# Summary

Variation in structure and composition of the DNA are found throughout our genome. All types of variation are collectively called 'genomic variation'. Identification and analysis of genomic variation is important to distinguish neutral variants ('non pathogenic') from variants involved in disease ('pathogenic'). Identification of new disease genes will increase our knowledge of the molecular pathogenesis of genetic disorders. Every technical advance in genetic analysis has revealed new levels of variation, ranging from single nucleotide differences to full chromosome changes. Chromosome changes larger than 5-10 Mb in size can be detected by conventional karyotyping. Other techniques, notably Fluorescent in situ hybridisation (FISH), Array comparative genome hybridisation (Array-CGH), and SNP-arrays but also amplification methods such as Multiplex ligation-dependent probe amplification (MLPA) allow the detection of aberrations below the resolution of conventional karyotyping. These submicrosopic changes are called copy number variations (CNVs). High resolution melting curve analysis (HRMA) and Sanger sequencing are used to study alterations at the single nucleotide level (sequence variation). The application of new technologies, whole exome sequencing (WES) and whole genome sequencing (WGS) aid in bridging the technical divide between quantitative (CNV) and qualitative DNA differences (sequence variants). As new DNA methods are applied, increasing numbers of variants with unclear significance to disease (UVs) are identified and choices have to be made regarding the variants that deserve follow-up work. When the pathogenic consequence of a variant is unclear, the effect has to be studied in detail at other levels (functional studies, RNA studies, in silico analysis tools, and databases). This research described in this thesis outlines the rapid development and application of molecular techniques for detecting (pathogenic) genomic variation in the context of genetic disorders.

**Chapter 1** provides a general introduction to human genetics and the different types of genomic variation, in relation to health and disease. An overview of cytogenetic and molecular techniques was given, and the advantages and disadvantages of several techniques were discussed. Options for interpretation and classifications of variants are summarized.

In **Chapter 2**, methods to detect CNVs in the human genome are reviewed. A 1400-plex targeted CNV assay was designed to study 320 patients with intellectual disability. In 9% of the analysed cases, a pathogenic CNV was found. Furthermore, HRMA was used as a new method to confirm the presence of CNVs detected using SNP-arrays.

In **Chapter 3**, Array-CGH and SNP-array were used to delineate a *de novo* chromosome 19p deletion in a patient with intellectual disability, Split Hand Foot Malformation (SHFM) and Tetralogy of Fallot. In addition to this case study, MLPA was used to study the presence of copy number changes in genes within the deletion interval in 21 other SHFM patients. Based on its function, EPS15L1 was postulated as a candidate gene for Split-Hand-Foot-Malformation (SHFM).

In the following chapters **(4 and 5)** several applications of HRMA are discussed and show the succes of HRMA as a presequence screening technique. HRMA was used to identify MBTPS2 as the causative gene for Keratosis Follicularis Spinulosa Decalvans (KFSD). In three families (Dutch, American and English), a missense mutation was identified. Functional studies have demonstrated this mutation to be pathogenic. Studying X-inactivation (XCI), differences in allelic expression could be correlated to the clinical phenotype of carrier females. Mutations in MBTPS2 have also been described in Ichthyosis Follicularis Atrichia Photophobia (IFAP) syndrome, a syndrome that shows clinical overlap with KFSD.

**Chapter 6** describes three families with Terminal Osseous Dysplasia (TOD) where X-exome sequencing was applied to identify the causative gene, FLNA, in the linkage interval on Xq27.3-q28. The missense variant was predicted to affect splicing. Initial RNA analysis in patient-derived fibroblasts did not detect expression of the mutant allele. Additional RNA analysis in 15-year-old fibroma tissue showed alternative splicing confirming activation of a cryptic splice site, and results in an in-frame deletion. To study the pathogenetic mechanism of the FLNA mutation in TOD, a 3D protein model was built indicating that the deleted region affects or prevents important protein-protein interactions. XCI patterns in patients were established and suggest early skewing with preferential inactivation of the mutant allele is key to disease development.

**Chapter 7** illustrates the use of WES to identify the causative gene in a family with Aarskog-Scott syndrome. A branch point variant (-35 bp intronic) in FGD1 was identified that was missed with conventional Sanger sequencing. RNA analysis confirmed an effect on splicing. This work shows, although WES is not specifically designed to do so, deeper intronic variants can be detected.

In **Chapter 8** mutations in the SWI/SNF chromatin remodelling complex ARID1B were identified as the genetic cause of Coffin-Siris syndrome (CSS). One case-parent trio and two sporadic patients were studied according to a recessive inheritance model using WES. After initial analysis did not reveal causative variants, a dominant inheritance model identified De novo truncating mutations in all patients. By excluding an autosomal recessive inheritance pattern for CSS, the recurrence risk for parents with a child diagnosed with CSS could be reduced to 1-2%. Mutations in other genes affecting the SWI/SNF complex have been published in CSS patients (Tsurusaki et al, Nat Genet. 2012; 44(4):376-378), in other syndromes (Clapier et al, Annu Rev Biochem. 2009;78:273-304) and are associated with tumorigenesis (Wilson et al, Nat Rev Cancer. 2011;11(7):481-92 and Wang et al, Nat Genet. 2011;43(12):1219-23 ), indicating that mutations in chromatin remodelling factors contribute significantly to human diseases.

Finally, in **Chapter 9** the evolution of techniques, strategies for gene identification and variant interpretation are discussed.


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