [28]. This revealed that the cell fate is nevertheless dependent on signal duration, but it also involves the capacity of IEGs to function as signal sensors (Fig. 3). It was shown that C-FOS is sensitive to degradation, and without phosphorylation by ERK, C-FOS is a very unstable protein, however efficient phosphorylation and activation of C-FOS occurred only if and when the signal reached a certain strength and duration. The subsequent activation of downstream target genes of C-FOS is therefore dependent on signal strength and duration. In contrast, when there is transient ERK signalling, C-FOS is rapidly degraded. Like C-FOS, other IEG products (C-MYC, JUNB etc.), which are substrates of ERK, require different signal strength and duration thresholds to become activated. The threshold differences of IEGs with respect to the signal strength and duration required to become activated are one of the factors determining the cellular response [26, 29].
The importance of ERK signal location, timing, distribution, duration and intensity for different cellular processes, which subsequently form a multicellular organism, has been demonstrated during mouse embryogenesis. It was shown that the spatial and temporal ERK activation was distinct with clear boundaries in the developing mouse. Furthermore, some domains of the developing embryo showed a prolonged ERK activation, such as limb buds, branchial arches, midbrain-hindbrain etc., whereas the peripheral nervous system, developing blood vessels, primordium of the eye, ears and heart showed a transient activation [30].

Regulation of the RAS-MAPK pathway

Several different mechanisms exist to regulate the signal duration, location and strength, and thereby the cell fate. Here, a few examples are described representing different categories of regulatory mechanisms (Fig. 4) [20].

One mechanism to regulate the signal is through the different posttranslational modifications of RAS proteins that localize the RAS proteins to different subcellular compartments. The different localized RAS proteins have consequences for the signal duration, the strength and the targets [21]. Both negative and positive feedback loops are involved in regulating the RAS-MAPK signal, and among the proteins involved in this type of regulation are some of the ERK substrates, e.g. the phosphatases DUSP6 and
DUSP1. DUSP6, a cytosolic protein representing a positive feedback loop, inactivates ERK through dephosphorylations. However, when ERK is activated, it phosphorylates DUSP6 and thereby DUSP6 becomes degraded. The opposite has been shown with DUSP1, which becomes stabilized upon phosphorylations by ERK, hence it can thereafter dephosphorylate ERK and inactivate it. Other negative feedback loops, which limit the signal duration, are the phosphorylations of ERK on MEK, RAF and SOS. Through different subsequent mechanisms, these phosphorylated upstream signal components are thus prevented from continuing the signalling, e.g. phosphorylation of SOS inhibits the interaction with GRB2 and consequently SOS is no longer recruited to the plasma membrane, which is required for its interaction with RAS [20].

![Diagram](image)

*Figure 4.* Some examples of how the signal of the RAS-MAPK pathway is regulated.

Other mechanisms to regulate the signal are scaffolding and anchoring, and a large number of these regulators have been identified. A scaffold has several different functions; e.g. it can bring two or more proteins in close contact, which thereby facilitates their interaction, it can target key components to a certain cellular location and thereby function as an anchor, or it can affect the extent of the signal. One example is the KSR proteins that bind RAF, MEK, ERK and 14-3-3 and is involved in the signalling at the plasma membrane. When KSR is present in excess in the cell compared to its binding
targets, it inhibits activation of ERK, whereas at optimal levels of KSR it enhances ERK activation. In contrast to the plasma membrane localization of KSR-mediated ERK signalling, MP1 is another scaffold that targets MEK1 and ERK1 and localizes/anchors the signal to endosomes. Another scaffold that targets BRAF is IQGAP1, which facilitates the activation upon EGF stimulation. Also the phosphatases are capable of anchoring the signal, e.g. DUSP6 anchors ERK2 to the cytoplasm [20, 21].

A final mechanism to regulate the signal is through inhibitors that are not phosphatases and among these inhibitors are the SPRY and SPRED proteins that can inhibit the activation of either RAS or RAF [21, 31]. One other group of molecules involved in the negative regulation of the signalling are the GTPase activating proteins, e.g. Neurofibromin. The RAS proteins have an intrinsic GTPase activity that is low, and therefore the GTP-GDP hydrolysis is slow. GTPase activating proteins stimulate this release of GTP from RAS, which is increased \(10^5\) times compared to the intrinsic GTPase activity. This conversion of RAS-GTP to RAS-GDP switches off the signalling in the cell [12].

The involvement of the phosphatase SHP2 in regulating RAS-MAPK is of great importance as it has been shown that SHP2 is required for full activation of the pathway and in certain conditions required for prolonged signalling. It is not fully understood what the exact mechanism is, how it regulates the activation and what the targets of SHP2 are, and several models are proposed (Fig. 4). One mechanism is the dephosphorylation of binding sites on the receptors or the docking protein GAB1, which are involved in the binding of GTPase activating proteins. Thereby the recruitment of these GTPase activating proteins to the plasma membrane is prevented or delayed, and consequently the signalling is prolonged. Another proposed mechanism is the dephosphorylation of SPRY which then is not able to inhibit the activation of the RAS-MAPK pathway [21, 32].

**RAS-MAPK and cancer**

Given the importance of the RAS-MAPK pathway in cell proliferation, differentiation and apoptosis, it is not surprising that disturbed RAS-MAPK signalling is one important component of human cancers. The involvement of the pathway in cancer has been known for about 30 years, and today several proteins in the RAS-MAPK pathway are associated with a number of different human cancers (Table 1) [23].
Table 1. Mutation in RAS-MAPK genes in human cancer

<table>
<thead>
<tr>
<th>Tissue</th>
<th>KRAS</th>
<th>HRAS</th>
<th>NRAS</th>
<th>BRAF</th>
<th>NF1</th>
<th>PTPN1</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS</td>
<td>0%</td>
<td>0%</td>
<td>18%</td>
<td>54%</td>
<td>0%</td>
<td>10%</td>
</tr>
<tr>
<td>adrenal gland</td>
<td>0%</td>
<td>0%</td>
<td>5%</td>
<td>2%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>autonomic ganglia</td>
<td>3%</td>
<td>0%</td>
<td>8%</td>
<td>1%</td>
<td>4%</td>
<td>1%</td>
</tr>
<tr>
<td>biliary tract</td>
<td>32%</td>
<td>0%</td>
<td>1%</td>
<td>14%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>bone</td>
<td>1%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>breast</td>
<td>5%</td>
<td>1%</td>
<td>1%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>central nervous system</td>
<td>1%</td>
<td>0%</td>
<td>2%</td>
<td>4%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>cervix</td>
<td>8%</td>
<td>9%</td>
<td>1%</td>
<td>1%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>endometrium</td>
<td>14%</td>
<td>1%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>eye</td>
<td>4%</td>
<td>0%</td>
<td>1%</td>
<td>7%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>GI tract (site indeter-</td>
<td>19%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>minate)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>genital tract</td>
<td>8%</td>
<td>ND</td>
<td>ND</td>
<td>0%</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>H L tissue1</td>
<td>5%</td>
<td>0%</td>
<td>12%</td>
<td>1%</td>
<td>1%</td>
<td>7%</td>
</tr>
<tr>
<td>kidney</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>large intestine</td>
<td>32%</td>
<td>0%</td>
<td>3%</td>
<td>14%</td>
<td>19%</td>
<td>0%</td>
</tr>
<tr>
<td>liver</td>
<td>6%</td>
<td>0%</td>
<td>4%</td>
<td>3%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>lung</td>
<td>17%</td>
<td>0%</td>
<td>1%</td>
<td>2%</td>
<td>0%</td>
<td>1%</td>
</tr>
<tr>
<td>oesophagus</td>
<td>4%</td>
<td>1%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>ovary</td>
<td>14%</td>
<td>0%</td>
<td>4%</td>
<td>12%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>pancreas</td>
<td>60%</td>
<td>0%</td>
<td>2%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>penis</td>
<td>4%</td>
<td>7%</td>
<td>0%</td>
<td>0%</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>peritoneum</td>
<td>6%</td>
<td>0%</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>pituitary</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>1%</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>prostate</td>
<td>8%</td>
<td>6%</td>
<td>1%</td>
<td>8%</td>
<td>17%</td>
<td>0%</td>
</tr>
<tr>
<td>salivary gland</td>
<td>3%</td>
<td>17%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>skin</td>
<td>3%</td>
<td>5%</td>
<td>18%</td>
<td>41%</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td>small intestine</td>
<td>20%</td>
<td>0%</td>
<td>25%</td>
<td>3%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>soft tissue</td>
<td>10%</td>
<td>7%</td>
<td>4%</td>
<td>3%</td>
<td>21%</td>
<td>2%</td>
</tr>
<tr>
<td>stomach</td>
<td>6%</td>
<td>4%</td>
<td>2%</td>
<td>1%</td>
<td>5%</td>
<td>0%</td>
</tr>
<tr>
<td>testis</td>
<td>5%</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>ND</td>
</tr>
<tr>
<td>thymus</td>
<td>15%</td>
<td>0%</td>
<td>0%</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>thyroid</td>
<td>3%</td>
<td>4%</td>
<td>7%</td>
<td>38%</td>
<td>8%</td>
<td>0%</td>
</tr>
<tr>
<td>UADT2</td>
<td>4%</td>
<td>9%</td>
<td>3%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>urinary tract</td>
<td>4%</td>
<td>12%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>
Together, the three RAS genes harbour activating mutations in 30% of human cancers, although different cancer types are associated with the three genes. The vast majority of cancer-associated mutations occur at amino acid position 12, 13 and 61 of the RAS genes, and these mutations have revealed a defective intrinsic GTPase activity and a resistance towards GTPase activating proteins. Hence, the cancer-associated mutations are causing a gain-of-function [23].

BRAF is involved in about 7% of human malignancies, and one single mutation, p.V600E, is responsible for 90% of all BRAF mutations [33]. One of the most frequently affected tissues is the skin, and up to 70% of melanomas are assigned to BRAF mutations. Interestingly, mutations in BRAF are also present in 80% of the benign skin naevi. The effect of the mutations on BRAF kinase activity varies, some display a high enhanced activity and some an impaired kinase activity. A similar gain-of-function mechanism, however, has been demonstrated with the impaired mutants, which is attributed to the capability of BRAF to interact with and activate RAF1. Consequently, both BRAF and RAF1 can activate MEK and transmit the signal. In the impaired kinase mutants, the direct BRAF-to-MEK1 signal is not efficient, however MEK1 is still being activated via the activation by RAF1 (Fig. 5) [33, 34]. The involvement of RAF1 in cancer is, contrary to BRAF, rare.

Figure 5. A different mechanism underlies the stimulation of MEK and ERK in activated or impaired BRAF mutants (adapted from Wan et al. [34]). RAF1 is denoted C-RAF here.
SHP2, encoded by *PTPN11*, is rarely involved in human cancer. Hitherto, the haematopoietic and lymphoid system is most often affected by cancer-associated mutations in *PTPN11*, especially in children. In juvenile myelomonocytic leukaemia, JMML, 35% of patients harbour mutations in *PTPN11*. These mutations are clustered around amino acid residue 60-76 and 502-503, and a high phosphatase activity has been observed as a result of these mutations, *i.e.* these are gain-of-function alleles [35]. Regarding JMML, totally 85% of the samples show mutations in either *NF1*, *PTPN11*, *KRAS* or *NRAS*. Hence, the involvement of RAS-MAPK in leukaemia is substantial [23].

*NF1* has been regarded as a tumour suppressor gene, and in the dominant cancer predisposition syndrome Neurofibromatosis type I, which is one of the investigated RAS-MAPK syndromes, somatic loss of the wildtype allele has been found in tumours in these patients [36]. Although the majority of *NF1* mutations reported in the COSMIC database are derived from tumours in Neurofibromatosis type I patients, *NF1* mutations also contribute to sporadic tumours not associated with Neurofibromatosis type I, *e.g.* colorectal carcinomas.

Involvement of copy number changes of RAS-MAPK-related genes has also been shown in cancer, in addition to point mutations. Large gains including *BRAF*, *NRAS*, *HRAS* or *KRAS* as well as other genes have been identified in about half of all patients with gliomas [37]. Furthermore, in about half of all patients with astrocytomas, a smaller gain of ~1 Mb encompassing *BRAF* has been found [38].

The involvement of several RAS-MAPK-related genes in many human cancers causing a constitutive activation of the pathway has prompted the search for specific inhibitors as candidate targets for cancer therapy. Several inhibitors against RAF kinases have been evaluated, and one is approved for use in renal and hepatocellular carcinomas. Although the MEK proteins have no role by themselves in human cancers, they have been considered tempting candidate targets to inhibit the constitutive active RAS-MAPK signalling in cancers. The reason to target MEK is its unique specificity towards the ERK and hence their unique role in transmitting the increased RAS-MAPK signal. Therefore, the RAS-MAPK pathway is the only pathway that would be targeted by MEK inhibitors. Several MEK inhibitors have been developed and are under clinical trials [33].
Different congenital disorders often share clinical features. With increasing knowledge about the genetic aetiology of disorders, genotype-phenotype conclusions can be drawn and the phenotypic overlap between disorders can be explained in terms of their sharing the same type of pathogenetic mechanism. One such example is the group of disorders that includes Noonan syndrome (NS), cardio-facio-cutaneous syndrome (CFC), Costello syndrome (CS), LEOPARD syndrome (LS), Neurofibromatosis type I (NF1) and Neurofibromatosis type I-like syndrome (NFLS). These are collectively denoted either as neuro-cardio-facial-cutaneous syndromes (NCFC), after their shared main characteristic features, or RAS-MAPK syndromes after their common pathogenesis, dysregulation of the RAS-MAPK pathway [39].

In the present thesis, three of these disorders have been studied and will be explained in more detail, NS, CFC and the NF1 variant Neurofibromatosis-Noonan syndrome (NFNS), while CS, LS and NFLS will be explained in brief. Furthermore, a detailed description of the genes (PTPN11, SOS1, KRAS, RAF1, BRAF, MEK1, MEK2 and NF1) and the molecular defects associated with NS, CFC and NFNS are presented separately below, together with a brief overview of the genetic defects associated with CS, LS and NFLS.

Clinical description

Noonan syndrome

The name Noonan syndrome, previously denoted male Turner syndrome, is an eponym after Jacqueline Noonan, who was one of the first to give a more comprehensive description of this group of patients in 1963. NS (OMIM 163950) is a relatively common condition occurring in about 1 in 1000-2500 births [40, 41]. Approximately 30-75% of the cases are familial, which display an autosomal dominant inheritance pattern, and remaining cases are sporadic [40]. NS is clinically variable, even within families, and also a changing phenotype has been observed with increasing age, thus familial cases could resemble an autosomal recessive inheritance pattern [14]. NS affects both females and males, with a sex-ratio distortion, favouring males by a factor of 2:1 [42]. The diagnosis NS is based on clinical observations
and a few scoring systems, *e.g.* the scoring system by van der Burgt *et al.*, presented in 1994 [41], have been proposed to facilitate the diagnosis of these patients. However, in the majority of cases the diagnosis is based on observation of main manifestations.

The main characteristics of NS include congenital cardiac malformation, short stature and facial dysmorphism [40, 41]. The spectrum of cardiac defects is wide, where pulmonic stenosis is the most prevalent type with a frequency of 50%-65%. The second most common type is hypertrophic cardiomyopathy, described in 20% of the patients. Other types of cardiac defects are *e.g.* atrioventricular canal and atrial and ventricular septal defects [40, 41, 43, 44]. The weight and length at birth is often normal. Feeding difficulties in infancy are present in about 63-75%, and this is often resolved by itself within 18 months. Regarding stature, only about 30% reach a height within the normal range in adulthood [40, 41, 45]. The final height of males is 162.5-169.8 cm and of females 152.7-153.3 cm, although it has been shown that growth hormone (GH) treatment in NS patients improves the final height [41, 44, 46, 47]. The main facial features in the newborn period include broad forehead, hypertelorism, down-slanting palpebral fissures, low-set posteriorly rotated ears with thick helices, a deep philtrum, high arched palate, a low posterior hairline and excess nuchal skin. The face as well as the body stature changes with age, *e.g.* the shape of the face becomes more triangular, the neck appears longer and webbing of the neck occurs. The facial features become less prominent in adulthood [14].

A number of associated features have been reported such as skeletal malformations, *e.g.* pectus carinatum superiorly or excavatum inferiorly present in 70-95% of individuals and scoliosis in 15%, cryptorchidism (60-77%), coagulation abnormalities (55%), lymphatic dysplasias (20%) and ophthalmic abnormalities among others [40, 41]. Different types of skin manifestations occur such as pigmented naevi (25%), café-au-lait (CAL) spots (10%), lentigines (2%) and a severe follicular keratosis (ulerythema ophryogenes) in 14% of patients, which may cause absence of eyebrows [40, 41, 43]. Mild mental retardation is reported in 15-35% of patients. Two studies reported a mean full-scale intelligence quotient (IQ) of 85 in NS, compared with the mean full-scale IQ of 90-109 of control individuals in the two studies. The IQ range in NS was, however, very wide [48, 49]. A mild delay in motor milestones has been observed with a mean age for sitting of 10 months and a mean age of walking alone of 21 months [43]. A delay in language development occurs in 20% of patients, with a mean age of 31 months for twoword sentences, which may be an effect of *e.g.* mild hearing loss or articulation deficiency [40, 43].

Patients with NS have a predisposition to cancer, albeit with a low risk. The reported malignancies include *e.g.* JMML, acute lymphoblastic leukaemia (ALL), rhabdomyosarcoma and neuroblastoma [41, 42]. A benign tumour that is associated with NS is multiple giant-cell lesions or giant-cell
granulomas, which affect the maxilla and mandible and other bone and soft tissues [41, 44].

Cardio-facio-cutaneous syndrome
CFC (OMIM 115150) was first described as a syndrome in 1986, but until 2001 there was often a question of whether or not CFC was a variant of NS, as it shows substantial phenotypic overlap with NS as well as with CS [50, 51]. Molecular investigations since 2001, when the discovery of the first NS associated gene was reported, have begun to explain the complex relationship between NS and CFC (see the genetics section).

CFC is a rare autosomal dominant disorder affecting both men and women with equal frequencies. In contrast to NS, CFC occurs only sporadically with a prevalence that has been estimated to a few hundred cases worldwide (http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gene) [51]. Like the other RAS-MAPK-related disorders, CFC displays clinical variability, with the main features consisting of short stature, congenital heart defect, facial dysmorphism, ectodermal abnormalities and mental retardation. The neonatal period is characterized by severe feeding problems, which are prolonged in comparison to NS [52]. Short stature at adulthood is present in more than 70% [52-54]. Cardiac defect is present in 62-84%, the most observed types being pulmonic stenosis, hypertrophic cardiomyopathy and atrial septal defect, among others. Notably, there is a higher frequency of hypertrophic cardiomyopathy of approximately 40% in comparison with NS [51-55]. Several of the facial features in CFC are similar to NS, e.g. epicanthal folds, ptosis, hypertelorism, down-slanting palpebral fissures, and relative macrocephaly. However, differences in facial appearance between CFC and NS include a less triangular face shape in CFC and instead a broader and longer face that is coarser than in NS. In addition, CFC presents with bitemporal narrowing and a short nose with a broad base [51]. The ectodermal manifestations of CFC are cardinal features that are present in more or less every patient. The skin findings include keratosis pilaris, hyperkeratosis, ichthyosis, pigmented naevi, hemangioma and CAL spots, among others. The hair is affected in the majority of cases, which manifests as sparse, thin and curly hair and absence of or sparse eyebrows and eyelashes [51, 52, 54, 55]. Follow-up studies of CFC patients have shown that some of these ectodermal features may improve with age, while others may not [51]. Neurological problems are reported in up to 80-100% of patients, thus, far more common than in NS [51, 52, 54-56]. Milestones in motor development are considerably delayed, with a mean age for sitting independently at 14-34 months and walking without assistance at 36 months. Language development is likewise delayed, with a mean age of first words at 25-28 months [52, 56]. Other frequently occurring complications are hypotonia, seizures,
and different MRI findings such as hydrocephalus, ventriculomegaly, Chiari I malformation, etc. [52, 56].

In addition to the main characteristics of CFC, a number of associated features occur, e.g. ophthalmology findings (e.g. optic nerve hypoplasia), genitourinary abnormalities (e.g. cryptorchidism 38%-44%), skeletal defects and gastrointestinal problems (e.g. constipation, vomiting, and swallowing difficulties), among others [52, 57]. However, bleeding disorders are not frequently observed in comparison to NS [51]. Regarding malignancies, CFC is considered to have no or low increased risk of tumours. To date, two cases with ALL and one case with hepatoblastoma have been reported and molecularly characterized [53, 54, 58].

CFC has hitherto been a clinical diagnosis, and important distinguishing features have included the degree of mental retardation and the ectodermal manifestations. In 2002, a CFC index was reported that could be used in the diagnostic process until a CFC diagnosis could be molecularly validated [57]. Recent findings emphasized the frequently occurring prolonged feeding difficulties, central nervous findings, seizures and optic nerve hypoplasia, which can also be used as diagnostic hints [52].

Costello syndrome
CS (OMIM 218040) is a rare, autosomal dominant disorder with only sporadic occurrence [59]. The condition is clinically similar to NS and CFC and is distinguished by facial manifestations, mental retardation (100%), short stature, heart defects (arrhythmia 53% and hypertrophic cardiomyopathy 61%), and prenatal overgrowth, which is followed by postnatal feeding problems, musculoskeletal abnormalities and skin manifestations. The facial appearance overlaps with that of NS and CFC, but is generally more “coarse” with additional features such as a large mouth, thick lips and full cheeks. The skin manifestations observed in nearly all patients include e.g. loose skin of the hands and feet and deep palmar and plantar creases. Multiple papillomata, which are small wart-like lesions, are the second most common ectodermal symptom present in about half of patients. The hair is often curly and sparse, but eyebrows are more commonly thick [59]. Predisposition to tumours, including especially rhabdomyosarcoma, is reported in approximately 15-20% of patients [60].

LEOPARD syndrome
The name LEOPARD syndrome (OMIM 151100) is an acronym formed by its main characteristics: Lentigines, ECG abnormalities, Ocular hypertelorism/Obstructive cardiomyopathy, Pulmonic stenosis, Abnormalities of genitalia in males, Retardation of growth and Deafness. LS is a rare autosomal dominant disorder with both sporadic and familial occurrence. Main features
of LEOPARD syndrome are the skin manifestations, hypertrophic cardiomyopathy and deafness (present in 15-25% of patients), which occurs more rarely in the other RAS-MAPK syndromes. The skin manifestations are mainly of two types, multiple lentiginosis and CAL spots. The multiple lentiginosis, usually found on the face, neck and upper part of the trunk, generally appear at 4-5 years of age and increase in number (~1000) until puberty. The CAL spots are found in about half of patients and are present at birth or appear before the lentigines. Hypertrophic cardiomyopathy is found in as many as 80% of patients and sometimes causes death, while pulmonic stenosis is present in about 40% of patients. Mild learning difficulties are diagnosed in ~30% of the cases, while hypotonia, which can delay motor development, is common in infants [61].

LS display extensive clinical overlap with NS and also with the Neurofibromatosis-Noonan syndrome and the late development of the characteristic multiple lentiginosis complicates the diagnosis in infants.

Neurofibromatosis type I and its variants

NF1 (OMIM 162200) was recognized and described in 1882 by von Recklinghausen [62]. NF1 is one of the most common autosomal dominant disorders, with an incidence of approximately 1 in 2500-3000. About half of NF1 cases are the first to be affected in their family, hence such cases have a de novo mutation. The phenotype involves the skin, bone and the nervous system and can generally be divided into tumour and non-tumour manifestations. In 1987, the National Institutes of Health Consensus Development Conference proposed the diagnostic criteria for NF1, which were reviewed in 1997 (see list below) [63]. Despite the identification of the causative gene behind NF1, the NF1 diagnosis is still based mainly on clinical evidence where two or more of the following features are required:

- Six or more café-au-lait spots
  - ≥0.5 cm in children
  - ≥1.5 cm in adults
- Two or more cutaneous/subcutaneous neurofibromas or one or more plexiform neurofibromas
- Freckling in the axilla or groin
- Optic glioma
- Two or more Lisch nodules
- Bony dysplasia (*i.e.* dysplasia of the sphenoid bone, dysplasia or thinning of the long bone, pseudoarthrosis)
- First degree relative with NF1
The tumour manifestations, the neurofibromas, are one of the characteristic hallmarks of NF1. These are benign tumours subdivided in cutaneous neurofibromas, subcutaneous neurofibromas and plexiform neurofibromas. The frequency of cutaneous neurofibroma in NF1 patients is >99%, while plexiform neurofibroma is less common with a frequency of ~30% that are visible and 50% that are found on imaging. Lisch nodule is another type of benign tumour that affects the iris of NF1 patients with a very high frequency (90-95%). Other benign tumours frequently observed in NF1 are gliomas, which affect the central nervous system e.g. the optic pathway (frequency of 15%), the brainstem and the cerebellum [62].

About 5% of NF1 patients develop malignant tumours [63, 64]. There are several different malignant tumours associated with NF1 such as leukaemia, paechromocytoma and rhabdomyosarcoma, although glioma is the most frequently observed. In addition, NF1 patients have a lifetime risk of 8-13% of developing malignant peripheral nerve sheath tumours (MPNST), which originate mainly from plexiform peripheral nerve sheath tumours (MPNST), which show a poor outcome [62].

The non-tumour manifestations of NF1 include the presence of >6 CAL spots and axillary and/or groin freckling. These symptoms occur with a very high frequency in NF1, >99% and 85%, respectively. Several other non-tumour manifestations are associated with NF1 that are more or less frequently observed. Cognitive impairment is a very common complication in NF1, with a frequency of 30-60% [62]. However, the level of impairment is only slightly below normal and an IQ<70 is rare. The cognitive problems observed in NF1 include motor deficits, reading/writing problems, visuospatial problems, memory deficits and attention-deficit hyperactivity disorder [62, 65]. The relationship between the cognitive profile in NF1 patients and the magnetic resonance findings of abnormally bright areas of the brain and spinal cord, so-called “unidentified bright objects” (UBOs), in many NF1 patients is not conclusive [65]. Orthopaedic lesions have a prevalence of 30-70% in NF1 and include non-specific symptoms such as scoliosis and short stature and the highly NF1-specific symptoms: sphenoid wing dysplasia and thinning of long bones (usually tibia and/or fibula) [62, 66]. Other features of NF1 include epilepsy (frequency of 6-7%), congenital heart defect (2%) (mainly pulmonic stenosis), hypertension, gastrointestinal problems, and macrocephaly [62, 64, 66].

NF1 is clinically extremely variable, both inter- and intrafamiliarily [12, 13, 67]. Although, the penetrance of NF1 is 100%, the severity of the symptoms in an offspring is unpredictable [13]. In addition, the ages of onset for the various symptoms differ, which is important to consider when diagnosing a patient and when performing genotype-phenotype correlation. Usually, the order of the most common diagnostic features to appear is: CAL spots, axillary freckling, Lisch nodules and neurofibroma. By the age of eight
years, approximately 90% of NF1 patients meet the diagnostic criteria of having two or more manifestations [15].

The clinical variability within NF1 and the overlap between NF1 and NS are illustrated by the two variants Watson syndrome (WS) and NFNS. WS is characterized by pulmonic stenosis, borderline intelligence, multiple CAL spots and freckling. In a review of some of the first WS families described, the WS phenotype was expanded to also include relative macrocephaly, Lisch nodules and neurofibromas in one third of the family members [68]. Furthermore, the study demonstrated that the clinical difference between WS and NF1 was in the prevalence of symptoms, e.g. Lisch nodules present in 58% in WS and in nearly all NF1 patients [68].

The name Neurofibromatosis-Noonan syndrome (OMIM 601321) was suggested in 1985 for patients displaying both NF1 and NS symptoms and later it was reported that NS features occur in a few percent of NF1 patients [64]. Some symptoms, e.g. short stature, mild MR, skeletal defects, seizures and pulmonic stenosis, overlap between NF1 and NS, while others are specific to each syndrome (NF1 - neurofibromas and Lisch nodules; NS – facial dysmorphism). In 1985, four possible explanations for these patients were presented, as summarized by Carey et al. in 1998 [69, 70]:

- NFNS a coincidental event of two autosomal dominant disorders, NF1 and NS
- NFNS being a variant of NF1
- NFNS being a variant of NS
- NFNS being a separate distinct disorder

Patients reported as having NFNS generally have a low incidence of plexiform neurofibromas, skeletal anomalies and internal tumours, but have hypertelorism, ptosis, low-set ears and cardiac defect [71]. Previously, similar findings had been reported, but in addition NFNS patients were suggested to lack Lisch nodules [70]. The genetic aetiology behind NFNS has been debated ever since 1985; however, in recent years, most of the evidence supports the idea that NFNS is a variant of the clinically variable NF1.

Multiple CAL spots are also associated with other syndromes in addition to WS, NFNS and NFLS (described below) as well as with cancer [62]. One other syndrome is autosomal dominant multiple CAL spots (OMIM 114030), which is allelic to NF1 in some cases [72]. Patients described with this phenotype only have multiple CAL spots and no other symptoms. In addition, it is important to remember that one or two CAL spots are also present in 10% of the general population [62].
Neurofibromatosis type 1-like syndrome

NFLS is a rare autosomal dominant disorder first described in 2007 [73]. The patients presented with multiple CAL spots, axillary freckling and macrocephaly. In addition, some individuals displayed NS-like faces and an attention deficit/learning difficulties resembling NF1. Several affected adults were affected by subcutaneous lipomas. The diagnostic criteria for NF1 were fulfilled in several affected patients, however NF1-specific tumour manifestations such as neurofibromas, central nervous tumours or Lisch nodules were absent. In contrast to WS and NFNS, the NFLS syndrome was surprisingly shown to be a distinct disorder, hence linked to a different gene than NF1, WS and NFNS – despite the extensive clinical overlap with these disorders.

Genetic studies

Genetic knowledge about the RAS-MAPK syndromes dates back two decades, although the major genetic contribution hitherto has taken place very recently (Fig. 6).

Summarizing the genetic knowledge from a clinical point of view would be as follows: NS is caused by mutations in PTPN11, SOS1, KRAS and RAF1; CFC is associated with mutations in BRAF, MEK1, MEK2 and KRAS; CS is linked to defects in HRAS; LS is caused by mutations in PTPN11 and RAF1; NF1 and its variants WS and NFNS are associated with defects in NF1; and finally NFLS is linked to mutations in SPRED1.

![Figure 6. Timeline of genetic findings on the RAS-MAPK syndromes. The years are indicated below the timeline; the RAS-MAPK syndromes are indicated above followed by the gene/genes each syndrome has been associated with (see below for references).](image-url)
In a chronological order, the causative genetic factor behind NF1 was identified in 1990 and the gene was denoted NF1 [36]. Subsequently, NF1 was also shown to be involved in WS [36]. The first NS-associated locus, 12q24, was localized through linkage analysis of two large families [74, 75]. Concurrently, genetic heterogeneity was reported [74]. In 2001, the causative gene within the NS candidate region was identified, PTPN11 [76]. Shortly thereafter, the PTPN11 gene was also linked to LS [77, 78]. However, it was excluded from involvement in the pathogenesis of CS and CFC, despite the clinical overlap with NS [60, 79, 80]. Regarding CFC, a criticized suggestion of a candidate region on chromosome 12, based on the presence of a deletion in a few patients with a CFC phenotype, was the only clue for several years [81-83].

The progress in gene identification in recent years has been based on a candidate gene approach, which was possible due to the accumulating knowledge about NF1 and PTPN11, the link of these genes to the RAS-MAPK pathway and to JMML. In 2005, CS was shown to be caused by mutations in HRAS [84]. The same year, it was demonstrated that NFNS was caused by mutations in NF1 in the majority of cases [71]. Next, four different genes, KRAS, BRAF, MEKI and MEK2, were identified in the pathogenesis of CFC [58, 85]. In 2006 and 2007, three additional genes involved in the pathogenesis of NS were identified, KRAS, SOS1 and RAF1 [86-91]. Also in 2007, LS was associated with mutations in RAF1 [87], and NFLS was linked to the SPRED1 gene [73].

A summary of the genetic information from a genetic point of view is presented in Table 2, which illustrates the general opinion about the genetic defects in the RAS-MAPK syndromes [12, 36, 42, 71, 73, 92]. This will be further discussed in the context of the studies in the thesis.
Table 2. Summary of the genes involved in the RAS-MAPK syndromes and the frequency of involvement*

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome band</th>
<th>Associated disorder</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTPN11</td>
<td>12q24</td>
<td>NS</td>
<td>~50%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LS</td>
<td>~90%</td>
</tr>
<tr>
<td>SOS1</td>
<td>2p22.1</td>
<td>NS</td>
<td>~13%</td>
</tr>
<tr>
<td>RAF1</td>
<td>3p25</td>
<td>NS</td>
<td>~3-17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LS</td>
<td>~3%</td>
</tr>
<tr>
<td>KRAS</td>
<td>12p12.1</td>
<td>NS</td>
<td>&lt;2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CFC</td>
<td>&lt;2%</td>
</tr>
<tr>
<td>BRAF</td>
<td>7q34</td>
<td>CFC</td>
<td>~75%</td>
</tr>
<tr>
<td>MEK1 and MEK2</td>
<td>15p22.31, 19p13.3</td>
<td>CFC</td>
<td>~20%</td>
</tr>
<tr>
<td>HRAS</td>
<td>11p15.5</td>
<td>CS</td>
<td>~100%</td>
</tr>
<tr>
<td>NF1</td>
<td>17q11.2</td>
<td>NF1</td>
<td>~100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NFNS</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WS</td>
<td>?</td>
</tr>
<tr>
<td>SPRED1</td>
<td>15q14</td>
<td>NFLS</td>
<td>?</td>
</tr>
</tbody>
</table>

*See text for further details and references

PTPN11

The PTPN11 gene consists of 16 exons and encodes a ubiquitously expressed cytoplasmic protein tyrosine phosphatase denoted SHP2. The protein contains two tandemly arranged amino-terminal SH2-domains (N-SH2 and C-SH2), which are followed by the catalytic protein tyrosine phosphatase domain (PTP). The protein circulates between two different forms, one inactive where the N-SH2 domain is bound to the PTP domain thereby blocking the active site and one active where the N-SH2 domain has bound a phosphopeptide, which subsequently releases the binding to the PTP domain, rendering the SHP2 active (Fig. 7) [32].

Figure 7. Structure of the 15 protein coding exons of the PTPN11 gene is shown above. The SHP2 protein is illustrated below with the three functional domains (adapted from Tartaglia et al., 2001 [76]).
**PTPN11** is the major gene involved in NS, responsible for about 50% of the patients [42, 76]. To date, 61 different missense mutations, two small in-frame deletions and one indel have been reported in the HGMD 2008.4. The mutations are located in exon 1-4, 7, 8 and 11-14, however exon 3 and 8 harbour 72% of all mutations (Fig. 8). The most frequently involved amino acid residue is p.N308 in exon 8, which is affected in nearly 25% of patients, where the substitution p.N308D is present in 19% of all patients. Two other commonly observed mutations both reside in exon 3, p.Y63C, affected in 9% of patients, and p.Q79R affected in 6% of patients. The majority of the aberrations affect amino acids involved in the autoinhibitory binding between the N-SH2 and PTP-domain or in close proximity to this interaction interface [39].

**Figure 8.** Distribution of PTPN11 mutations in totally 424 NS patients as reviewed in Aoki et al., [39]. The different grey areas correspond to different mutations, where a few mutations are indicated that are specifically discussed in the text.

Studies to address the question of genotype-phenotype correlation in NS with respect to PTPN11 mutations have revealed a significant relationship between the presence of a PTPN11 mutation and pulmonic stenosis. In contrast, a significantly lower frequency of hypertrophic cardiomyopathy was observed in PTPN11-positive individuals [93, 94]. A significantly higher occurrence of mutations in PTPN11 was also observed among familial cases than among sporadic patients [93, 94]. Conflicting data regarding other NS features have been presented. The study by Zenker et al. [94] showed a significant association between a PTPN11 mutation and short stature, easy bruising and thorax deformities and a lower prevalence of speech de-
lay/learning disabilities in PTPN11-positive patients; this was, however, not significant. The study by Tartaglia et al. [93] reported no difference in PTPN11-positive patients vs. PTPN11-negative patients regarding short stature, pectus anomalies, cryptorchidism and developmental delay, although the degree of developmental delay in patients having p.N308D was suggested to be mild, as none of the 17 patients in the study required special education [93]. This was confirmed in six additional patients who had either p.N308D or p.N308S [95]. Regarding short stature, other studies have reported both a significant difference and no difference in PTPN11-positive patients vs. PTPN11-negative patients [45]. Studies investigating the GH status in NS patients have revealed a significantly higher GH secretion and lower levels of IGF-1 and IGFBP-3, two proteins that are normally stimulated upon GH secretion, in PTPN11-positive patients vs. PTPN11-negative patients. Hence, a mild growth hormone resistance has been suggested in these PTPN11-positive patients [45].

The p.T73I mutation has been suggested to predispose to JMML, as about 40% of NS patients having this mutation presented with this malignancy [39, 96]. Germline mutations in PTPN11 are often different than the somatic mutations involved in cancer. Hitherto, 15 germline mutations show the same substitution as a somatic mutation and another 13 germline mutations affect the same codon as a cancer-associated mutation, but with a different substitution (COSMIC) (Table 3) [39].

Table 3. Summary of germline PTPN11 mutations that overlap with somatic mutations involved in cancer

<table>
<thead>
<tr>
<th>NS-associated mutations</th>
<th>Cancer-associated mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>p.Y63C</td>
<td>p.Y63C</td>
</tr>
<tr>
<td>p.F71I,L</td>
<td>p.F71K,L</td>
</tr>
<tr>
<td>p.T73I</td>
<td>p.T73I</td>
</tr>
<tr>
<td>p.E139D</td>
<td>p.E139D</td>
</tr>
<tr>
<td>p.T507K</td>
<td>p.T507K</td>
</tr>
</tbody>
</table>

Data collected from HGMD 2008.4 and from COSMIC
A gain-of-function mechanism in NS was proposed as the clustering of the mutations and structural analysis of them predicted that the autoinhibition of SHP2 would be impaired [76]. Functional studies of several different mutants, including p.N308D, demonstrated an increased basal phosphatase activity, which varied among mutants and was even more enhanced in the presence of EGF. Furthermore, the mutants prolonged the signalling through RAS-MAPK compared to wildtype, which was dependent on a ligand (e.g. EGF) and the docking protein GAB1. This prolonged RAS-MAPK signalling resulted in increased cell proliferation [96, 97]. In comparison with NS-associated SHP2 mutants, leukaemia-associated SHP2 mutants have demonstrated a higher phosphatase activity than mutants associated with NS [39]. In addition to the SHP2 mutants displaying changes in phosphatase activities, some mutants have an altered phosphotyrosine-binding capacity or display differences in substrate specificity instead of having enhanced phosphatase activity. Hence, the pathogenesis of NS is not only explained by increased phosphatase activity [98].

Different animal models of Shp2 mutants have been generated. While the homozygote knock-out mouse is lethal during embryogenesis due to defective gastrulation, a heterozygote knock-in of p.D61G demonstrated a short stature, craniofacial abnormalities, heart defects and myeloproliferative disease, thus all major NS features were present [39, 99]. The role of SHP2 in NS-associated heart defects has been further investigated using an inducible knock-in mouse model, which demonstrated that all cardiac defects associated with NS derive from the endocardium (the innermost layer of tissue in the heart and made of endothelial cells). However, the NS facial features were attributed to aberrant signalling in the neural crest cells [100]. Several different knock-in mice mutants were also created demonstrating that the specific *Ptpn11* allele as well as the genetic background was important to phenotypic outcome. Regarding the specific *Ptpn11* allele, the severity of the phenotype was correlated with the degree of Shp2’s phosphatase activity, *i.e.* the defects observed in mice were less severe with a heterozygote knock-in of p.N308D, which displays less elevated Shp2 activity than the other mutants tested. Finally, the study by Araki *et al.*, showed that the cardiac defects in mice were attributed to increased ERK activation [100].

Mutations in *PTPN11* are also responsible for the vast majority of patients with LS [77, 78]. The distribution of mutations, however, differs in comparison with NS-associated *PTPN11* mutations. In LS, only nine different missense mutations have been reported in HGMD 2008.4 affecting exon 7, 12 and 13, where p.Y279C and p.T468M account for ~50% and ~37% of the patients, respectively [39]. Although other exons are hot-spots for mutations in LS in comparison to NS, several of the LS-associated mutations have also been identified in NS patients, *e.g.* p.Y279C and p.T468M (Fig. 8).

Interestingly, functional characterization revealed that LS-associated mutations (*e.g.* p.Y279C and p.T468M) have reduced phosphatase activity, and
a dominant negative effect has been suggested that interferes with the RAS-MAPK pathway (reviewed in [101]). However, a recent transgenic fruit fly model investigating these two mutants, p.Y279C and p.T468M, revealed that despite the reduced phosphatase activity; the fruit fly demonstrated a gain-of-function effect. Furthermore, these mutants increased ERK activity, which was attributed to the residual phosphatase activity that is still present [101].

**SOS1**

*SOS1*, at locus 2p22.1, is a 23-exon large gene encoding SOS1, a ubiquitously expressed guanine exchange factor of RAS. The protein contains several important functional domains (Fig. 9), the histone-like domain (HD), the Dbl-homology domain (DH), the plekstrin-homology domain (PH), the RAS exchange domain (REM) and the CDC-homology domain (CDC25). The C-terminal is involved in the binding of SOS1 to GRB2, while the PH-domain is responsible for the alternative recruitment of SOS1 to the membrane. SOS1 has two RAS binding sites, one allosteric formed by the REM and CDC25 and one catalytic located in CDC25. SOS1 is autoinhibited by the DH-PH domain, which binds to the allosteric site and thereby suppresses the catalytic activity. The autoinhibition is abrogated when RAS binds to the allosteric site and induces a conformational change of SOS1 so that the catalytic site is released [22].

*SOS1* is the second most common gene involved in NS, responsible for about 13% of the cases [42, 89, 91]. Today, 23 missense mutations and one small in-frame indel have been reported (Fig. 9) [89, 91, 102, 103].

![Figure 9](image_url)

*Figure 9.* The structure of the SOS1 protein with the functional domains and the mutations indicated. Two of the 23 missense mutations result in the same amino acid substitution, p.G434R, hence only 22 mutations are depicted. In parentheses, the number of patients reported with the mutation.
The majority of mutations cluster within or close to three domains, DH, PH and REM, which are predicted to be involved in the maintenance of the autoinhibition of SOS1 [89, 91]. To date, two codons are more often affected, where different changes in p.R552 are most commonly involved in NS, accounting for about a third of the SOS1-positive NS patients, and the mutation p.E846K is responsible for about 13% of the SOS1-positive NS patients [89, 91, 103].

Genotype-phenotype correlations comparing SOS1-positive patients to PTPN11-positive patients revealed that SOS1-positive patients significantly more often had a thorax deformity (83% vs. 64%), but less often easy bruising (23% vs. 51%), mental retardation (18% vs. 43%), a short stature with a height of <3rd centile (36% vs. 71%) and cryptorchidism (48% vs. 80%) [103]. In addition, more ectodermal manifestations, e.g. keratosis pilaris, hyperkeratotic skin and curly hair, were present in SOS1-positive patients than in PTPN11-positive cohorts with great resemblance to the CFC phenotype [91, 103]. SOS1 is not regarded as a cancer-associated gene, and therefore it has been suggested that the risk of cancer for RAS-MAPK patients with a SOS1 mutation is most likely low (COSMIC) [42].

Based on the clustering of mutations and the structural modelling, an abrogation of the autoinhibition of the RAS-guanine exchange factor (RAS-GEF) activity of SOS1 was suggested. This was confirmed functionally through studies of the p.R552G mutation as well as some other representative mutations, which revealed an enhanced basal RAS and ERK activation and a prolonged RAS and ERK activation upon EGF-stimulation compared with wildtype [89, 91]. No NS-associated animal model has been generated; however, the homozygous Sos1 knock-out in mice was lethal during embryogenesis. In contrast, the highly homologous and ubiquitously expressed Sos2 has been shown not to be essential to mouse embryonic development [22].

**RAF1**

The ubiquitously expressed RAF1 gene (or CRAF) contains 17 exons and is located at 3p25. RAF1 shares three conserved regions (CR1, CR2 and CR3) with the two other known RAF proteins (BRAF and ARAF) (Fig. 10). The CR1 contains a RAS-binding domain (RBD) and a cysteine-rich domain (CRD), which are responsible for the recruitment of RAF to the membrane. CR2 constitutes a regulatory domain and binds to 14-3-3, an adaptor protein. CR3 contains the kinase domain and in the middle of CR3, the activation segment is located, which contains important regulatory phosphorylation sites. In addition to membrane recruitment, the N-terminal domain (CR1 and CR2) is involved in autoinhibition of RAF, as it interacts with the kinase domain. Furthermore, in BRAF, a glycine-rich loop (G loop) within the CR3 has been identified that is also involved in trapping BRAF in an autoinhibi-
ited conformation. It has been assumed that RAF1 has a similar mechanism. This autoinhibited conformation is stabilized by 14-3-3. However, an additional binding site for 14-3-3 is present in the C-terminal end, and this site has a positive effect on RAF because binding of 14-3-3 is essential to the activation of RAF [104].

RAF1 is involved in approximately 3-17% of NS [42, 87, 88]. In total, 15 different missense mutations have been identified and all of them cluster in three exons, 7, 14 and 17, which constitute parts of the CR2 and CR3 (Fig. 10) [87, 88, 102].

The mutations in exon 7, which are the most frequently occurring, affect any of the residues from 256-263. The single most common mutation is the c.770C>T, p.S257L, identified in about 40% of patients [87, 88, 102]. Mutations in RAF1 have also been identified in 2/6 patients affected by LS [87]. One of the two mutations was the most common one reported also in NS, p.S257L, and the other was p.L613V, which has been identified in NS patients as well [88].

The phenotype in RAF1-positive NS patients was characterized by a significantly higher frequency of hypertrophic cardiomyopathy than in NS in general (~95% vs. 20%) [87]. Specifically, mutations in exon 7 as well as 17 (p.S612T and p.L613V) were associated more with hypertrophic cardiomyopathy than were mutations in exon 14 [87, 88, 102]. Additionally, both patients with LS presented with hypertrophic cardiomyopathy. RAF1 is rarely involved in cancer, and hitherto only one of the somatic mutations involved in cancer, p.S259A (from ovarian tissue), affects the same codon as

Figure 10. The structure of the RAF1 protein with CR1, 2 and 3 indicated. Below the RBD, CRD, G loop and the activation segment are depicted, and the overlapping known somatic cancer-associated mutation is in italics. The germline mutations associated with RAS-MAPK syndromes are shown above. The number of affected NS patients having the mutation and the two LS-patients are specified in parentheses.
in NS, although the somatic substitution is different from the germline substitutions involved in NS (Fig. 10) [87].

As expected by the mutational clustering where some of the mutated residues are involved in the maintenance of the inactive conformation, all aberrations in exon 7 resulted in a gain-of-function with increased kinase activity and enhanced ERK activation, likewise the mutations in exon 17 [87, 88]. However, the mutations in exon 14 (residue p.D486 and p.T491) in CR3, which were not associated with hypertrophic cardiomyopathy, showed kinase impairment [87]. These results demonstrate the involvement of increased RAS activation in the pathogenesis of hypertrophic cardiomyopathy [87]. Through knock-down experiments in zebrafish of raf1 and of ptpn11, the distinct roles of raf1 and ptpn11 in heart development were elucidated, as the knock-down of raf1 resulted in an enlarged heart tube that could not be observed in the ptpn11 knock-down. This study correlates with the phenotypic differences seen in patients with respect to pulmonic stenosis and hypertrophic cardiomyopathy [88].

**KRAS**

The *KRAS* gene, located at 12p12.1, consists of six exons and encodes two splice variants, KRASA and KRASB, of which KRASB is the predominant isoform with ubiquitous expression. All RAS genes share a highly homologous G-domain, which has three important regions; the P loop that binds the GTP and the switch I and II regions that regulate the binding to regulatory components and downstream molecules. Furthermore, the RAS proteins have a hypervariable region at the C-termini, which is responsible for post-translational modifications and cellular localization (Fig. 11) [23].

*KRAS* is involved in less than 2% of patients with NS [42, 86, 90] and in less than 2% of patients with CFC [42, 58, 90]. Totally, 16 different missense mutations have been identified (Fig. 11) [55, 58, 86, 90, 102, 105-107].
Altogether, two mutations are more frequently observed than the others, p.V14I in ~24% of patients and p.D153V in ~18% of patients. The mutation p.V14I has, hitherto, only been identified in NS patients, while p.D153V has been associated with both NS and CFC. Among the patients, a few were described as having CS or as having features overlapping with CS. However, the diagnoses of these patients were revised and changed to NS or CFC, as a mutation in HRAS was thereafter considered to be required for the diagnosis CS [53, 60, 108].

The rare occurrence of KRAS mutations in NS and CFC impedes proper genotype-phenotype correlations. It has been observed though that clinical variability is extensive among KRAS-positive patients with features overlapping between NS, CFC and CS [55, 86, 107]. Interestingly, a higher prevalence of mental retardation in NS patients without PTPN11 mutation had previously been observed, and the same group noted that in their cohort of KRAS-mutation-positive patients mental retardation was present in all of them [94, 107]. The p.T58I has been associated with JMML in one case of totally three patients with the same mutation [55, 90]. Six of the hitherto identified germline mutations in KRAS have previously been reported as somatic mutations involved in cancer, either as the same or as a different substitution (Fig. 11). Various substitutions of p.G12 accounts for ~87% of
all the somatic mutations in KRAS and the p.G12S is one of the frequently occurring cancer-associated mutations (COSMIC).

Structural and functional analyses of several mutations, p.V14I, p.P34R, p.T58I, p.D153V p.V152G and p.F156L, have identified complex and variable gain-of-function effects (not for p.D153V) through different mechanisms: impaired intrinsic GTPase hydrolysis, reduced response to GAPs and an increased dissociation rate of GDP/GTP [86, 90, 109]. A mutation causing an increased GDP/GTP dissociation rate of KRAS would consequently result in preferred GTP binding, as the concentration of GTP in the cytoplasm is higher than that of GDP. There is therefore no need for a GEF, which normally enhances the release of GDP [86]. The effect of these germline mutations was also compared with the oncogenic p.G12D, which generally displayed a more potent gain-of-function effect [90, 109]. In contrast to the gain-of-function effect of the above investigated KRAS mutations, the p.G60R mutation did not activate ELK transcription factor, indicating an opposite effect on the RAS-MAPK pathway [58].

An animal model of an NS/CFC/CS-associated KRAS mutation has not yet been generated, but a homozygous knock-out of Kras in mice is lethal during embryogenesis and displays fetal liver defects and anemia. Due to the high involvement of KRAS in cancer, several knock-in models of the two most frequently occurring cancer mutations, p.G12V and p.G12D, have been generated, although not p.G12S that was found in one CFC patient [39].

**BRAF**

The BRAF gene, located at 7q34, consists of 18 exons and encodes BRAF, a cytoplasmic serine threonine kinase that is activated upon binding to RAS. BRAF is expressed in a variety of tissues with predominant expression in neuronal tissue. Like RAF1, BRAF is composed of three conserved regions (CR1, CR2 and CR3) with similar domains and functions (Fig. 12) [104].

BRAF accounts for the majority of mutations in CFC (~75%) [42], and to date, 31 missense mutations have been observed and two small in-frame deletions (Fig. 12) [53-55, 58, 85, 110-112]. A clustering of mutations has been observed in the CR1 domain in exon 6, which contains about half of all mutations, and another clustering is located in the CR3 domain in exon 11-16 [42]. Overall, the most often occurring mutation is p.Q257R, which has been found in about a third of all BRAF-associated CFC patients [53-55, 58, 85, 110-112].
Figure 12. The structure of BRAF with CR1, 2 and 3 indicated. Below the RBD, CRD, G loop and the activation segment are depicted. Also below the overlapping known somatic missense mutations involved in cancer are in italics with the different substitutions following the codon number. The germline mutations associated with RAS-MAPK syndromes are shown above. The number of affected CFC patients reported with the mutation is in parentheses. The 18 patients with CS or with a CS/CFC phenotype are indicated as “CS”. Two of the 31 missense mutations result in the same amino acid substitution, p.F595L, hence only 30 mutations are depicted.

BRAF has been regarded as a CFC-specific gene, although 18 patients with an initial clinical diagnosis of CS or a CS-like phenotype have also been associated with mutations in BRAF. Today, all these CS patients have been rediagnosed as having CFC, similarly as CS-like patients with a mutation in KRAS (Fig. 11) [53, 60, 108]. The initial genotype-phenotype correlation, where BRAF-positive patients displayed more ectodermal manifestations than KRAS-positive patients, was disregarded in a later study by the same group [54, 58]. Of the 31 identified missense mutations, only three amino acid substitutions have been reported in cancer, p.G464R, p.G469E and p.F595L, thus most of the mutations are unique to CFC (COSMIC). Hitherto, in the two CFC patients presenting with ALL, two different mutations in BRAF were identified: p.E501G and p.G469E.

Functional studies of different mutations have shown a variable effect. Several mutations, including the p.Q257R, demonstrated increased kinase activity and also increased levels of downstream molecules such as MEK,
ERK and the ELK transcription factor [58, 85]. These mutations showed variable BRAF kinase activity levels and some displayed kinase activities comparable to the most common cancer-associated BRAF-mutation, p.V600E. In contrast, a number of other mutations, e.g. p.E501K and p.E501G, were kinase-impaired and hence did not activate MEK, ERK or ELK [58, 85]. Animal models specifically investigating CFC-associated BRAF mutants have hitherto not been created. However, the homozygous knock-out of Braf in mice is embryonically lethal with vascular defects, and the knock-in in mice of the commonest cancer-associated mutation, p.V600E, is similarly embryonically lethal when ubiquitously expressed [39].

**MEK1 and MEK2**

The MEK1 gene, which is often referred to as MAP2K1, is found at 15q22.31 and the MEK2 gene, or MAP2K2, is located on 19p13.3. Each gene contains 11 exons, and they encode the mitogen-activated protein kinase 1 or 2, MEK1 or MEK2, which are threonine/tyrosine kinases downstream of RAF. The MEK proteins are composed of a large protein kinase domain with a central catalytic loop (Fig. 13). Furthermore, the N-terminal region is involved in regulating kinase activity and binding to the MEK substrates, ERK1 and ERK2 (reviewed in [113]).

Mutations in MEK1 or MEK2 account for approximately 20% of patients with CFC, although MEK1 defects are more common, with approximately a 2:1 ratio to MEK2 [114]. The majority of aberrations are missense mutations and all, except one mutation, are clustered in exon 2 and 3 in both genes (Fig. 13) [53-56, 85, 110, 112, 114]. Rarely, small in-frame deletions have been observed in both genes [53, 55, 114]. The MEK1 p.Y130C is the most frequently occurring substitution, identified in about half of the patients with MEK1 mutations [53-56, 85, 110, 112, 114]. In MEK2, different substitutions of the residue p.F57 are responsible for approximately one third of patients with MEK2 mutations [56, 85, 112, 114]. Mutations of the corresponding amino acids, i.e. p.F53 in MEK1 and p.Y134 in MEK2, have also been identified, emphasizing the importance of these two residues in the function of MEK [56, 85, 110, 112].

The MEK genes have been considered to be CFC specific, like BRAF. Nevertheless a few patients with NS and CS have been reported with MEK1 mutations. Like CS-like patients with mutations in BRAF and KRAS, the MEK1-positive CS patients have been rediagnosed as CFC [53, 60, 108]. In general, the phenotype of mutation MEK-positive CFC patients is in agreement with a classical phenotype when comparing with the CFC index and with the studies by Roberts et al. and Armour and Allanson [51, 52, 57, 114]. Furthermore, initial studies reported no difference in CFC phenotype between BRAF- and MEK-positive patients [54, 112]. Although CFC is not a
cancer-predisposition syndrome, one patient with the most common mutation, **MEK1** p.Y130C, suffered from a hepatoblastoma [53]. Neither of the two **MEK** genes are reported as having cancer-associated mutations (COSMIC), although constitutively activated MEK mutants have been generated that showed transformation capacity, and a **MEK1** mutation, p.D67N, was identified in a cervical cancer cell line [42].

Functional characterization of different mutations affecting the most common residues, p.F53 or p.F57 and p.Y130 or p.Y134, in **MEK1** and **MEK2**, respectively, revealed a higher capacity of stimulating ERK phosphorylation than the wildtype [85, 110]. Furthermore, it has been demonstrated that **MEK1** mutations p.F53S and p.Y130C and **MEK2** mutation p.F57C cannot activate ERK if they are not phosphorylated by RAF, although the **MEK2** p.F57C was less dependent on RAF phosphorylation [115]. Animal models

---

**Figure 13.** The structure of MEK1 (above) and MEK2 (below). Below each protein, the ERK docking sites, the catalytic loop and the two phosphorylation sites “P” for RAF are indicated. Above each protein are the identified mutations. The number of patients reported with the mutation is found in parentheses, where a few patients affected by NS or by CS and a mutation in MEK1 are also indicated by “NS” or by “CS”. Below MEK1 the somatic cancer-associated mutation p.D67N is indicated.
of Mek have revealed very different results, whereas the homozygous Mek1 knock-out in mice was lethal during embryonic development, the Mek2-/mice were viable with normal development [39]. Although no CFC-specific animal model has been created yet, inducible knock-in mice models of constitutively active Mek1 have shown some interesting phenotypes related to CFC-specific symptoms [39].

**HRAS**

The HRAS gene is located on 11p15.5, and like KRAS it contains six exons and has a similar structure of the protein [23].

Patients with a missense mutation in HRAS are considered to have CS [53, 60, 108], and the gene is involved in approximately 100% of patients. To date, it is not clear whether or not HRAS is the only gene involved in CS [60]. There are 10 missense mutations and two small indels reported in association with CS in the HGMD 2008.4, although as many as 90% of patients harbour a p.G12S substitution. In total, different changes of codon 12 and 13 are involved in the majority of aberrations [60].

Genotype-phenotype studies conducted thus far have failed to establish a clear correlation. However, an indication of an association of higher risk for malignancies in patients with p.G12A in comparison with p.G12S has been reported. Furthermore, patients affected by mutations residing at codons other than 12 and 13 seem to present a more atypical phenotype [60]. In a comparison between HRAS-positive and -negative patients, it was demonstrated that patients with HRAS mutations more often displayed arrhythmias such as atrial tachycardia, a higher frequency of polyhydramnios, ulnar deviation, GH deficiency and >1 papillomata than did HRAS-negative patients (or patients with mutations in BRAF and MEK) [53].

All germline mutations hitherto affecting codon 12 and 13 have been identified as somatic mutations involved in cancer (COSMIC).

Functional studies of different HRAS mutants have revealed similar pathogenetic mechanisms as demonstrated in KRAS mutants involving decreased intrinsic and GAP induced GTPase activity and furthermore an increased nucleotide dissociation rate [42]. Recently, a transgenic mouse expressing the CS-associated p.G12V mutant demonstrated several features in line with CS, e.g. facial dysmorphisms, hypertrophic cardiomyopathy and bone defects. However, other CS features were not present, e.g. growth reduction, childhood tumours, skin defects, among others. Interestingly, the activation of the RAS-MAPK pathway and also the PI3K pathway in both the heterozygous and the homozygous p.G12V mice were similar as in wild-type, despite the fact that the mutant was shown to be bound to GTP (active form of HRAS). The reason for this apparent lack of involvement of the RAS-MAPK pathway, however, was not explained in the study [116].
The gene for NF1, denoted \(NF1\), is located at 17q11.2 and is one of the larger genes with 61 exons spread over a region of more than 300 kb. It encodes several different transcripts and the most common one encodes a 2818 amino acid long protein, neurofibromin. Neurofibromin has a central catalytic domain, the GAP-related domain (GRD), corresponding to residues 1125-1537 (exon 20-27a), which is responsible for the accelerated intrinsic hydrolysis of RAS-GTP to RAS-GDP. The other functionally characterized domain in neurofibromin is SEC14, amino acid residues 1560-1698, which is found in secretory proteins and lipid-regulated proteins [12]. Two putative functional domains in neurofibromin have been reported, one is denoted cysteine/serine-rich domain (CSRD), which has been indicated by a slight clustering of missense mutations and in-frame deletions in the region comprising exon 11-17 [117, 118], and finally a pleckstrin homology (PH)-like domain (amino acids 1715-1816) that might be involved in regulation of the neighbouring SEC14 domain [119].

The mutational screening of \(NF1\) has been hampered by its large size, presence of several \(NF1\) pseudogenes and lack of clear mutational hot spots or recurrent mutations. Nevertheless, as many as 1033 different mutations in \(NF1\) have hitherto been reported in HGMD 2008.4. The mutations are subdivided in different categories depending on the type of aberration (Fig. 14); however, the description of the mutations at the DNA level does not always directly correlate with the effect on mRNA and on the protein.

![Figure 14. Neurofibromatosis type I associated mutations in NF1 as reported in HGMD 2008.4.](image-url)

A variable, but high number (≈20-50%) of the mutations have been reported to affect the mRNA splicing, which stresses the importance of also characterizing the mutation on the RNA level [117, 120, 121]. In addition to the majority of these mutations that are located in the consensus splice site, also
frameshift, nonsense and missense mutations were found to affect the splicing [122]. Several studies have reported a high proportion, 77-94% of mutations resulting in or predicted to give a truncated protein [117, 120, 121, 123]. In contrast, the in-frame defects (missense mutations and single amino acid changes) occur at a rather low percentage, 6-10% [117, 120, 121, 123]. As evident by the numbers in HGMD, approximately 5-10% of the patients have large deletions in the 17q11 region, which includes several neighbouring genes in addition to NF1. The most frequent type of deletion is 1.5Mb large [124, 125]. Conclusively, the pathogenetic mechanism of NF1 is haploinsufficiency of neurofibromin.

Despite the lack of obvious mutational hot spots and specific recurrent mutations, a few exons have repeatedly been shown to harbour more mutations than what would be expected if the mutations were distributed randomly, e.g. exon 4b and 37 among others [117, 120, 123, 126, 127]. In addition, some mutations have been identified in several patients, although each recurrent mutation has a low individual frequency, e.g. p.R304X [120, 123], which in the study by Ars et al. occurred in 4% of the total of identified mutations. Some of these recurrent mutations affect a CpG dinucleotide [117, 118, 123], which possibly explains the recurrence, while for instance the recurrent c.2970-2972 del AAT has been predicted to be caused by two almost perfect inverted repeats flanking the AAT that can form a hairpin-loop, which is stabilized upon the deletion of AAT [128].

Numerous studies of genotype-phenotype correlations have been performed throughout the years with poor results. Although the statistical significance was borderline, Castle et al. reported a lower risk of developing Lisch nodules in individuals with missense mutations in comparison with patients with a nonsense or frameshift mutation [129]. Until recently, the only clear correlation was between patients with a large deletion of the NF1 gene and a more severe phenotype, e.g. early onset and large number of neurofibromas, different dysmorphic features, growth abnormalities and more often cognitive impairment [130]. In 2007, a 3-base pair in-frame deletion in exon 17, c.2970-2972 del AAT, was demonstrated to be clearly correlated with lack of cutaneous neurofibromas [128].

The NF1 variant WS has been associated with different lesions in NF1, e.g. c.2970-2972 del AAT [36, 129]. Also, the variant NFNS is caused by defects in NF1 in the majority of cases [71, 131-133], although rarely the NFNS phenotype can also be an additive effect of two molecular defects [134-136]. In the family reported by Bertola et al., the NFNS phenotype was molecularly proved to be associated with two mutations, one in PTPN11 and the other in NF1 [136]. Hitherto, a number of different NF1 mutations have been associated with NFNS, and in some cases the same mutation has also been identified in classical NF1 patients (see Paper IV for a review).

Different mechanisms responsible for clinical variability both inter- and intra-familiarly among NF1 patients have been suggested [36, 67], and mo-
Molecular evidence for some of these mechanisms has been demonstrated, e.g. a two-hit event in benign and malignant tumours and in melanocytes from CAL spots [36, 137]. Modifying genes is another mechanism that has been suggested, and a functional polymorphism in the 5′UTR (+61 A/G) of the EGF gene was associated with an earlier onset of CAL spots and appearance of Lisch nodules in NF1 patients with a GG genotype [138].

Several different animal models of NF1 have been created [12, 139]. The homozygous knock-out of Nf1 in mice is lethal during embryogenesis and displays cardiac defects caused by increased cell proliferation and decreased apoptosis. In contrast, heterozygous mice are viable, display increased risk for tumour development and learning disabilities.

**SPRED1**

SPRED1 is located at 15q14 and contains seven exons. It encodes SPRED1, a negative regulator of the RAS-MAPK pathway.

After exclusion of mutations in NF1 in a few families displaying great resemblance to NF1, linkage analysis revealed a candidate region encompassing SPRED1. Mutational analysis of SPRED1 identified 12 different loss-of-function mutations, e.g. nonsense, frameshift, a splice mutation, a missense mutation and a small in-frame deletion. Functional analysis demonstrated that SPRED1 mutant proteins were not able to inhibit RAS-MAPK signalling [73]. Furthermore, a second somatic mutation was demonstrated in melanocytes derived from CAL spots, just as in NF1-associated CAL spots [73, 137].
Present investigations

Background

Clinically, the RAS-MAPK syndromes were described decades ago, while knowledge about the genetic aetiologies behind these disorders has emerged more recently. Hence, there are a number of unanswered questions regarding the RAS-MAPK syndromes, and further research is required that bridges the two sides, i.e., the clinical and genetic aspects of RAS-MAPK syndromes. The studies in the present thesis focused generally on identifying genetic factors in patients with RAS-MAPK syndromes and on investigating genetic factors involved in clinical variability. The background to the studies is presented below.

When these studies were initiated, NS had recently been associated with mutations in KRAS and CFC with mutations in KRAS, BRAF, MEK1 and MEK2. The findings on the involvement of these novel genes in causing NS and CFC prompted the investigations in Paper I. However the SOS1 gene involved in NS was described towards the end of the investigations in Paper I, and the association of RAF1 with NS was described after the studies in Paper I had been finished. Based on these findings, the general view during the past year or two has been that NS is caused by mutations in PTPN11, SOS1 and RAF1 and that CFC is associated with mutations in BRAF, MEK1 and MEK2. Finally, KRAS was associated with a variable phenotype [42, 92]. Hence, it was believed that NS and CFC were distinct disorders for all genes except KRAS [42].

Despite the recent discoveries of novel disease-causing genes in NS, the number of NS patients that lack a genetic defect in known RAS-MAPK-syndrome-related genes is not negligible, ~15-30% [42]. In 2008, a patient with a phenotype resembling NS was reported which was caused by a 10 Mb duplication including PTPN11 among other genes [140]. It was suggested that copy number gains of PTPN11 could be responsible for some of the remaining NS patients. The PTPN11 duplication, in addition to the NF1 deletion in some NF1 patients and the known gene dosage imbalances of some RAS-MAPK-related genes causing some cancers, motivated the study in Paper II.

All of the RAS-MAPK syndromes display clinical variability, and a number of associated features have been described to occur at different frequencies. For instance, some less common features of NS are urinary tract mal-
formations and CAL spots, both occurring in about 10% of cases. Furthermore, Arnold Chiari type I malformation has hitherto been described in only six NS patients, as reviewed in Ferrero et al. [40, 41, 43, 141]. However, the previously reported cases with Arnold Chiari type I malformation were not molecularly investigated or a mutation could not be identified (reviewed in [141]). The clinical variability in NS, the Arnold Chiari type I malformation that in NS patients has not been genetically explained together with the fact that NS is regarded as a relatively mild condition, motivated the investigations in Paper III.

The genetic aetiology behind patients displaying NFNS has been discussed ever since it was first described in 1985 when the four hypotheses were suggested. Two of the hypotheses have been confirmed, where the suggestion that NFNS is an example of extreme clinical variability in NF1 is true in the majority of patients [71]. The other hypothesis of coincidental occurrence of two common autosomal dominant disorders has been molecularly proven in one patient. However, both these hypotheses have only considered \textit{PTPN11} as a second candidate gene for NFNS, as it was the only described gene for NS until recently. The involvement of the recently described genes associated with the RAS-MAPK syndromes has not been assessed in the context of NFNS. This limited genetic information behind NFNS, along with the fact that the number of molecular investigated NFNS patients is still low, formed the basis for the investigation in Paper IV.

Aims of the study

Given this background to the present studies, the specific aims were:

1. To investigate genetic causes, in patients with NS or CFC and to correlate the genotypes with the phenotypes (Paper I and II).

2. To genetically investigate clinical variability within NS by analysis of a severe and atypical NS patient (Paper III).

3. To genetically investigate NFNS and the two hypotheses; extreme clinical variability in NF1 resulting in the NFNS phenotype \textit{vs.} co-occurrence of two disorders, NF1 and NS. This question was addressed by analysis of a large family presenting with both NF1 and NS (Paper IV).

62
Paper I and II

Subjects

The patients included in Paper I and II were referred to Uppsala University Hospital with a clinical diagnosis of NS or CFC or with phenotypic resemblance to these syndromes. Patients were clinically re-evaluated and of these patients, 31 (23 NS and 8 CFC) were selected that had previously been excluded to harbour PTPN11 mutations using dHPLC or sequencing. In Paper I, all 31 patients were analysed for mutations in BRAF, KRAS, MEK1, MEK2 and SOS1 (some exons previously shown to harbour mutations [89, 91]) using bi-directional sequencing of genomic DNA. Subsequently, an extended mutation analysis using the same method of remaining exons of SOS1 and of the complete coding region of RAF1 was performed, as these genes were described in the literature in the late phase of the investigations in Paper I (unpublished results).

In Paper II, 19 (18 NS and 1 CFC) of the mutation-negative patients (from Paper I and the subsequent extended analysis) were screened for gene dosage imbalances using a synthetic MLPA assay. The MLPA assay was designed to target PTPN11, KRAS, BRAF, RAF1, SOS1, MEK1 and MEK2 with two probes for each gene. The probes were validated using 15 healthy control individuals before being used in the screening of mutation-negative patients.

Results and discussion

The mutational screening of BRAF, KRAS, MEK1, MEK2, SOS1 and RAF1 (Paper I and unpublished results) identified causative de novo mutations in seven of the eight CFC patients and in three NS patients (Table 4). The screening of copy number variants of PTPN11, KRAS, BRAF, RAF1, SOS1, MEK1 and MEK2 in the pathogenesis of NS or CFC revealed normal gene dosage of the assessed genes.
Table 4. The mutational screening of KRAS, BRAF, RAF1, SOS1, MEK1 and MEK2 identified a mutation in totally ten patients.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Clinical diagnosis</th>
<th>Affected gene</th>
<th>Nucleotide substitution</th>
<th>Amino acid substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;i&lt;/sup&gt;</td>
<td>CFC</td>
<td>KRAS</td>
<td>c.458A&gt;T</td>
<td>p.D153V</td>
</tr>
<tr>
<td>2&lt;sup&gt;i&lt;/sup&gt;</td>
<td>CFC</td>
<td>BRAF</td>
<td>c.770A&gt;G</td>
<td>p.O257R</td>
</tr>
<tr>
<td>3&lt;sup&gt;i&lt;/sup&gt;</td>
<td>CFC</td>
<td>SOS1</td>
<td>c.1297G&gt;A</td>
<td>p.E433K</td>
</tr>
<tr>
<td>4&lt;sup&gt;i&lt;/sup&gt;</td>
<td>CFC</td>
<td>MEK1</td>
<td>c.607G&gt;C*</td>
<td>p.E203Q*</td>
</tr>
<tr>
<td>5&lt;sup&gt;i&lt;/sup&gt;</td>
<td>CFC</td>
<td>MEK2</td>
<td>c.171T&gt;A*</td>
<td>p.F57L*</td>
</tr>
<tr>
<td>6&lt;sup&gt;i&lt;/sup&gt;</td>
<td>CFC</td>
<td>MEK2</td>
<td>c.400T&gt;C</td>
<td>p.Y134H</td>
</tr>
<tr>
<td>7&lt;sup&gt;i&lt;/sup&gt;</td>
<td>CFC</td>
<td>BRAF</td>
<td>c.736G&gt;C</td>
<td>p.A246P</td>
</tr>
<tr>
<td>8&lt;sup&gt;i&lt;/sup&gt;</td>
<td>NS</td>
<td>BRAF</td>
<td>c.1495A&gt;G</td>
<td>p.K499E</td>
</tr>
<tr>
<td>9&lt;sup&gt;UP&lt;/sup&gt;</td>
<td>NS</td>
<td>RAF1</td>
<td>c.770C&gt;T</td>
<td>p.P261S</td>
</tr>
<tr>
<td>10&lt;sup&gt;UP&lt;/sup&gt;</td>
<td>NS</td>
<td>RAF1</td>
<td>c.781C&gt;T</td>
<td>p.S257L</td>
</tr>
</tbody>
</table>

<sup>i</sup>Mutations identified in Paper I
<sup>UP</sup>Mutations identified in the unpublished extended analysis
<sup>*</sup>Novel mutations

Of the missense mutations that were identified, two were novel MEK1 p.E203Q and MEK2 p.F57L. The MEK1 p.E203Q resides in exon 6 and is still the only mutation identified in MEK1 located outside the hot-spots exons 2 and 3. As expected based on the current knowledge about mutational clustering in the six analysed genes, nine mutations were found within the hot-spots regions and several of the identified mutations affected codons that are among the most commonly involved, e.g. RAF1 p.S257L identified in patient 10 has hitherto been identified in ~40% of RAF1-positive patients. Of the ten mutations, the number of transitions exceeds the number of transversions (6 vs. 4), which corresponds to what has been observed among point mutations, and one mutation affected a CpG dinucleotide, the RAF1 p.S257L. The presence of a CpG dinucleotide in RAF1 p.S257L explains the high frequency of this mutation among RAF1-positive patients.

In the mutational analysis of the cohort of NS and CFC patients, a genetic aberration could be identified in 7/8 CFC (87%) patients, which corresponds well to the reported frequencies of CFC patients in other studies (~60-90%) [54, 110, 112, 114]. The frequency of NS patients with a mutation in one of these genes in the studied cohort was 3/23 = 13%, which is slightly lower than what has been reported (<20-34%) [42]. However, other studies have also reported a similar low frequency in small cohorts like the present one [142]. In any case, the number of genetically unexplained RAS-MAPK-syndrome cases is considerable. Clearly, the search for the genetic aetiology of the RAS-MAPK syndromes is not at an end, and several hypotheses exist. One hypothesis was assessed in Paper II, namely the involvement of copy number changes of known RAS-MAPK-related genes. The results revealing...
no copy number changes of *PTPN11, KRAS, RAF1, SOS1, MEK1* and *MEK2* in 18 NS patients demonstrated that this is a rare mechanism in NS. Furthermore, a *BRAF* gene dosage imbalance in the pathogenesis of NS is excluded as a major mechanism based on the results in Paper II and the study by Sarkozy et al., in which 50 NS patients were investigated for *BRAF* copy number changes using MLPA [143]. Regarding CFC, only one patient was investigated in Paper II, which is a reflection of the rare occurrence of this disorder. In Sweden, the number of CFC patients is low, and an international collaboration is needed to exclude the hypothesis of gene dosage imbalances in CFC.

The results in Paper I have contributed to a recent report about genotype-phenotype correlations, in which all, hitherto, CFC-positive patients with *MEK* mutations (61 patients) were compared with 32 *BRAF*-positive CFC patients. In this study, a slightly lower frequency of cardiac defect in *MEK*-positive patients than in *BRAF*-positive patients was observed. Furthermore, ocular abnormalities and naevi were statistically less commonly observed in *MEK*-positive patients than in *BRAF*-patients, while downslanting palpebral fissures were statistically more often observed in the *MEK* patient group [114]. Another study in which the patients from Paper I are also included has similar conclusions as well as some additional observations (personal communications Dr. J. Allanson).

**Genetic overlap among RAS-MAPK syndromes**

Given that the general opinion has been that NS and CFC are distinct disorders for all genes except *KRAS* [42, 92], the genetic investigations performed in this cohort of NS and CFC patients contribute important findings to the field of RAS-MAPK syndromes with respect to two genes: *BRAF* and *SOS1*.

Regarding *BRAF*, one NS patient was shown to harbour a *BRAF* p.K499E mutation, which has previously been reported in CFC patients only [58, 85]. This patient was the second NS patient reported with a *BRAF* mutation. The study by Razzaque et al. identified the first NS patient harbouring a *BRAF* mutation, p.E501K [88]. The results in Paper I suggested that *BRAF* was also involved in the pathogenesis of NS and not only in CFC. Consequently, NS and CFC would be allelic disorders for *BRAF* in addition to *KRAS*. This result was criticized [92], however, subsequently three other reports have identified *BRAF* mutations in NS, confirming the association of NS and *BRAF* [46, 95, 143]. To date, totally 9 NS patients have been associated with *BRAF* and six different mutations have been identified (Table 5).
The prevalence of \textit{BRAF} mutations in NS was recently estimated to <2\% [143]. The phenotype of the patient presented in Paper I is in agreement with findings in the five NS patients reported by Sarkozy \textit{et al.} [143] with regard to more pronounced mental retardation, feeding problems, hypotonia and pigmented naevi in \textit{BRAF}-associated NS patients than in \textit{PTPN11}- or \textit{SOS1}-associated NS patients. However, the patient reported by Pierpont \textit{et al.} [95] displayed a mental development within the normal/mild range. Hence, further studies are required to draw an appropriate genotype-phenotype conclusion as well as to investigate whether allele specificity exists, which was suggested by Sarkozy \textit{et al.} [143] with regard to NS and CFC mutations.

Regarding \textit{SOS1}, Paper I presented the first CFC patient associated with a \textit{SOS1} mutation, p.E433K. Previously, mutations in \textit{SOS1} had only been associated with NS and an overlapping NS/CFC phenotype, where ectodermal manifestations were present similar to those in CFC [89, 91, 103]. However, the study by Zenker \textit{et al.} excluded \textit{SOS1} mutations in the CFC pathogenesis [103]. Therefore, the results in Paper I suggested a complex genetic overlap in the pathogenesis of NS and CFC with respect to mutations in \textit{SOS1}. Also the suggestion of \textit{SOS1} involvement in CFC was criticized [92], however, one subsequent study confirmed our findings of \textit{SOS1} involvement in CFC, where one CFC patient harboured a novel mutation p.P481_p.G482insRLP [142]. The study by Narumi \textit{et al.} furthermore suggested that \textit{SOS1} mutations are associated with a phenotype ranging from NS to CFC. In summary, regarding \textit{SOS1} there is contradicting data on its involvement or lack of involvement in CFC, and this issue needs to be further investigated. However, a general consensus is that \textit{SOS1}-positive patients display an extensive clinical overlap between NS and CFC, as was illustrated by the patient in Paper I.

The results in Paper I demonstrating a greater genetic overlap than previously suggested are not the only evidence. One additional example of genetic overlap that is more extensive than previously thought is the identification of \textit{MEK1} mutations in a few NS patients [55]. Conclusively, the genetic overlap between NS and CFC includes \textit{BRAF}, \textit{SOS1} and \textit{MEK1} in addition to the previously described \textit{KRAS}. However, the total degree of genetic overlap is still unknown, as a similar genetic overlap might occur among the

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{Nucleotide substitution} & \textbf{Amino acid substitution} & \textbf{Number of patients} & \textbf{Disease} & \textbf{Reference} \\
\hline
\textit{c.1501G>A} & \textit{p.E501K} & 2 & CFC, NS & [46, 58, 88] \\
\hline
\textit{c.1495A>G} & \textit{p.K499E} & 1 & CFC, NS & [58, 85, 144] \\
\hline
\textit{c.1789C>G} & \textit{p.L597V} & 2 & NS & [95, 143] \\
\hline
\textit{c.722C>T} & \textit{p.T241M} & 1 & NS & [143] \\
\hline
\textit{c.722C>G} & \textit{p.T241R} & 1 & NS & [143] \\
\hline
\textit{c.1593G>C} & \textit{p.W531C} & 2 & NS & [143] \\
\hline
\end{tabular}
\caption{Summary of the hitherto identified \textit{BRAF} mutations in NS patients}
\end{table}
genetic defects that remain to be discovered in the NS and CFC patients who lack mutations in the known genes.

Clinical overlap among RAS-MAPK syndromes and how to handle it

The extensive genetic overlap presented in Paper I, together with the study in which MEK1 was associated with NS [55] and the variable phenotypes associated with KRAS [86, 90, 145], illustrates that among the RAS-MAPK syndromes there is a continuous clinical spectrum and that occasionally there is a complicated clinical overlap. Borderline cases exist that are clinically difficult to diagnose, especially during early childhood, as some of the syndromes are progressive and distinguishing features do not appear until later. Because there are important differences among the RAS-MAPK syndromes, e.g. different types of cardiac complications and risk for malignancy, a correct diagnosis is essential to predicting the prognosis and to offering suitable clinical management of each patient.

To facilitate diagnosis, there is an ongoing discussion about what the distinguishing features are in the various syndromes [52, 53]. Alternatively, a genetic investigation facilitates the decision of diagnosis. In fact, a gene-based diagnosis was suggested in 2007 for CS [53, 60]. This proposal, according to which a diagnosis of CS requires a HRAS mutation, was met with general agreement one year later [108]. Consequently, the borderline CS patients with a mutation in BRAF, KRAS or MEK1, which was summarized in Paper I, have all been re-diagnosed as CFC or NS. A gene-based diagnosis has also been discussed for NS and CFC [53, 55, 60], however, the reason a gene-based diagnosis was implemented early on in CS is attributed to HRAS-positive patients’ increased risk for cancer, entailing that these patients require specific clinical surveillance. Although the overall knowledge about genotype-phenotype correlation in RAS-MAPK syndromes still needs further investigation, there are already examples also regarding NS where the underlying genetic defect governs how the patient should be clinically surveyed, like in HRAS-positive CS patients. These examples justify the introduction of a gene-directed diagnosis also in other RAS-MAPK syndromes. For instance, NS patients with a PTPN11 mutation have a lower risk of developing hypertrophic cardiomyopathy than do NS patients with RAF1 mutations. Hence, RAF1-positive NS patients should follow a clinical management programme that takes into account the increased risk of developing hypertrophic cardiomyopathy. Furthermore, adding the knowledge about the specific mutation is also informative with respect to surveillance, e.g. the suggested predisposition of developing JMML in NS patients with a PTPN11 p.T73I mutation [39, 96]. The emerging information about which gene and/or mutation causes a specific complication, such as malignancy, mild or severe mental retardation or cardiac complications, is of great value for patients, their families and the health care system when it comes to surveillance and prognosis prediction.
The gene-based diagnosis in CS and the discussions about implementing this type of diagnosis also in NS and CFC were based on the general opinion that CS, NS and CFC are distinct disorders, hence a mutation in \( \text{e.g. } \text{BRAF} \) would lead to a CFC diagnosis \([53, 60]\). However, considering the results in Paper I, a gene-based diagnosis used in a similar way as in \( \text{HRAS} \)-positive CS patients is not possible in \( \text{BRAF} \) due to the complex genetic overlap. Therefore in Paper I, the discovered genetic complexity linking two disorders to the same gene, \( \text{e.g.} \) both NS and CFC associated with \( \text{BRAF} \), was taken into account when a different type of gene-based classification was suggested. Paper I proposed that a gene-based diagnosis of RAS-MAPK syndromes should include the affected gene after the molecular investigations have been performed, hence CS should become CS-\( \text{HRAS} \) associated and the various disorders related to defects in the \( \text{BRAF} \) gene should be denoted as NS-\( \text{BRAF} \) associated, CFC-\( \text{BRAF} \) associated, etc.

Molecular investigations that could form the basis for a gene-based diagnosis are facilitated by our increasing knowledge about genotype-phenotype correlation, as different genes can be prioritized in the screening. Furthermore, the identification of hot-spots regions in the various genes, \( \text{e.g.} \) \( \text{BRAF} \) exon 6 and 11-16, \( \text{RAFI} \) exon 7, 14 and 17, exon 2 and 3 of the \( \text{MEK} \) genes, the PH-, DH- and REM-domains of \( \text{SOS1} \) and \( \text{KRAS} \) exon 2, 3 and 6, also speeds up the process of molecular confirmation of the diagnosis. However, as illustrated in Paper I, it is important to screen additional exons to identify all aberrations.

**Paper III**

**Subjects**

In Paper III, a patient with an atypical and severe form of NS was investigated. The atypical findings include Arnold Chiari type I malformation, hypoplasia of the corpus callosum, syringomyelia, malrotation of the bowel and hydronephrosis. Furthermore, the patient as well as some otherwise healthy family members – the father, sister, uncle and cousin – presented with multiple CAL spots.

To investigate clinical variability in NS, which in this patient was manifested as a severe and atypical NS, investigations using sequencing, dHPLC, MLPA and haplotype analysis of several RAS-MAPK candidate genes were performed.

**Results and discussion**

Initially, \( \text{PTPN11} \) was screened for mutations and a previously reported mutation, c.853 T>C; p.F285L, was identified in the NS patient, though this
mutation was absent in the family members presenting with CAL spots only. Previous reports have demonstrated the p.F285L mutation in a few patients, and none of these patients presented the additional complicating features that affect the index patient described here. In fact, the patient in Paper III is the first reported NS patient with Arnold Chiari type I malformation in which a mutation has been identified.

Thus far, there have been no functional studies of the p.F285L mutation. However, one study classified a large number of mutations according to the predicted effect on protein function [146]. The p.F285 residue was categorized as having a role in maintaining the overall PTP structure together with, e.g., p.N308. The p.N308D was subsequently selected for functional assays as being representative of the group that included p.F285. The phosphatase activity of p.N308D under basal conditions as well as with stimulation was increased compared to wildtype, but was among the less active mutations, which has been repeatedly observed [97, 98, 100, 146]. Furthermore, the p.N308D mutation was shown to increase cell proliferation only upon EGF stimulation and not in unstimulated cells, whereas two other PTPN11 mutants increased cell proliferation also without stimulation [97]. Taken together, these studies indicate that the p.N308D mutation is a somewhat milder mutation, which was confirmed in the p.N308D knock-in mouse model displaying a mild phenotype [100]. Assuming that the classification of mutations in the study by Tartaglia et al. [146] is correct and that the effect of p.N308D is representative of the group, the p.F285 mutation should also have a milder effect than other PTPN11 mutations. Although highly speculative, these assumptions suggest that the severe phenotype of the present NS patient is not solely due to the p.F285L mutation.

Because p.F285L did not segregate with the CAL spots trait in the family, this indicated that the CAL spots in the NS patients were not an associated feature of NS, as they are in 10% of NS patients, but rather represented a different genetic trait. Hence, the study in Paper III demonstrates that clinical variability in NS can be attributed to several different genetic defects. The consequences of an additional genetic trait in the index patient could provide an explanation for her severe and unusual symptoms. With this aim in mind, that is, to explain the severe symptoms in the NS patient as well as to possibly identify the genetic defect underlying the skin manifestation, further investigations were performed. The hypothesis with regard to the CAL spots trait in the family included NF1-associated disorders, i.e. NF1, NFNS and autosomal dominant CAL spots trait, which has been linked to the NF1 gene in one family [72] as well as NFLS. Clinical examination of the family revealed that they did not fulfil the clinical criteria for NF1 [63]. The molecular investigations of the candidate genes NF1, SPRED1 and SPRED2, whose protein, SPRED2, forms heterodimers with SPRED1, suggested that these genes were not involved in the CAL spots trait. These results support previ-
ous suggestions concerning the existence of an autosomal dominant CAL spots trait that is not linked to the NF1 gene [147, 148].

Interestingly, in NF1 a previously reported rare polymorphism c.5425C>T; p.R1809C was identified in the index patient [120], although it did not segregate with the CAL spots trait. The affected arginine is a well-conserved amino acid among species and a computational prediction of the p.R1809C substitution suggested it to be damaging. Bearing this in mind, this polymorphism is noteworthy with respect to the repeated suggestions of modifying genes/various genetic backgrounds that are involved in the clinical variability in RAS-MAPK syndromes [67, 99, 100]. Possible modifying loci could include polymorphisms/mutations in known RAS-MAPK-syndrome genes, such as the NF1 p.R1809C variant that was identified in the index patient. This hypothesis has been illustrated in a patient displaying NFNS who was shown to harbour two different mutations, one in NF1 and the other in PTPN11 [136]. Alternative modifying factors could be polymorphisms/mutations in other RAS-MAPK-pathway genes or in other components involved in cell signalling. One such example has been demonstrated in which a functional polymorphism in the EGF gene was shown to modify the NF1 phenotype [138]. Because the GG-genotype in position +61 of EGF was associated with an earlier onset of CAL spots, this modifying locus was investigated in the family to possibly explain the CAL spots and the severity of the symptoms in the NS patient. However, the results showed no segregation of a GG-genotype in family members affected by CAL spots, excluding this polymorphism’s involvement in the phenotype in the family or in the NS patient (unpublished data).

Taken together, clinical variability in NS could be a consequence of a different genetic trait/defect present in the patient, as was shown in Paper III, where the CAL spots in the NS patient represented a different genetic entity. Furthermore, the results suggested that the atypical and severe phenotype in the NS-affected index patient is an additive effect that could have two explanations. One is the presence of two different genetic traits in the index patient, NS caused by the p.F285L mutation in PTPN11 and the autosomal dominant CAL spots trait. An alternative, speculative, hypothesis is that the NF1 variant, p.R1809C, acts as a modifier in the index patient.

Paper IV

Subjects

The family presented in Paper IV was first described in 1995 as having NS and CAL spots [149]. At that time, seven family members were investigated. In Paper IV, the family was extended and 12 family members were clinically and molecularly re-investigated.
In the same paper from 1995, an additional family was described as having LS. Molecular analysis of this family identified the recurrent LS mutation c.836A>G; p.Y279C in PTPN11 (unpublished results).

Results and discussion

The clinical re-investigation revealed NF1 characteristic symptoms such as Lisch nodules and axillary freckling in several family members. Thus, the family fulfilled NF1 criteria and the diagnosis was reconsidered as NFNS. The NFNS phenotype has been associated with a lower incidence of Lisch nodules, neurofibromas, internal tumours and axillary freckling and a higher incidence of heart defect and facial dysmorphism in comparison with NF1 [70, 71]. The present family’s symptoms support these correlations, although in a review of all reported NFNS cases in Paper IV, this family stands out with its complete absence of neurofibromas in all adults.

The re-examination of the family emphasizes the importance of using slit lamp examination in RAS-MAPK patients displaying CAL spots. The CAL spots are present to various extents in several of the RAS-MAPK syndromes, and regarding NFNS, NS, LS and NFLS, the clinical overlap complicates the diagnosis especially if other distinguishing features, e.g. neurofibromas, are not present. Another important aspect regarding the clinical re-examination of this family was the change in affection status in two individuals (described as II:10 and IV:3 in Paper IV). In the previous paper from 1995 [149], these individuals were considered unaffected, but in Paper IV several lines of evidence are presented that justify their being classified as affected. The reason they were considered unaffected in the previous paper could be a consequence of not having performed a slit lamp examination as well as the absence of congenital heart defect in both of them. It has in fact been discussed that there has been an underestimation of affected NS patients, because previously the presence of heart defect was often considered a prerequisite for the diagnosis [43].

The explanations behind NFNS involve four hypotheses [70], where the hypothesis that NFNS is a variant of the clinically variable NF1 has been confirmed in the majority of patients [71]. Therefore, screening by sequencing of the NF1 gene was prioritized. A novel missense mutation was identified in exon 24, c.4168C>T, p.L1390F, which segregated in the family. Based on cDNA analysis and previous functional characterizations (see Paper IV for references), the mutation was suggested to affect the catalytic process of neurofibromin.

To further investigate whether NFNS could be a consequence of two coincidental disorders, as has been proven rarely [136], further analysis was performed. Hitherto, only PTPN11 has been investigated in this context [71, 131-133]. Paper IV extended the investigation of this hypothesis to include additional candidate genes. However, the analysis of PTPN11, RAF1, KRAS,
SOS1, BRAF, MEK1, MEK2 and SPRED1 using sequencing or haplotype analysis revealed no additional aberrations in the family. Hence, the NFNS phenotype is most likely an effect of the p.L1390F mutation, supporting previous results [71] in which the NFNS phenotype was shown to be a variant of the clinically variable NF1.

Regarding the spectrum of genetic defects in NFNS, a difference in the type of aberration had previously been observed in comparison with NF1 [71]. The NFNS phenotype was more often associated with in-frame defects (missense or small in-frame deletions) than was NF1. Furthermore, a clustering of NFNS mutations in the GAP-domain of NF1 was observed [71]. The present paper reviewed all previously reported cases and confirmed these findings, where ~40% of NFNS are caused by in-frame defects vs. ~10% in-frame defects in NF1. Consequently, in Paper IV, a screening of NFNS patients was proposed that would target the GAP-domain first. Because the NF1 gene is very large and no clear hot-spots for mutations exist in the gene, studies like those presented in Paper IV are important clues for future possibilities to prioritize screening of smaller regions of the gene in patients with a certain clinical phenotype. One interesting group of patients that could be worth analysing for mutations in the GAP-domain are those displaying NS features and in addition CAL spots that lack mutations in the other RAS-MAPK-syndrome-related genes.
Concluding remarks and future perspectives

Although the past years of intense international efforts have greatly increased our knowledge about the RAS-MAPK syndromes with respect to both clinical and molecular aspects, much remains to be elucidated. The results presented in the present thesis have contributed some important findings about the RAS-MAPK syndromes.

First of all, *BRAF* was shown to be involved in NS and *SOS1* in the pathogenesis of CFC, hence an increased genetic complexity was demonstrated among the syndromes in comparison to what was previously believed [42, 92]. This genetic complexity, where several of the disorders have been shown to be allelic, thus far involves *KRAS*, *BRAF*, *SOS1* and *MEK1* (summarized in Fig. 15), but this might not be the final word. Paper I also illustrates the remarkable clinical complexity found among patients with NS and CFC. Based on the results in Paper I, a gene-based diagnosis was suggested that should include the gene in the name.

Second, the screening of genetic defects in NS in Paper II revealed that gene dosage imbalances of *PTPN11*, *KRAS*, *BRAF*, *RAF1*, *SOS1*, *MEK1* and *MEK2* are not a major contributor or are, at the most, a rare event in the pathogenesis of mutation-negative NS patients.

The next aspect dealt with here involved clinical variability in the RAS-MAPK syndromes, and in Paper III and IV, this issue was investigated from a genetic point of view.

In Paper III, it was shown that CAL spots, reported to be present in 10% of NS cases [40], are not necessarily a part of the NS phenotype, but can be a distinct genetic entity. Furthermore it was suggested that the severe phenotype of the investigated NS patient was attributed to other genetic aberrations/modifying loci in addition to the identified *PTPN11* mutation.

Finally, in Paper IV, the hypothesis of the involvement of additional candidate RAS-MAPK-syndrome-related genes in causing the combined phenotype of NF1 and NS was assessed for the first time. The results suggested that the investigated genes are not involved in the NFNS phenotype. The identification of a novel *NF1* mutation in the GAP-domain confirmed the study by De Luca et al. showing that NFNS was caused by defects in *NF1* in the majority of cases [71]. Hence, NFNS represents a variant of the clinically variable NF1. The involvement of the GAP-domain in the pathogenesis of NFNS was emphasized in Paper IV and mutational screening of NFNS patients was suggested to start with this domain.
As mentioned previously, despite our increased understanding of RAS-MAPK syndromes, to which the present studies have contributed some aspects, there are a number of unanswered questions.

First, we still lack knowledge about the genotype-phenotype correlation which is of crucial importance to predicting the prognosis of patients more accurately and to offering suitable surveillance.

Second, what are the genetic causes in remaining patients with the RAS-MAPK syndromes? In addition to the hypothesis of gene dosage imbalances assessed in Paper II, which must be further investigated before it can be excluded, other alternative hypotheses exist. One is the involvement of regulatory sequences, promoters, etc., in the RAS-MAPK syndromes. This has hitherto not been assessed as a possible explanation with one exception, the 5’UTR of \( NF1 \) [150]. In a large cohort of NF1 patients, the promoter and the 5’flanking region of \( NF1 \) were investigated, but this revealed no pathological mutations. Another hypothesis explaining remaining RAS-MAPK patients is the involvement of other RAS-MAPK-related genes, which has partly been assessed by several research groups, as summarized in Paper II. One of the previously screened candidate genes is \( DUSP6 \) [86], which is highly expressed in heart tissue and encodes a negative regulator of ERK. Interestingly, the chromosomal location of \( DUSP6 \) is 12q21.33, which is within the deleted region that was identified in a few patients resembling CFC. This region, 14-19 Mb large at 12q21.2-q22, was suggested to be a CFC candidate region, prior to the identification of mutations in RAS-MAPK-related genes [82, 83]. The fact that deletions covering the \( DUSP6 \) gene display a
CFC-like phenotype fits the model of a loss-of-function mutation in a negative regulator of ERK, which could consequently have an increased RAS-MAPK signalling similar to loss-of-function mutations in NF1 and SPRED1. Hitherto the involvement of DUSP6 has been investigated in 87 NS patients and in 8 CFC patients (previously screened only for PTPN11 mutations) [86]. Hence the involvement of DUSP6 is still an unresolved question. The number of other possible candidate genes for the RAS-MAPK syndromes is substantial when observing the complex involvement and interactions of several molecules in the RAS-MAPK pathway, e.g. KSR, SPRY, SPRED2, GRB2 and IQGAP, among others. A screening of candidate genes would be feasible with careful selection of genes based on, e.g., expression pattern and intracellular involvement in the RAS-MAPK cascade. The questions is how many genetic factors are involved in causing the remaining RAS-MAPK patients, and will there be further genetic overlap between various RAS-MAPK syndromes in these hitherto unknown genetic factors?

A third interesting aspect, which needs further investigation, is the functional characterization of the hitherto identified mutations. The overall pathogenetic mechanism seems to be a gain-of-function of RAS-MAPK signalling, where positive RAS-MAPK proteins (SHP2, SOS1, KRAS, HRAS, BRAF, RAF1, MEK1 and MEK2) enhance their function and negative RAS-MAPK proteins (Neurofibromin and SPRED1) lose their function, which consequently leads to increased RAS-MAPK signalling. Interestingly, the question about LS caused by PTPN11 mutations, which was long considered to have a loss-of-function effect, was finally resolved when a gain-of-function phenotype was detected despite decreased enzyme activity [101]. Despite the general gain-of-function effect in RAS-MAPK syndromes, a few mutants in different genes, KRAS, BRAF and RAF1, showed kinase impairment or no activation of downstream molecules [58, 85, 87]. One suggested explanation of the kinase-impaired BRAF mutants has been that these could signal via RAF1, as has been shown in cancer [34, 42]. Further investigations are needed to clarify whether or not these are true loss-of-function mutants.

Overall, the functional studies of RAS-MAPK syndrome mutants are still at a basal level, except for NF1 and the animal models of the cardiac defects caused by PTPN11 [12, 99, 100, 139]. To date, most of the functional assays have primarily focused on the molecules within the RAS-MAPK pathway, and these studies have repeatedly observed a prolonged and increased RAS-MAPK signalling with various mutants associated with the RAS-MAPK syndromes [87, 89-91, 97, 99, 143]. Interestingly, the various organs affected in the RAS-MAPK syndromes correlate partly with the organs that demonstrated transient ERK activation during mouse embryogenesis, i.e. the peripheral nervous system, developing blood vessels and primordium of the eye, ears and heart [30]. However, a more detailed explanation of the mechanism and the consequences of the disturbed signal duration and
strength in RAS-MAPK syndromes is an issue for further research. Such functional studies should address the various cell types involved in different aspects of the RAS-MAPK phenotype, which substrates mediate the effect between the increased/prolonged ERK signalling and the phenotype in RAS-MAPK syndromes, and finally the specific effect in the studied cell types, e.g. increased cell proliferation.

Another important aspect of the RAS-MAPK syndromes is the fact that the RAS-MAPK pathway is implicated in many types of cancer. The vast majority of germline mutations involved in RAS-MAPK syndromes do not overlap with somatic mutations associated with cancer, and functional studies have shown a general higher enzyme activity with the somatic mutants than with germline mutants. Traditionally, NF1, CS and NS have been regarded as cancer predisposition syndromes, although the latter only has a low risk. Nevertheless, careful observation of all RAS-MAPK patients with a mutation that has also been reported in cancer, or if the genetic defect affects a well-known cancer gene such as *BRAF*, would be wise until more knowledge has been gained about the association between cancer and RAS-MAPK syndromes in general.

The fact that the same pathway is associated with cancer is also of benefit with respect to the enormous number of studies aimed at understanding this pathway and the consequences of disturbed signalling. For instance, several inhibitors against increased signalling are under development in the context of cancer therapy [33]. It will be truly exciting to see whether these inhibitors can be used to treat some of the progressive symptoms associated with RAS-MAPK syndromes, as has been suggested [42, 115]. However, caution must be taken not to induce high expectations in patients and families. Initial investigations to test the possibility of treatment for CFC have recently been conducted where the inhibitor U0126, which selectively targets MEK1 and MEK2, was used against a few CFC-associated MEK mutations. The study revealed promising results, as all the tested MEK mutants displayed a reduced level of phosphorylated ERK [115]. A possible treatment for the learning disabilities in NF1 was suggested in a mouse model that did not involve cancer inhibitors [139]. This is an interesting issue for future research, and perhaps it will be of importance also for the other RAS-MAPK syndromes. The possibility to ameliorate the reduced stature in NS patients using GH treatment has shown a positive effect on final height [46, 47]. However, when taken into account the involved gene in NS patients, conflicting results exist as to whether or not there is a reduced response of GH therapy in *PTPN11*-positive patients vs. *PTPN11*-negative patients [45, 46]. It has also been suggested that in NS patients with mild GH resistance (normal GH secretion but low IGF-1 and IGFB3 levels), a different therapeutic strategy involving IGF-1 could be worth testing [45]. Altogether, the issue of genotype-phenotype correlation with respect to short stature and the importance of genotype for the outcome of growth hormone therapy require fur-
ther investigation, both from an economical point of view as well as for genetic counselling and prognosis prediction.

In conclusion, the studies presented in the present thesis have demonstrated that the RAS-MAPK syndromes are remarkably complex both clinically and genetically, further exemplifying that “simple” monogenic disorders are not that simple. Recent research has increased our understanding of both clinical and molecular aspects of RAS-MAPK syndromes, enabling a genetic confirmation of the clinical diagnosis in a large number of patients in which the diagnosis was previously uncertain. However, the knowledge we have today is most likely only the tip of the iceberg.
Populärvetenskaplig svensk sammanfattning

I denna avhandling har en grupp av sjukdomar studerats, de s.k. RAS-MAPK-syndromen, som orsakas av medfödda skador i arvsmassan. RAS-MAPK-syndromen inkluderar sex syndrom som uppvisar stora likheter med varandra vad det gäller symptombilden, men även viktiga skillnader. Man har kunnat påvisa skillnader i cancerrisk, olika svårighetsgrader av mental utvecklingsstörning och olika typer av hjärtfel bland dessa syndrom. Likheten bland de sex syndromen försvårar bedömningen av patienterna, särskilt hos små barn och i många fall har man ej varit säker på vilket av dessa sex syndrom som barnet drabbats av. Att få en korrekt diagnos är av största vikt för att kunna erbjuda särskilda vårdprogram och kontroller, där man t.ex. tar hänsyn till risken för svår/lätt utvecklingsstörning och för att tidigt upptäcka cancer. För att underlätta diagnosbedömningen och i framtiden kunna uttala sig om prognosen med större säkerhet för dessa patienter och deras familjer, krävs det fortsatta studier om RAS-MAPK-syndromen.


De senaste åren har en mängd framsteg gjorts om hur man ska klassificera och särskilja/gruppera dessa syndrom och studierna i denna avhandling har bidragit till detta. När dessa studier påbörjades hade man nyligen visat att RAS-MAPK syndromen beror på skador i en och samma livsnödvändiga process, cellsignalering. Under fosterutvecklingen sker en rad olika processer som så småningom ska leda till en komplett människa med fungerande organ och vävnader. Dessa processer är bland annat beroende av att cellerna...
i embryot kan kommunicera med varandra och reagera på yttre stimulering, t.ex. olika hormoner eller tillväxtfaktorer. När en cell stimuleras av yttre faktorer sker en överföring av informationen från cellens yta till dess inre, cellkärnan, där arvsmassan finns. Arvsmassan i cellkärnan reagerar på den yttre stimuleringen och svarar genom att t.ex. producera nya proteiner. De nya proteinerna leder till att cellen förändras, t.ex. att den utvecklas till en nervcell. Om signalöverföringen från cellens yta till dess arvsmassa inte fungerar som den ska p.g.a. skador i olika komponenter, så utvecklas inte cellen som den bör.

Att RAS-MAPK-syndromen orsakas av skador i en och samma signalringsprocess förklarar de kliniska likheterna hos syndromen. Däremot är olika komponenter i denna signalringsprocess inblandade i de olika syndromen och detta förklarar delvis de skillanden som finns hos syndromen. När studierna i denna avhandling påbörjades var det bland annat känt att Noonans syndrom orsakas av skador i fyra olika arvsanlag som bildar fyra proteiner i signalringsprocessen och att kardio-facio-kutant syndrom orsakas av skador i tre andra arvsanlag samt ett av de arvsanlag som även var skadat i Noonans syndrom. Vidare var det känt att Neurofibromatos typ I orsakas av ytterligare ett annat inblandat arvsanlag.


I arbete III och IV studerades orsaker till kliniska variationer hos patienter med t.ex. Noonans syndrom och de som uppvisar både Neurofibromatos typ I och Noonans syndrom. Resultaten visade att hos vissa patienter med Noonans syndrom kan de olika symptomen orsakas av skador i flera arvsanlag. Däremot hos de patienter som har en blandad symptombild av Neurofibromatos typ I och Noonans syndrom bekräftade resultaten att majoriteten av dessa patienter bär på skador endast i det kända arvsanlaget för Neurofibromatos typ I.

Sammanfattningsvis har resultaten från dessa arbeten bidragit med fler bitar till RAS-MAPK-pusslet. Fler patienter har fått sin diagnos säkerställd och i framtiden kommer genetiska studier, såsom i denna avhandling, att öka kunskapen om vilken skada på arvsanlagen som leder till vilka komplikationer, t.ex. cancer eller svår utvecklingsstörning, och därmed leda till en bättre och säkrare vård av patienterna.
I would like to express my sincere gratitude to Uppsala University and the Department of Genetics and Pathology, Rudbecklaboratory, for creating a stimulating research environment.

I am very grateful for all the financial support to these studies and to myself by the foundation of Sävstaholm, the Swedish Research Council, the Borgströms foundation, foundations at the Medical faculty of Uppsala University and the Göransson-Sandviken foundation through Gästrike-Hälsinge Nation.

I also wish to thank all the patients and their families for participating in these studies.

Many people at Rudbeck have contributed to this thesis in many ways. You make Rudbeck a wonderful place to work at, thank you all! In particular I would like to thank:

My supervisor Marie-Louise Bondeson for giving me the opportunity to study the interesting field of clinical genetics, for believing in me and encouraging me. You have contributed to my development and independence in several ways and I would like to thank you for being a wonderful supervisor and a nice person. Thank you also for supporting me in combining science and family life.

My co-supervisor Göran Annerén for giving me the possibility to study clinical genetics, for encouragement and for believing in me. Thank you for being a nice and supportive person, for all the clinical input in my studies, for giving me the chance to follow your interesting work in real life and for giving me the opportunity to present my work in various occasions.

Present and former members of the group; Christina, for being a good friend and reliable person that I could share good moments with and less good moments with, both on a professional and private level. Sara, for being a nice and reliable person from the first day you came to the group and for our uncomplicated collaborations. You have a true “hök-öga” which I have appreciated very much, especially when reading this thesis! Lotta Thuresson, for sharing your knowledge with me and teaching me so much about array-CGH and MLPA. Thank you also for being a nice person, all the chats and company in Leiden. To my former students/project workers, Calle S, Emelie S, Ulrika G and Malin E for being good and nice students that have helped me in my projects and learned me a lot.
All the people in the Clinical Genetics section for interest in my work and a friendly atmosphere. A special thanks to all nice people in the “DNA-lab”, Marita J and Ann-Christine A for helping me with all sorts of things, answering all my questions, for nice chats and caring about me. This thesis would not have been possible without you!

Thanks to present and former members of the Dahlgroup and Niklas Dahl, for adopting me to your group meetings and for inviting me to dinners and kick-offs. A special thanks to Ann-Sofie, Miriam and Malin for all the nice breaks and lunches and for sharing ups and downs as a PhD-student and as a mother. Johanna, Jocke, Jens, Jitendra, Hasse and Larry for nice chats and sharing your knowledge. A special thanks to Larry for nice company in Bertinoro and for taking so good care of me during the trip. Present and former members in my room – Anh-Nhi for being a good teacher when I did my master degree, for becoming a good friend and for sharing nice evenings with our families. Mehdi, for being such a nice and reliable person and for always trying to help me and answer my questions. Alvaro and Patricia for nice company and sharing the life of a PhD-student and nice dinners in the early days. Ola W, for being so nice and caring. Madhu, for being friendly and for lending me your computer when I was close to panic. Ola S for asking me if I wanted to join for a coffee at early mornings and nice time at the ESHG-meeting in Barcelona. Lena Åslund for being a nice person and for taking me as a master student once upon a time. All administrators for always being so helpful and kind. Members of the core facility/ the Genome Centre that has helped me throughout the years and also the Array facility at the hospital.

Thank you all my other co-authors for good collaborations and for sharing your knowledge with me; in particular I want to thank, Bo Strömberg, Jan Gustafsson, Gerd Holmström, Judith Allanson and Masood Kamali-Moghaddam.

Friends outside the lab, from my undergraduate studies and later on also their partners, Hanna F, for being a true friend, for all our deep-talks and for sharing so many unforgettable moments both early on and lately with our families. I miss you here in Uppsala and I will always remember the years in Flogsta/Edinburgh. Lina S for being a true friend that I can always share my thoughts with. Thank you for all the nice times both during the first years and lately with our families. Malin E for the time together in Edinburgh and together with Hanna J sharing the years at Rudbeck. Cissi P and Malin H, for all the nice times these years. Janne O, for being a good teacher and a nice friend outside the lab.

To my new family, Molin/Antlind, for love and support. In particular my mother in law, Siv, for being so nice to me, for all the help during hectic periods and for encouraging and supporting me to do my things. My father in law, Sven-Erik for your interest in my work and so many nice and relax-
ing days at “Flygplansgatan”. My sisters in law, Eva and Ingrid, and their families for sharing so many nice occasions and for supporting me.

To my Nyström family, I love you! My parents Lisa and Leif, for endless love, encouragement and support, ever since I was born. I would not be here without you. A special thanks for all “skytteltrafik” these last months that saved me. To my sisters, Lara and Marilena, and their families for all your love, support, help with all sorts of things, encouragement and always standing at my side! Thanks for understanding me these last months when I repeatedly said “I cannot talk now”!

To my own Molin/Nyström family: Anders, thank you for everything from the moment we first met, words are not enough! I ♥ Y. My angels Nathalie and Leonora – thank you for showing me what life is all about, you are my ♥, my ☼ and my ☽!
References


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 452

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)