From DEPARTMENT OF DENTAL MEDICINE
Karolinska Institutet, Stockholm, Sweden

Genomic and Dental Investigations of Individuals and Families with Non-syndromic Cleft Lip and/or Cleft Palate, Van der Woude and Popliteal Pterygium Syndromes

Marie Pegelow

Stockholm 2012
Opponent
Dacid Rice, Professor in Orthodontics, Institute of Dentistry, University of Helsinki, Helsinki, Finland

Examinining Committee
Filip Farnebo, Professor, Department of Reconstructive and Plastic surgery, Karolinska University Hospital, Stockholm, Sweden

Maria Ransjö, Professor, Department of Odontology/ Orthodontics, Sahlgrenska Academy, Göteborg, Sweden.

Catharina Lavebratt, Associate professor, Department of Molecular Medicin and Surgery, Karolinska University Hospital, Stockholm, Sweden

Main supervisor
Agneta L-A Karsten, Associate professor, Department of Dental Medicine/ Orthodontics, Karolinska Institutet, Huddinge, Sweden

Co supervisor
Myriam Peyrard-Janvid, PhD, Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden
**ABSTRACT**

Van der Woude syndrome (VWS) is the most common oral cleft syndrome, accounting for two percent of all cleft lip and/or cleft palate (CL/P) cases. The main characteristics of VWS are lip pits (LP), cleft lip (CL), cleft lip and palate (CLP), cleft palate (CP) and/or hypodontia (H). Popliteal pterygium syndrome (PPS) has the same orofacial characteristics as VWS, combined with systemic anomalies. In 2002, the gene responsible for VWS/PPS was found to be the interferon regulatory factor 6 gene (**IRF6**), located on chromosome 1, regio q32.2. The gene encodes for a transcription factor containing both a DNA binding domain and a protein binding domain. Non-syndromic (NS) cleft lip with or without cleft palate (CL±P) occurs without associated malformations in any other organs in approximately 70% of the affected subjects. The inheritance pattern for NSCL/P is complex, with a probable co-segregation of several genes and environmental factors. The **IRF6** gene has been proposed to be part of the cause of NSCL/P.

**The aim** of this thesis was to investigate individuals and families, mostly of Swedish and Finnish origin, with NSCL/P, VWS and PPS, with regard to their phenotypes, including dental anomalies, to detect mutations of the **IRF6** gene in the syndromic cases and to investigate whether the **IRF6** gene is responsible for the cleft phenotype in the non-syndromic individuals.

In **Study I**, 129 individuals affected with NS unilateral (U) CL±P were analysed for dental characteristics. Malformed lateral incisors were common in NSUCL, while hypodontia was more common in the NSUCLP phenotype. In the total material, hypodontia was found in 29.5% inside and in 15.5% outside the region of the cleft. Most of the existing lateral incisors were positioned distal to the cleft in both the primary and the permanent dentition. **Study II** revealed **IRF6** gene mutation in 59% of the 17 VWS/PPS families studied using direct sequencing of all exons of the gene.

In **Study III**, the **IRF6** gene was investigated in 17 Swedish NSCL/P families, using direct sequencing of the gene, in one affected and one healthy individual of each family. We could not detect any mutation in the protein-coding region of the gene. However, two non-coding SNPs – rs861019, a non-coding polymorphism in exon 2, and rs7552506, located in intron 3 – showed an association with the NSCLP phenotype.

In **Study IV**, we tested two SNPs of **IRF6**, rs642961 in the promoter and rs2235371 in exon 7 (Val274Ile), for association with our entire sample set of NSCL/P, VWS and PPS families (119 families). In all but the Finnish VWS/PPS families, the “A” allele of rs642961 was identified as a risk allele; transmission to an affected child occurred in a large majority on the same chromosome as the detected **IRF6** mutation. The SNP rs642961, located in the AP-2a binding site in the promoter of the **IRF6** gene, has previously been shown to be associated with NSCL±P but our results do not support this. However, we did find a significant risk (p=0.013) for transmission of the G-C haplotype (rs642961-rs2235371) to affected individuals in the NSCP subgroup of Swedish families. Of the 16 VWS/PPS families found to have a mutation in **IRF6** (**Studies II and IV**), 31% had a de novo mutation, that is, a mutation occurring in the proband only and not in the healthy parents.

**To conclude**, NSCL/P is a complex anomaly, where disturbed dental development is a frequent finding. NSCL/P is not dependent on a single gene, as in Mendelian inherited VWS. Dividing our material into sub-phenotypes resulted in rather small groups, but we did find a significant risk with a haplotype of **IRF6** in the NSCP group, and also an association for two SNPs of the **IRF6** gene with NSCLP. Our results on NSCL/P emphasize the need for additional evaluation of the **IRF6** gene and other genes/modifiers, to further clarify their roles in the development of the NSCL/P phenotype.
LIST OF PUBLICATIONS


IV. M. Pegelow, H. Koillinen, I. Fransson, J. Kere, A. Karsten and M. Peyrard-Janvid. Association and mutation analyses of the IRF6 gene in families with non-syndromic and syndromic cleft lip and/or cleft palate. Submitted.

These papers are reprinted with kind permission from the publishers at Oxford University Press (Paper I and III) and Nature publishing group (Paper II).
4 Results .................................................................................................................................................. 24

4.1 Study I. The prevalence of various dental characteristics in the primary and mixed dentition in patients born with non-syndromic unilateral cleft lip with or without cleft palate.

4.2 Study II. Novel and de novo mutations of the IRF6 gene detected in patients with Van der Woude or popliteal pterygium syndrome.

4.3 Study III. Familial non-syndromic cleft lip and palate analysis of the IRF6 gene and clinical phenotypes.

4.4 Study IV. Association and mutation analyses of the IRF6 gene in families with non-syndromic and syndromic cleft lip and/or cleft palate.

4.5 Combined Results Studies I-IV ......................................................................................................... 29

4.5.1 Hypodontia
4.5.2 De novo mutation
4.5.3 Clinical characteristics
4.5.4 Mixed cleft type

5 Discussion ............................................................................................................................................... 31

Dental anomalies
IRF6 gene mutation
De novo mutation
VWS/PPS mutation and genotype/allele at the promoter, SNP rs642961
VWS/PPS cases with no detected IRF6 mutation
Variable phenotypical characteristics in VWS/PPS individuals
IRF6 role in NSCL/P

6 Future perspectives ............................................................................................................................... 38

7 Populärvetenskaplig sammanfattning ................................................................................................. 39

8 Acknowledgements ............................................................................................................................. 42

9 References ............................................................................................................................................... 44
LIST OF ABBREVIATIONS

VWS  Van der Woude Syndrome
PPS  Popliteal pterygium Syndrome
LP  Lip pits
CLP  Cleft lip and palate
BCLP  Bilateral cleft lip and palate
UCLP  Unilateral cleft lip and palate
CL  Cleft lip
CLA  Cleft lip and alveolus
CP  Cleft palate
NSCL/P  Non-syndromic cleft lip and/or cleft palate
NSCL±P  Non-syndromic cleft lip with or without cleft palate
H, Hy  Hypodontia
OMIM  Online inheritance in man (www.omim.org)
DNA  Deoxyribonucleid acid
Aa  Amino acid
Bp  Base-pair
Kb  Kilo-base=1 000 Bp
Mbp  Million base pairs
SNP  Single nucleotid polymorphism
IRF6  Interferon regulatory factor 6 gene
AP2α  Retinoic-acid responsive gene
DNA-BD  DNA binding domain
Prot-BD  Protein binding domain
ORF  Open reading frame
De novo  New mutation, not inherited from parents
UTR  Untranslated region
CDS  Coding sequence
PCR  Polymeras chain reaction
LD  Linkage disequilibrium
TDT  Transmission disequilibrium test
PDT  Pedigree disequilibrium test
MAF  Minor allele frequency
OR  Odds ratio
Chr²  Chi square analysis
1 INTRODUCTION

Many characteristics of our body, as well as some diseases, are mainly or partially inherited. While most individuals have the same set of chromosomes, there are some differences in the genome that make us unique and can be passed on from generation to generation. Families and individuals with non-syndromic cleft lip and/ or cleft palate (NSCL/P), Van der Woude syndrome (VWS) and popliteal pterygium syndrome (PPS) have been studied. These congenital anomalies, their phenotypes, including dental development, their different inheritance patterns and the influence of the IRF6 gene are the foci of this thesis.

1.1 Phenotypes

1.1.1 Cleft lip and palate (CLP)

A cleft lip (CL) can vary from a small notch in the vermillion border to a total cleft of the lip, alveolus and palate. A cleft of the palate can vary from a bifid uvula to a total cleft palate (CP) extending up to the foramen incisivum. The phenotypes for total CL, CLP and CP are shown in Figure 1. A picture of a boy born with CLP is shown at 5 months of age before lip and nose surgery, and at 1 year of age, after surgical correction (palatoplasty) of the CP (Fig. 2 and 3).

![Figure 1. Phenotypes for CL/P](image-url)
Figure 2. A child born with CLP, before lip surgery (with a plate for nasal moulding in the mouth).

Figure 3. After palatoplasty, at one year of age.

Dental anomalies are common findings in individuals born with non-syndromic unilateral cleft lip with or without cleft palate (NSUCL±P). The development and eruption pattern of the permanent lateral incisor in the cleft region are altered and hypodontia (H) (52–58.5 per cent) (Tsai et al., 1998; Shapira et al., 2000) or a supernumerary lateral incisor is often found (30 per cent), (Hansen & Mehdinia, 2002). Hypodontia, i.e. missing teeth, outside the region of the cleft has been found in 27.8–35 per cent of individuals affected with NSCP (Eerens et al., 2001; Karsten et al., 2005), mainly second premolars. Maxillary hypoplasia that results in an anterior crossbite is often found in CLP and can depend on the scarring after palatal surgery according to a study from Sri Lanka (Mars et al., 1990). The lateral segment often has a diverse position on the cleft side, resulting in an altered transverse occlusal relation, i.e. a crossbite on the cleft side (DiBiase et al., 2002), (Fig. 4).

Figure 4. Rotated central incisor (21), hypodontia of the lateral incisor and crossbite on the cleft side.
1.1.2 Van der Woude and popliteal pterygium syndromes

Following Ann Van der Woude’s (1954) description of a syndrome with lip pits and oral clefts in five pedigrees, the syndrome has been named after her (Online Mendelian Inheritance in Man (OMIM) #119300). It is the most common oral cleft syndrome, accounting for two percent of all CL/P cases (Rintala & Ranta, 1981; Schutte et al., 1996). The genetically predisposed patients affected with VWS (Kondo et al., 2002) develop lip pits (LP), CL, CLP, CP and/or H. The CL/P phenotype is found in approximately 50 per cent of the cases and LP in approximately 80 per cent.

Hypodontia can be found as either the sole characteristic or together with LP and CL/P in approximately 25 per cent of the individuals (10–69 per cent; reviewed in Rizos & Spyropoulos, 2004). The clinical appearance of the LP located on the lower lip varies from an asymptomatic slight depression on the vermilion border to deep fistulas, sometimes communicating with the ducts of the underlying minor salivary glands. The LPs are usually symmetric bilateral paramedian sinuses. They can however occur as bilateral asymmetric, microform, median, or unilateral (Van Der Woude, 1954; Janku et al., 1980; Schinzel & Klausler, 1986; Rintala & Ranta, 1981; Rizos & Spyropoulos, 2004).

The distinction between clefts in the secondary palate only and those involving the primary palate and including CL±P is valid on genetic (Hilliard et al., 2005) as well as embryologic grounds, since the primary and secondary palates form a series of steps that can be influenced separately. The VWS can present a mixture of cleft types, i.e., CP only in some individuals and CL±P in others, with both phenotypes occurring within the same family (Burdick et al., 1985; Sander et al., 1993), (Fig. 5).

Figure 5. Pedigree of a family with VWS showing a mixture of cleft-types and presence or absence of LP in affected individuals, illustrating the variable expressivity of the phenotype.
The popliteal pterygium syndrome has the same orofacial characteristics as VWS, combined with systemic anomalies such as skin webb, syndactylies of hands/feet, syngnathia (intraoral tissue bands), nail deformities, ankyloblepharon (tissue bands between the eyelids) or genital deformities (OMIM #119500). The skin webb is positioned on the back of the leg and can originate from the ischial tuberosities to the Achilles’ tendons, sometimes reducing the possibility of straightening the leg (Lees et al., 1999). The syndrome is rare, with a prevalence of 1:300,000 (Froster-Iskenius, 1990).

1.2 Embryology

An understanding of normal development is necessary when studying patients with CLP. Development of the lip and nose begins during early embryogenesis. Below is an overview of the normal embryological development in the region.

Development of the human face begins with the migration of cranial neural crest cells from the dorsal region of the anterior neural tube into the facial region. During week four of gestation, five distinct facial primordia are established from the neural crest and the first pharyngeal arch; they comprise an unpaired fronto-nasal prominence that gives rise to a pair of lateral and medial nasal processes during week five in utero, the paired bilateral maxillary prominences that form the upper jaw, and the paired mandibular prominences that develop into the lower jaw. Lip pits represent vestigial remnants of the lateral sulci of the mandible; these remnants can sometimes communicate with a duct of labial glands and form a congenital fistula of the lip, starting at the end of week five in utero (Kitamura, 1989, reviewed in Rizos & Spyropoulos, 2004).

The maxillary processes emerge and during the following two to three weeks they join with the medial and fronto-nasal processes to form the primary palate and the central part of the upper lip and nose; the lateral nasal process forms the alae of the nose (Fig. 6).

Figure 6. The developing face (modified after Langman’s medical embryology).
While the primary palate is derived from the inter-maxillary segment, the main part of the secondary palate is formed by two shelf-like outgrowths from the maxillary prominence. During gestation weeks 8–12, the secondary palatal shelves tend to attain a horizontal position above the tongue, elevate and fuse in the anterior-posterior direction to form the hard and soft palates (Sadler, 2010). Failure at any of these joining points will result in a cleft lip and/or cleft palate (CL/P). In cleft lip with or without cleft palate (CL±P), associated malformations can occur, such as H or supernumerary teeth, since failure of the dental epithelia’s fusion (which normally occurs 4–6 days later than the rest of the facial processes) may lead to one tooth-bud on either side of the cleft or no tooth at all (Meikle, 2002; Hovorakova et al., 2006).

Teeth form from the ectoderm of the first pharyngeal arch and the fronto-nasal prominence and from the underlying mesenchyme (derived from the neural crest cells). Individual teeth develop from the dental lamina that starts to develop as a thickened epithelial stripe at the sites of the future dental arches of the maxilla and mandible (Thesleff, 2006). The initial development of teeth takes the form of ectodermal placodes, small thickenings of the epithelium at the site of each tooth family (Thesleff, 2003).

1.3 Epidemiology

A CL/P is one of the most common orofacial birth defects (Marazita & Mooney, 2004). The worldwide incidence is one to two babies per 1000 live births but the geographic variation is considerable (Gorlin et al., 2001; Lidral & Murray, 2004). In Sweden, the incidence of CL/P is 1.7–2.0 per 1000 live births; for CL alone the incidence is 0.4, for CL and alveolus (CLA) 0.2, for CLP and for CP respectively 0.7 per 1000 live births. Milerad et al. (1997) showed that in Sweden, 21 per cent of individuals with a cleft have associated malformations (so-called syndromic CL/P) (Milerad et al., 1997; Hagberg et al., 1998). There are over ~300 known syndromes that include CL/P (OMIM, www.ncbi.nlm.nih.gov/Omim/).

A CL±P occurs without associated malformations in any other organs, i.e., as non-syndromic, in approximately 70 per cent of cases. For CP alone the corresponding figure is 50 per cent (Tolarova & Cervenka, 1998; Schutte & Murray, 1999). A unilateral cleft is more common than a bilateral cleft according to Hagberg et al. (1998)
and, as in Study I, they show that CL and CLP occur more often in boys than in girls.

1.4 Treatment

In the developed countries, the first surgical correction of a cleft lip is done at the age of 3–5 months, to reduce the social stigma for the child, but further treatment is needed throughout childhood and adolescence, with surgery, speech improvement and orthodontics. The goal is always the child’s wellbeing. The child should be able to breath well, eat well, speak normally, grow normally and have looks that contribute to good self-esteem. In Stockholm, a CL defect is surgically corrected at around 4–5 months of age and a CP at one year. When the permanent tooth is developing and needs bone to support eruption, a bone graft is performed to the cleft-region of the alveolus. Sometimes the palate will be too short for good speech, in which case it is lengthened surgically (velopharyngeal flap) at around five years of age, combined with speech training by a speech therapist. Orthodontics is often needed if a cleft of the alveolus is present, both before bone grafting and when all permanent teeth have erupted. For VWS and PPS, surgery is also needed for correction of some of the characteristics, such as removal of lip pits. In some individuals with a retrognathic maxilla, the occlusion can be corrected with orthognathic surgery when growth has ceased, with forward movement of the maxilla (Le Fort I surgery).

1.5 Genetics

1.5.1 Mode of inheritance

Inheritance and environmental factors

The inheritance pattern for NSCL/P is complex, with several genes and environmental factors that contribute to the disease, while the effect of each of them is probably relatively small. This is known as a multifactorial effect. Environmental factors that in large-scale studies have been shown to be involved in the development of CL/P are smoking, folate-vitamin (B9) supplementation and maternal alcohol consumption (binge-level drinking) during the first trimester of pregnancy (Chung et al., 2000; Little et al., 2004a; Little et al., 2004b; Wilcox et al., 2007; DeRoo et al., 2008).

VWS and PPS are Mendelian disorders, i.e., caused by a mutation in a single gene. In these syndromes, the inheritance pattern is autosomal dominant; mutation of just one of the two alleles inherited from the parents is sufficient for the individual to develop the disease. Both VWS and PPS show inter- and intra-familial variability in the expression
of the characteristics/phenotype (variable expressivity; see Fig. 5 for example). The penetrance (the frequency with which a genotype manifests itself in a given phenotype) of VWS/PPS is high but not total (known to be 96.7 per cent in VWS) (Janku et al., 1980). The two syndromes are allelic, i.e. the same alleles can be responsible for both anomalies. The gene responsible for VWS/PPS is the interferon regulatory factor 6 (IRF6) (Froster-Iskenius, 1990; Gorlin et al., 2001; Kondo et al., 2002).

Approximately 14 per cent of clinically detected VWS/PPS cases lack the typical lower lip pits (Burdick et al., 1985) and are liable to be incorrectly classified as NSCL/P. The calculation of recurrence risk within the family will then be faulty, since VWS/PPS has a Mendelian dominant inheritance pattern. Based on these findings, families with apparent NSCL/P but segregating CP and CL/P within the family (as a mixed cleft type is common in VWS/PPS) might be screened for the IRF6 gene (Jehee et al., 2009).

Relative risk in NSCL/P
Falconer’s threshold theory helps to explain how recurrence risks vary among NSCL/P families that have a non-Mendelian, complex inheritance pattern. The relative risk (RR) compares the incidence in the general population with the risk in a family carrying high risk factors for the disease. Families with one affected child face an increased risk of their future children being affected (reviewed in Fraser, 1970). The liability for NSCL/P is polygenic and normally distributed. The larger proportion of shared risk genes among the sibs of an already affected child shifts the distribution of liability (above the threshold) towards a higher likelihood of developing the anomaly (Fig. 7), (Falconer, 1965).

Figure 7. Falconer’s threshold model (Falconer, 1965).
Odds Ratio (OR) is the ratio of two risks (probabilities), as an estimate of the likelihood that people sharing the same combination of genotypes will develop the condition in question compared to a group without this combination of genotypes. An OR of one indicates no difference between the two groups in the risk of developing the condition. If the prevalence of a known disease in the population is low, OR and RR will be almost equal. In Sweden, the prevalence of CLP is 0.02.

1.5.2 The human genome

The human genome is organized into chromosomes, densely packing the genetic material in the nucleus. It was established back in the 1950s that human somatic cells have 46 chromosomes (Tjio & Puck, 1958): 22 pairs of autosomal chromosomes and one pair of sex chromosomes (XX or XY). A single copy is inherited from each parent to form each pair.

The DNA (deoxyribonucleic acid) molecule is built as an anti-parallel double-helical structure (Watson & Crick, 1953), including four types of bases, purines such as adenine (A) and guanine (G) and pyrimidines such as thymine (T) and cytosine (C). Complementary bases from both strands (A pairs with T and G pairs with C) of the double helix are held together by hydrogen bonds. A specific part of the DNA, a gene, serves as a template for transcription and a ribonucleic acid (RNA) molecule is synthesised. A gene consists of a variable number of exons separated by introns; the introns are removed by splicing so that the final messenger RNA (mRNA) contains only the exonic parts of the gene (Strachan & Read, 2010). Via the translation step, the mRNA serves as a template for protein production, using a “three-letter” code, a codon, whereby the three neighbouring bases in the mRNA give the identity of the amino acid in the protein product (Fig. 8).
Figure 8. (A) A putative two-exon gene (shown single-stranded), (B) transcribed and spliced into an mRNA (where the T bases are replaced with U for Uracil) whose coding region (CDS), (exon 1 + 2), is then translated into a protein (C).

UTR= untranslated region, i.e., no protein-coding region.
A=Genomic DNA, B=mRNA, and C=Protein made of 6 amino acids.

The International Human Genome Project was launched in 1990 with the goal of sequencing the entire human genome. The project was completed and published in 2004, the number of genes was estimated to be 20 000–25 000 and the total genome length ~3 000 Mbp (International Human Genome Sequencing Consortium, 2004).

Variations in the genome
As regards their DNA sequence, humans are identical to ~99.5 per cent (Levy et al., 2007; Baye et al., 2010) but many genetic variations have been observed in the human genome. Such variations between individuals can be common or rare; the former are defined as genetic variants with minor allele frequencies (MAF) of at least 1 per cent in the population (generally named polymorphisms), and the latter as those with an MAF of less than 1 per cent (generally defined as mutations). Genetic variation includes single nucleotide polymorphisms (SNPs), sequence repeats (micro- or mini-satellites), as well as structural variants that include insertion and deletions, inversions and copy number variations (CNVs) (reviewed in Frazer et al., 2009).

SNPs are polymorphisms resulting in one bp substitution, insertion or deletion. An individual carries on average 3 million SNPs according to individual genome sequencing, with on average one SNP every 1000 bp of the genome and constituting 75 per cent of all genetic variation in an individual (Levy et al., 2007; Wheeler et al.,
Normally, SNPs are bi-allelic, which means that only two different bases (alleles) can be observed for that locus (Fig. 4). An individual is homozygous (HoZ) when the two alleles at a locus are the same (i.e. CC or TT), and heterozygous (HeZ) if the two alleles differ (i.e. CT) (reviewed in Frazer et al., 2009; Fig. 9).

![Diagram of SNP alleles and genotypes]

Figure 9. Example of the three possible combinations (genotypes) of the two alleles of a bi-allelic SNP.

Although SNPs affect only one bp, they can have drastic functional consequences. An SNP that occurs in a protein-coding region can change an amino acid (amino acid substitution) or introduce a premature stop codon in a protein product. Further, SNPs can affect splicing or transcription factor binding affinities in promoter, enhancer or silencer sequences. SNPs are popular genetic markers in large-scale genotyping projects investigating heritable traits (Cardon & Palmer, 2003).

Closely located genetic markers that show strong intermarker relationships are said to be in linkage disequilibrium (LD) (if they are often transmitted together, as a block) (Daly et al., 2001). The HapMap consortium studied SNPs at the population level, including allele frequencies and LD patterns in different populations, such as Europeans, Africans, Chinese and Japanese (Frazer et al., 2009; www.hapmap.org). Linkage disequilibrium is stronger in Caucasian and Asian populations than in African. This reflects the fact that some haplotype patterns across the genome were lost in population bottlenecks associated with human migration out of Africa (Frazer et al., 2009). HapMap includes information about allele frequencies and mapping positions for over 16 million SNPs (www.ncbi.nlm.nih.gov/SNP/snp_summary.cgi) in a database of common variations and maps of LD patterns. We have used this database when choosing SNPs for our ongoing genotyping analysis in our regions of interest.
1.5.3 Methods for analysing the human genome in complex diseases

Linkage analysis

Linkage studies are performed on families where a certain phenotype is observed in affected individuals. Each individual of the family for which DNA is available, is genotyped for genetic markers evenly spread throughout the genome (genome-wide) or a large chromosomal region of interest. If a region is found to have a number of shared alleles among the affected individuals that exceeds what is expected by chance at some of the genotyped loci, then those loci are said to be linked to the disease (Kruglyak et al., 1996). Linkage is a specifically genetic relationship in a family, while association (see below) is a statistical observation on the population basis (though all humans are related when traced back sufficiently far).

Association analysis

An association study is a common approach in genome analysis, at the whole genome level, (genome-wide association study, GWAS) or at the level of a candidate region. Many study designs are available for association analyses; family-based designs (extended pedigrees, parent-child trios and nuclear families) and non-family-based studies (case-control and cohort) (Cardon & Palmer, 2003). Association is based on statistics and in population case-control studies, unrelated affected and unrelated healthy are collected from a population and genotyped for a set of markers. The genotype or allele frequencies between cases and controls are compared, searching for altered frequencies between each group and differing from what would be expected by chance (if there were no association between the markers and the phenotype of interest), (Martin et al., 2000; McCarthy & Hirschhorn, 2008).

Hardy–Weinberg equilibrium (HWE) states that allele and genotype frequencies in a population remain constant, i.e. they are in equilibrium, from generation to generation, unless disturbing influences are introduced (Crow, 1988; Wigginton et al., 2005). Genetic equilibrium is an ideal state that can provide a baseline against which to measure change. Deviations from HWE can indicate inbreeding or population stratification, and even problems in genotyping.

The transmission disequilibrium test (TDT) can be used to test transmission of alleles at a specific position for a specific SNP, from the parents to the child. TDT uses the
information from heterozygous parental genotypes and its transmission (or not) to the affected child (Spielman et al., 1993). TDT compares the frequencies of transmitted vs. untransmitted alleles in affected offsprings and uses untransmitted parental alleles as controls. However, TDT has some limitations, since transmissions from homozygote parents are not useable and TDT cannot be used for extended pedigrees. On the other hand, an extension of TDT, pedigree disequilibrium test (PDT), can evaluate transmission in extended pedigrees by using data from all family members, including sibs (Martin et al., 2000).

1.5.4 Previous and current molecular studies in syndromic and non-syndromic CL/P

Syndromic: VWS/PPS

The localisation of the VWS/PPS locus to the long arm of chromosome 1, a region of 900 kb on 1q32-q41, was found in earlier linkage studies (Murray et al., 1990; Sander et al., 1993; Schutte et al., 1996). To identify the gene responsible for VWS, a direct sequence analysis was carried out on more than 20 known genes in the 900 kb critical region (Schutte et al., 2000). The mutation responsible for the VWS phenotype of the IRF6 gene was found in a pair of monozygotic twins discordant for the VWS phenotype whose parents were unaffected (Kondo et al., 2002). A nonsense mutation in exon 4 of IRF6, which was absent in both parents and the unaffected twin, was identified in the affected twin. The implication of the IRF6 gene as the VWS gene was confirmed in the same publication by the finding of mutations in this gene in 45 additional unrelated VWS families and in 13 PPS families. Mutations at the same bp in the gene could be seen in both syndromes, thus confirming that VWS and PPS were allelic.

The IRF genes are a family of genes coding for nine transcription factors that share a highly conserved DNA-binding domain (DNA-BD) and a more variable protein-binding domain (Prot-BD) and are thought to act as regulators of interferon expression following viral infection (Taniguchi et al., 2001). The IRF6 gene has a genomic size of 18 kb and contains nine exons; its mRNA is 4.4 kb long and its protein product consists of 467 amino acids, including the two conserved domains (Fig. 10; study II).
Figure 10. Structure of the IRF6 gene from the chromosomal region 1q32.2. The genomic structure of the gene is shown in (A) with its coding exons marked in grey, while UTR (untranslated) regions are shown in white. The protein (B) contains a DNA-binding domain (DNA-BD, exons 3 and 4) and a protein-binding domain (Prot-BD, exons 7 and 8), both marked in dark grey. The amino and carboxy termini of the protein are labelled N- and C-, respectively.

The first study by Kondo et al. (2002), demonstrated that the VWS mutations were equally spread between the two conserved domains of the protein, but were more often located in the DNA-BD in PPS (Kondo et al., 2002). These findings were confirmed when mutations of IRF6 were identified in 68 per cent of families with VWS and in 97 per cent of families with PPS, in total 106 novel disease-causing variants (de Lima et al., 2009). Mutations in exons 3, 4, 7, and 9 accounted for 80 per cent of the mutations (not random distribution). The VWS mutations were found to be equally spread between the two domains but more specifically found in the DNA-BD in PPS. In such a large collection of families, de Lima et al. could see that the types of mutation in VWS were evenly divided between protein truncation and missense mutations, whereas in the PPS families most of the identified mutations were of the missense type.

Molecular studies of animal models have contributed to the understanding of normal development. Mouse studies have shown that Irf6 is expressed in the epithelium of the maxillary and mandibular processes of the first pharyngeal arch – with strong expression detected in the epithelial fusion zone between the medial and lateral nasal processes and the maxillary process (Knight et al., 2006) – and in the medial edges of the paired palatal shelves immediately before, and during, their fusion (Kondo et al., 2002). Mice deficient for Irf6 have abnormal skin, limb and craniofacial morphogenesis. Irf6 has a significant role in regulating proliferation and differentiation of keratinocytes (Ingraham et al., 2006; Richardson et al., 2006). IRF6 is also
expressed in human tissue from the fusing secondary palate (Kondo et al., 2002); moreover, the phenotypes for VWS and NSCL/P resemble each other. A syndromic Mendelian disorder could possibly be a shortcut for finding genes responsible for NSCL/P, and perhaps these genes contribute to the population of non-syndromic clefts, through variable penetrance or the action of different modifiers (Stanier & Moore, 2004).

**Non-syndromic: CL/P**

To date, many loci have been suggested to affect the development of NSCL/P. Three to fourteen co-segregating genes have been statistically suggested, indicating that CL/P is a heterogeneous disorder (Schliekelman & Slatkin, 2002). Table 1 presents a summary of 14 genes (loci OFC1 to OFC13 as well as the *FOXE1* gene) shown in earlier studies to be implicated in the development of NSCL/P (www.ncbi.nlm.nih.gov/omim). The OFC6 (containing the *IRF6* gene) is the most replicated and has been verified in many studies (Table 1).

The first genome-wide scan done in NSCL/P families (with microsatellite markers) revealed a suggestive linkage to OFC1 and OFC2 (Prescott et al., 2000). A large candidate gene association study, combining Danish and Norwegian samples, containing families affected with NSCL/P, showed association to *FGF12* (mutated in cranio-synostosis), *ADH1C* (alcohol dehydrogenase) and to *IRF6* (Jugessur et al., 2009). New methods for sequencing the exons of the genes in a candidate region or at the whole genome level, so-called exome sequencing, will facilitate further research (Bamshad et al., 2011).
Table 1. Column 1: The chromosomal region, gene name (when loci in a gene) and number in OMIM for orofacial genes (OFC) 1-13 and FOXE1; Column 2: Method of analysis; Column 3: The year and author of the first published identified connection to NSCL/P.

<table>
<thead>
<tr>
<th>Locus, Gene, OMIM</th>
<th>Method of analysis</th>
<th>Identified in NSCL/P</th>
</tr>
</thead>
<tbody>
<tr>
<td>6p24 (OFC1; 119530)</td>
<td>Bacterial cloning</td>
<td>(Davies et al., 1995)</td>
</tr>
<tr>
<td>2p13 TGF(\alpha) (OFC2; 602966)</td>
<td>Association</td>
<td>(Ardinger et al., 1989)</td>
</tr>
<tr>
<td>19q13 BCL3 (OFC3; 600757)</td>
<td>Link. of cand. region</td>
<td>(Stein et al., 1995)</td>
</tr>
<tr>
<td>4q21-q31 (OFC4; 608371)</td>
<td>Linkage</td>
<td>(Beiraghi et al., 1994)</td>
</tr>
<tr>
<td>4p16.1 MSX1 (OFC5; 608874)</td>
<td>Association</td>
<td>(Lidral et al., 1998)</td>
</tr>
<tr>
<td>1q32.2 IRF6 (OFC6; 608864)</td>
<td>Association</td>
<td>(Sozen et al., 2001)</td>
</tr>
<tr>
<td>11q23.3 PVRL1 (OFC7; 600644)</td>
<td>Association</td>
<td>(Radhakrishna et al., 2006)</td>
</tr>
<tr>
<td>3q28 TP63 (OFC8; 603273)</td>
<td>Association</td>
<td>(Blanton et al., 2005; Scapoli et al., 2005)</td>
</tr>
<tr>
<td>13q33.1-q34 (OFC9; 610361)</td>
<td>Genome-wide linkage</td>
<td>(Ghassibe-Sabbagh et al., 2011)</td>
</tr>
<tr>
<td>2q33 SUMO1 (OFC10; 601912)</td>
<td>Transloc (chr2 and 8)</td>
<td>(Birnbaum et al., 2009)</td>
</tr>
<tr>
<td>14q22 BMP4 (OFC11; 600625)</td>
<td>Seq. of cand. region</td>
<td>(Suzuki et al., 2009)</td>
</tr>
<tr>
<td>8q24.3 (OFC12; 612858)</td>
<td>Genome-wide assoc.</td>
<td>(Zucchero et al., 2004)</td>
</tr>
<tr>
<td>1p33 (OFC13; 613857)</td>
<td>Association</td>
<td>(Leoyklang et al., 2006)</td>
</tr>
<tr>
<td>9q22.33-q41 FOXE1 (602617)</td>
<td>Association</td>
<td>(Birnbaum et al., 2009)</td>
</tr>
</tbody>
</table>

OFC6 in NSCL/P

In a study of the IRF6 gene (Zucchero et al., 2004), 8000 individuals (6755 from affected families) from 10 populations in Asia, Europe, and South America were screened for the SNP rs2235371, changing amino acid 274 from valine to isoleucine (C>T, Val274Ile, in exon 7). Strong evidence was found for over-transmission of the C-allele from the parent to the affected child in NSCL/P families, and individuals heterozygous for this specific polymorphism (CT genotype) were shown to have a lower risk of recurrence of CL/P than individuals homozygote for the C allele (CC genotype). The authors concluded that variations of IRF6 account for 12 per cent of the genetic contribution to CL/P and entail a threefold increased risk of recurrence in families who already had one affected child. In Asian populations, this marker (rs2235371) is more polymorphic than in populations from Europe or Africa (MAF of the T-allele of 22 per cent compared to 3 per cent in populations of European descent) (Zucchero et al., 2004).

Two studies, from Italy and Texas, respectively, showed LD between two markers of IRF6 (markers rs2013162 in exon 5 of IRF6 and rs2235375 in intron 6 of IRF6) and association of the SNPs with the NSCL±P phenotype (Blanton et al., 2005; Scapoli et al., 2005). Another study by Ghassibe et al. (2005) tested the same marker in exon 5
(rs2013162) and rs2235543 outside the \textit{IRF6} gene that were in LD with SNP rs2235371 (Val274Ile, see above, Zucchero et al., 2004), since their material (190 trios and 5 families affected with NSCL±P) did not include enough individuals carrying the Val274Ile polymorphism (low MAF) for an analysis of this particular SNP. Ghassibe et al. (2005) showed significant association with rs2013162 and NSCL±P, and also NSCL±P and the haplotype of the two common alleles at SNP rs2013162 and rs2235543.

Rahimov et al. (2008) highlighted SNP rs642961 (A-allele) as a risk allele that disrupts the binding of \textit{AP-2a} transcription factor to the \textit{IRF6}. They also showed association between NSCL±P and this SNP (G>A). Rs642961 is located in the \textit{IRF6} promoter/enhancer region. Furthermore, they demonstrated a haplotype (A-C) association (rs642961 and rs2235371) with NSCL±P, but not with NSCP. The results for rs642961 were most significant for families that included affected individuals with NSCL. In two European case/control cohorts from Denmark and Norway (in the same study), the A allele was found significantly more often in cases than controls, with the highest OR of 1.99 (95 per cent CI 1.54–2.57, p ~10–7) for the CL phenotype. These authors concluded that the presence of this common polymorphism (rs642961) within the \textit{IRF6} enhancer confers an 18 per cent attributed risk of isolated CL. To summarize, both Zucchero et al. (2004) and Rahimov et al. (2008) showed recurrences for NSCL/P with SNPs of the \textit{IRF6} gene that were higher than in the normal population (i.e. 0.1–0.2 per cent), (Gorlin et al., 2001; Lidral & Murray, 2004).
2 AIMS OF THESIS

The overall aim of this thesis was to investigate individuals and families with non-syndromic cleft lip and/or cleft palate, Van der Woude syndrome and popliteal pterygium syndrome, to diagnose the phenotypes, including dental anomalies, to detect mutations of the IRF6 gene in the syndromic cases of VWS/PPS and to investigate whether the IRF6 gene is responsible for the cleft phenotype in the NSCL/P individuals.

Specific aims

I: To follow the dental development and to investigate anomalies in the primary and mixed dentition in Swedish children born with NS UCL, NS UCLA, and NS UCLP; including the eruption pattern for the lateral incisor on the cleft side, hypodontia or supernumerary teeth both in and outside the cleft region, and to compare the transverse occlusal relation of the buccal segments on the cleft versus the non-cleft side.

II: To identify mutations of the IRF6 gene in VWS and PPS patients from 17 kindreds originating from Finland, Sweden, Norway, Thailand and Singapore, and to establish the status of novel mutations appearing in only one individual and in the last generation (de novo mutation).

III: To test the IRF6 gene for sequence variants in 17 Swedish families with NSCL/P and to relate any found IRF6 polymorphism to the cleft phenotypes or hypodontia.

IV: To detect IRF6 mutations in seven newly recruited VWS and PPS families of mixed origin. To test for association of SNP rs642961 from the IRF6 enhancer AP-2α (alpha) region, alone or as haplotype with the IRF6 coding SNP rs2235371 (Val274Ile) with individuals in NSCL/P and VWS/PPS families.
3 MATERIAL AND METHODS

3.1 Samples
All the NSCL/P patients included in this thesis were recruited by the Stockholm Craniofacial Team, Karolinska University hospital, Stockholm, Sweden (Studies I–IV) and by the Cleft Center, University Hospital of Helsinki, Helsinki, Finland (Studies II and IV).

The VWS and PPS families in Studies II and IV were recruited not only by the above-mentioned clinics but also by: Sahlgrenska University Hospital, Göteborg, Sweden; the Department of Molecular Medicine, Karolinska University hospital, Sweden; Rikshospitalet, Oslo, Norway; the National Dental Centre in Singapore; the Clinical Genetic Services, Schneider Children's Medical Center and the Raphael Recanati Genetic Institute, Beilinson Hospital Petach Tikva, Israel.

A consecutive series of 129 children born in Sweden affected with NS unilateral (U) CL±P was included in Study I. The material consisted of radiographs, dental casts, and photos from the medical records, taken at the ages of 5, 7, and 10 years. For the analysis of the transverse occlusion, 30 patients had to be excluded because of incomplete dental casts (29 individuals) and bilateral scissors bite (one individual).

In Study II, 16 families with VWS and one with PPS were analysed. In each family, individuals with one or more of the following features were diagnosed as affected: lower LP, CP, CL, CLP or hypodontia. For six families, the clinical information and positive linkage to the VWS locus on 1q32 had been reported previously (VWS-8, VWS-11 and VWS-59–VWS-62; Wong et al., 1999), while the other 11 families were newly recruited and had not undergone any linkage study.

Study III included 17 multiplex families with NSCL/P, all recruited in Sweden. One affected and one healthy individual from each family were analysed. The diagnoses, UCL±P, BCL±P, CP or H, were based on clinical examination, medical records, telephone interviews and information from other family members (as also in Studies II and IV).
In Study IV, 119 families were included: 33 multiplex families with NSCL/P and 38 trios (healthy parents and one affected child) with NSCL/P, all recruited in Sweden, 24 multiplex families with NSCP recruited in Finland, and 24 families of different origin affected with VWS or PPS (comprising 17 families analysed in Study II and seven newly recruited families). For analysis of allelic frequencies at the rs642961 locus, affected but all unrelated individuals were chosen from the above sample set of 119 families: 67 NSCL±P or NSCP (Swedish), 27 NSCP (Finnish) and 24 VWS/PPS affected individuals. Two control groups, one for each population (Swedish and Finnish) of unrelated healthy individuals, were also extracted from the same family material (35 individuals in the Swedish and 21 in the Finnish group). If an affected individual was related to all healthy individuals in the family, only the affected member in the family was chosen. Many of the individuals in the control group were incorporated in the family by marriage (i.e. married to a brother or sister of the proband’s parents).

3.2 Methods

3.2.1 Analysis of material from medical records

In Study I, x-rays, study-models and photos from medical records were analysed, x-rays were also used in Study III. The dental development and dental anomalies in patients were followed longitudinally from the primary to the mixed dentition. The modified Huddart and Bodenham index (Huddart and Bodenham, 1972; Mossey et al., 2003; Gray and Mossey, 2005) was calculated for the transverse and anterior-posterior relationships.

3.2.2 DNA preparation and genotyping

In Studies II–IV, genomic DNA was extracted from peripheral leukocytes, using a standard non-enzymatic extraction method based on salting out the cellular proteins by dehydration and precipitation with a saturated NaCl solution. Exons 1–9 of the IRF6 gene were amplified by polymeras chain reaction (PCR) as nine different fragments measuring 373-622 bp. Polymeras chain reaction is a technique for amplifying a single or a few copies of a piece of DNA to generate thousands of copies of a particular DNA sequence (Mullis et al., 1986). The PCR method relies on thermal melting, involving repeated cycles of heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA. Primers (short DNA fragments), containing sequences complementary to the target region along with a DNA polymeras (after which the
method is named), are key components to enable selective and repeated amplification. As PCR progresses, the generated DNA is itself used as a template for replication, setting in motion a chain reaction in which the DNA template is exponentially amplified.

All exons were sequenced on both strands, using the DYEnamic ET Dye terminator kit. The sequences were separated by the fluorescence-based MegaBACE1000 Automated capillary DNA sequencing system and sequences were analysed with the Sequence Analyser v3.0 software (Amersham Biosciences), (Twyman & Primrose, 2003). The DNA sequencing approach is based on dideoxy terminated primer-directed replication of single stranded DNA (Sanger et al., 1977). The method has been improved by the use of fluorescent labelled dideoxy nucleotides to enable parallel reactions for all four nucleotides (Strauss et al., 1986), followed by electrophoretic separation of products and light intensities.

To detect sequence variations, the chromatograms were inspected visually by two independent readers and the sequences were aligned using the Pregap and Gap4 softwares from the Staden package (www.staden.sourceforge.net).

To test the inheritance in trios with an affected child and healthy parents, we genotyped those trios for 20–40 SNPs with the Sequenom (Sequenom Inc.) matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry method (Jurinke et al., 2002; Twyman & Primrose, 2003). In the Sequenom analysis, the allele specific products are heated with a laser pulse into the gas phase where ionisation is achieved. Ions are then accelerated towards the detector and the time of flight (TOF) is measured (Fig. 11). The TOF is thereafter related to the molecular mass/charge ratio of the alleles. Up to 40 SNPs can be measured in one well (mass spectrum). SNPs on chromosomes 2 (Study II), 6, 7 and 15, which were previously shown to be polymorphic in the HapMap database in European individuals, were used (Studies II and IV).
In **Study IV**, to verify the identity and exact sequence of the inserted and deleted DNA fragments from the mutations detected in affected individuals from VWS-20 and VWS-21, we sub-cloned their corresponding PCR products (exons 4 and 7, respectively), using the pGem®-T Easy Vector System I, (Art. nr A3600, Promega, U.S.A). For each mutation, three mutated clones and three wild type clones were picked and further sequenced.

In the total sample in **Study IV**, SNPs rs642961 and rs2235371 were genotyped using TaqMan SNP Genotyping Assays (Livak *et al.*, 1995). PCR amplification was performed in 96-well plates and in a total volume of 10 µl. The Taqman genotyping probe (target specific for SNPs) acts with a fluorophore (5’end) and a quencher (3’end) in proximity (Fig. 12). While the specific sequence is amplified by PCR, the quencher inhibits any fluorescence signals until the degradation of the probe (cleavage of the sequence) releases the fluorophore and breaks the proximity to the quencher, thus stopping the inhibitory effect and allowing fluorescence of the fluorophore to be emitted and detected by the analysis software (Livak *et al.*, 1995).
3.2.3 Ethical approvals
Approval for Studies I-IV, (Diarienr: 2006/1243-31/4; 163/93 with complements in 2004-05-14; 2008/1434/31/3), was obtained from the ethical committee in Stockholm.

In Studies II to IV, all blood samples were collected with informed consent. Since material from Finland was included also in Studies II and IV, approval was obtained from the local ethical committees in both Finland and Sweden.

3.3 Statistical analysis
The Statistical Software (version 8; StatSoft, Inc., USA) was used for analysis of the recorded data in Study I. Differences in the proportions of the various characteristics recorded in the three groups were tested for statistical significance using the chi-square test for one sample and Pearson’s chi-square test (Chi²). In Table 11 (the transverse relation), the differences were analysed with the McNemar test and the Fisher exact test, in Table 13 (the developmental pattern for the lateral incisor) with the Sign test.

In Study III, possible LD between SNPs and building of the subsequent block structure were performed with the Haploview software (www.broad.mit.edu/mpg/haploview; Barrett et al., 2005). Marker and haplotype disease associations were tested using the family-based association test software FBAT (Lange et al., 2003). Association analysis with independent affected and healthy family members was performed. The Chi² test was used to compare cleft phenotypes and frequencies of hypodontia.
In **Study IV**, association analysis to single marker or two-marker haplotypes was performed with TDT for trios and with PDT (PDT phase v.2.4) for pedigrees (Spielman *et al.*, 1993; Dudbridge, 2003).

Allele frequencies and Hardy-Weinberg equilibrium were calculated. Odds ratios (OR) and Chi$^2$ and/or Fisher exact analysis tests were performed using the SPSS software (IBM Statistics SPSS version 19).
4 RESULTS

4.1 Study I. The prevalence of various dental characteristics in the primary and mixed dentition in patients born with non-syndromic unilateral cleft lip with or without cleft palate.

The different NS cleft phenotypes studied exhibited differences in dental expression. In patients with clefts that include the palate (UCLP), hypodontia was a frequent finding (43 per cent). Of those, 21 per cent also had hypodontia outside the cleft region. The absence, the position and the presence of a supernumerary lateral incisor in the region of the cleft are shown in Table 2.

The lateral incisor on the cleft side had an abnormal shape in 44 per cent of the individuals in the total sample. Combining information regarding the shape of the lateral incisor and the enamel of the central incisor showed that the enamel of the central was normal when the shape of the lateral was abnormal; this relationship was significant (p<0.01). Central incisor enamel defects were found in 48 per cent of the total sample.

We also found that the central incisor was significantly (p<0.05) often rotated when the lateral incisor was absent in the premaxilla, i.e. when the lateral incisor was missing or positioned in the distal segment of the cleft.

In UCLP patients, the transverse occlusal relation was altered because of the cleft. The lateral segment had a diverse position on the cleft side, compared to the non-cleft-side, resulting in a crossbite on the cleft side. The modified Huddart and Bodenham index showed a mean value of 3.9 on the cleft side, compared to 1.6 on the non-cleft side.
<table>
<thead>
<tr>
<th>Lateral incisor in the cleft region (%)</th>
<th>UCL</th>
<th>UCLA</th>
<th>UCLP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primary dentition</td>
<td>Permanent dentition</td>
<td>Primary dentition</td>
</tr>
<tr>
<td>Hypodontia</td>
<td>2</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Medial segment</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Distal segment</td>
<td>-</td>
<td>-</td>
<td>53</td>
</tr>
<tr>
<td>Supernumerary</td>
<td>45</td>
<td>21</td>
<td>47</td>
</tr>
<tr>
<td>Individuals (n)</td>
<td>n=47</td>
<td>n=47</td>
<td>n=15</td>
</tr>
</tbody>
</table>

Table 2: * In UCL, 53 per cent had a normally erupting lateral incisor in the primary dentition and 66 per cent in the permanent dentition (not shown in the table).

Hypodontia in the cleft region was found more often in UCLP than in UCL/ UCLA. Supernumerary lateral incisors were found more often in UCL/ UCLA than in UCLP.

The transition was followed individually from the primary to the permanent dentition for the lateral incisor in the region of the cleft. No significant pattern (p=0.15) was seen regarding the transition from the primary to the succeeding permanent lateral incisor, though most of the existing lateral incisors were positioned distal to the cleft.

4.2 Study II. Novel and de novo mutations of the IRF6 gene detected in patients with Van der Woude or popliteal pterygium syndrome.

IRF6 gene mutation was found in 59 per cent of the families studied with clinical symptoms of VWS/ PPS. All the detected mutations were different; six of them had not to our knowledge been reported previously. The ten mutations were all heterozygous and were found on exons 3, 4, 7, 8 and 9 of the IRF6 gene (Fig. 16).

Fifteen of the 17 families studied are of Nordic origin (two Norwegian, nine Swedish and four Finnish families). Mutations were found in the two Norwegian families and in six of the nine Swedish families. In the four Finnish families, no mutations were found in the parts of the gene we screened.

The p.Arg250Gln mutation found in VWS-5 (Fig. 13 and 14) and the p.Gln359X mutation found in family VWS-6 are affecting the Prot-BD. The DNA binding domain is affected in VWS-1 and PPS-2, with the mutations p.Lys34Glu and p.Arg84Cys, respectively (Fig. 16).
Figure 13. Pedigree of family VWS-5, ++ = mutation detected, – = no mutation found.

Figure 14. Sequencing chromatograms in exon 7 of IRF6 for the three individuals from family VWS-5: affected proband 526 (middle two panels), his father (left two panels) and mother (right two panels). A mutation is found in exon 7 of IRF6 with the three letter-code altered CGA/CAA, i.e. Arginin to Glycine (p.Arg250Gln), in individuals 525 and 526, the father and the proband, shown as a double contour in graph G. The father and proband are heterozygote GA, while the healthy mother (524) is homozygote GG. For clarity, the blue line is located at the base before the mutation. The upper panels show sequencing from the reverse DNA strand (r), the lower ones show sequencing on the forward DNA strand (f).

4.3 Study III. Familial non-syndromic cleft lip and palate analysis of the IRF6 gene and clinical phenotypes.

In the analysis of IRF6 in NSCL/P no protein coding IRF6 mutations were determined in the families examined. A total of seven polymorphic SNPs were found (Fig. 15 and 16). SNP 1, a newly detected SNP, was reported to the SNP database and named rs34743335. SNPs 1–2 were located in the promoter (-50 and -39 bp from the start codon respectively) of
SNP 3 and SNP 4 were found to be associated with the NSCLP phenotype; SNP 3, rs861019, is a non-coding polymorphism in exon 2 and SNP 4, rs7552506, is located in intron 3. Odds ratios, that the polymorphisms were more common in the affected individuals with CLP, were calculated to 3.1 for rs861019 and 5.45 for rs7552506. SNP 5, Ser153Ser, is a silent mutation in exon 5. SNP 6 and 8 are located in intron 6 and intron 7, respectively. SNPs 5 and 6 were shown to be in high LD (red squares; Fig. 16). SNP 7, rs2235371, p.Val274, is not shown in the block structure since all individuals were homozygous for the most common C-allele.

Figure 15. Linkage disequilibrium (LD) plot of the genomic region of the IRF6 gene. Red squares show 100% LD between markers, while light blue squares are non-informative. In the white squares, the numbers denote percentage LD. SNP 7 (rs2235371) is not shown in the block structure since all individuals were homozygous for the most common C-allele.

The two polymorphisms rs2235371 and rs861019 were also identified in the VWS families (Study II). The p.Val274Ile polymorphism (rs2235371) was found in four families, including three unaffected individuals. In exon 2, rs861019, the non-coding SNP was detected with frequencies similar to those reported in the SNP database (www.ncbi.nlm.nih.gov/snp).

4.4 Study IV. Association and mutation analyses of the IRF6 gene in families with nonsyndromic and syndromic cleft lip and/or cleft palate.

Mutations in the IRF6 gene with either a missense or a frameshift effect were detected in six out of seven VWS/PPS families. To find the exact sequences in two of the VWS families (VWS-20 and VWS-21), we sub-cloned the mutation products. An 8 bp
insertion in exon 4 and an 8 bp deletion in exon 7 were found, respectively, in these two families.

The *IRF6* gene was analysed for two specific SNPs (rs642961 and rs2235371; Fig. 16) that in earlier studies had been found to be associated with NSCL/P. We analysed the total sample set (Finnish and Swedish sets) and also divided the material into separate subgroups for each sub-phenotype: NSCL, NSCLP and NSCP. A significant association was detected in the sub-sample set of Swedish NSCP families with the A-C haplotype for rs642961-rs2235371, showing to be protective, (p=0.02), and the G-C haplotype, showing to be at risk (p=0.013). For the total Swedish NSCL/P sample set, no significant results were obtained when testing for transmission of single alleles or haplotypes.

A group of unrelated individuals (n=174) was gathered from the family cohort and analysed for allelic frequencies at rs642961. No individuals in the control groups were a descendant or ascendant to the proband, i.e. they were healthy and unrelated to any affected in the family. A comparison of MAF with frequencies in HapMap samples of European descent showed only minor differences (MAF of 0.243 and 0.223 respectively). Phenotypic subgroups of NSCL, NSCLP, NSCP were also analysed, in both the Finnish and the Swedish materials. No significant differences were observed for allelic and genotype frequencies between affected and unaffected in the groups of unrelated individuals.

Association analysis of our entire VWS/PPS sample set revealed the “A” allele from SNP rs642961 in the promoter of *IRF6* to be a risk allele (p=0.02), in all but the Finnish families, and when transmitted to an affected child, this occurred in a clear majority (≈80 per cent, nine out of eleven) on the same chromosome as the detected *IRF6* mutation.
Figure 16. Summary of all detected mutations and SNPs in or around the IRF6 gene from Studies II–IV. Rs642961 is located ~14 500 bp upstream of the start codon in the promoter region of IRF6.

4.5 Combined Results Studies I–IV

4.5.1 Hypodontia
Hypodontia was significantly more common in NSCL/P affected individuals (26–29 per cent: Study I and Study III), as compared with the family members without a cleft (1 per cent: Study III). In the VWS affected individuals hypodontia was a frequent finding (50 per cent: Study II).

4.5.2 De novo mutation
A de novo mutation was found in ~31 per cent of the probands in the VWS/PPS families in which an IRF6 mutation was detected (Studies II and IV).

4.5.3 Clinical characteristics
The phenotype is expressed variably in VWS/PPS. Approximately 31 per cent of the affected VWS/PPS individuals in our 24 families do not display lip pits as phenotype. In addition, four (VWS) affected individuals have hypodontia or toe-syndactylies as their only observable clinical symptom and still have a detected mutation of IRF6 (Studies II and IV).
4.5.4 Mixed cleft type

Affected individuals in the 24 VWS/PPS families (Studies II and IV) display a mixed cleft-type, with the phenotypes CP and CL±P occurring within the same family in 37.5 per cent of the material. In study IV, the corresponding figure in the 33 multiplex NSCL/P families is 6.1 per cent.
5 DISCUSSION

The *IRF6* gene has been linked to the VWS/PPS syndrome (Kondo *et al*., 2002) but has also been associated with NSCL/P (Houdayer *et al*., 2001; Zucchero *et al*., 2004; Rahimov *et al*., 2008; Jugessur *et al*., 2009). The aim of this thesis was to investigate individuals and families with NSCL/P, VWS and PPS, to diagnose the phenotypes, including dental anomalies, to detect mutations of the *IRF6* gene in the syndromic cases of VWS/PPS and to investigate whether the *IRF6* gene is responsible for the cleft phenotype in the NSCL/P individuals.

Dental anomalies

Dental anomalies were frequently found in the patients affected with NSCL/P in Study I. A peg-shaped permanent lateral incisor was found significantly more often in the UCL phenotype than in the UCLP phenotype, where hypodontia or a normal tooth was more frequent. This is in agreement with Rawadesh and Sirdaneh (2009). Disturbed tooth formation in the cleft region could be a result of the cleft itself, as this anomaly may adversely alter the tissue anatomy and the blood and nerve supply in the area of the developing tooth bud (Kjaer *et al*., 1994); this could also be related to disturbed enamel formation, which was often found on the permanent central incisor. We found that normal enamel of the central incisor and shape anomalies of the lateral incisor often coincided \( p<0.01 \), but we could not prove that there was a relation between a combination of hypodontia of the lateral incisor and disturbed enamel of the central incisor; with a combination of shape anomalies of the lateral incisor and normal enamel of the central incisor.

A missing contralateral incisor, which was found in 3.1 per cent of our material, might indicate a bilateral developmental disturbance, albeit less severe on the non-cleft side (Letra *et al*., 2007).

The clinical cleft characteristics of patients with NSCL/P resemble those of VWS and hypodontia is a common finding also in NSCL/P. The frequency of hypodontia in Swedish schoolchildren between 7 and 9 years of age was 7.4 per cent (Bergström, 1977, Bäckman and Wallin, 2001). Hypodontia in NSCL/P (26–29 per cent) was found more frequently in the affected individuals (in VWS/PPS 50 per cent, Studies I–III) than in healthy family members and not only in the cleft region. Hypodontia outside the region of the cleft was found more often in the NSCLP phenotype than in NSUCL. An
interpretation of this result could be that the frequency of hypodontia outside the region of the cleft is related to the degree to which the palate is affected (Karsten et al., 2005).

A supernumerary lateral incisor was found more frequently in the primary dentition and hypodontia more frequently in the permanent dentition, as also reported by Ringkvist and Thilander (1969). We also found that in the primary dentition, a supernumerary lateral incisor was more common in the UCL and UCLA phenotype than in UCLP. Formation of supernumerary teeth could be a result of non-fusion of the medial nasal and maxillary processes, which also leads to separation of the dental epithelia (Hovorakova et al., 2006).

The central incisor was significantly often rotated when the lateral incisor was absent in the premaxilla. This could to some extent be explained by the reduced amount of alveolar bone distal to the root of the central incisor in such cases; this situation already exists when the tooth is erupting. However, Ranta (1971), who also studied the occurrence of rotated central incisors in UCLP children, did not find any relation between such rotations and the presence or absence of the lateral incisor.

The altered transverse occlusal relation often found in individuals affected with UCLP, i.e. a crossbite, could be attributed to the surgical repair (DiBiase et al., 2002) or the disturbed anatomy. Maxillary hypoplasia frequently found in individuals with UCLP (~39 per cent with an anterior crossbite in study I) can also depend on the palatal surgical repair (Mars et al., 1990). The Huddart and Bodenham index (Huddart & Bodenham, 1972) used has been shown in former studies to be a useful tool of calculating the presence of crossbites and anterior-posterior relationships (Mossey et al., 2003, Gray & Mossey, 2005).

The time of eruption of a present permanent lateral incisor on the cleft side predisposes for the timing of auto-transplantation of bone to the cleft region (in UCLA±P); for that purpose it would be useful to be able to predict the presence of the permanent lateral incisor in the region. However, a study of each individual longitudinally failed to reveal any significant pattern in the cleft area as regards the transition from the primary to the succeeding permanent lateral incisor (p=0.15), but a common finding was that most of the existing lateral incisors were positioned distal to the cleft (Ranta, 1971; Tsai et al., 1998).
Previous studies have found that information regarding sub-phenotypes such as dental anomalies is an important tool in genetic analyses (Menezes et al., 2008 a, b). The frequency of dental anomalies in non-cleft siblings of patients with clefts has been found to be significantly higher than in a non-cleft population (Schroeder & Green, 1975). An extended clinical description of the patients that includes information concerning dental anomalies will provide new opportunities to map susceptibility loci for clefts (Vieira et al., 2008). This was a retrospective study, but future prospectively collected material will include genetic analyses combined with studies of sub-phenotypes such as dental anomalies.

**IRF6 gene mutation**

Study II reveals IRF6 gene mutation in 59 per cent of the families with clinical symptoms of VWS/PPS, which is in accordance with the original investigation by Kondo et al. (2002) reporting mutations in 55 per cent of the screened kindreds. Both studies used the Sanger dideoxy-nucleotide sequencing method (Sanger et al., 1977). The identified mutations in Kondo et al. were located in exons 3, 4, 7, 8 or 9 of IRF6, in families with both VWS and PPS. This confirms the involvement of the same gene (IRF6) in the pathogenesis of these allelic syndromes. In a later study (de Lima et al., 2009), a majority of the missense mutations and in-frame insertions and deletions found in VWS (109/207 families) were affecting the DNA-BD (64 families), (amino acid (aa) 13–113, exons 3 and 4), or the Prot-BD (45 families), (aa 226–394, exons 7 and 8), while the nonsense mutations were evenly spread over the gene (80/207 families). These authors identified five apparent hotspots in VWS/PPS. In our study, in the 16 VWS/PPS families showing IRF6 gene mutation (Studies II and IV) the mutations were observed in exons 3, 4, 7, 8 or 9, which supports the previous mutation reports for hotspots of mutations (de Lima et al., 2009). As in de Lima et al. (2009), the missense and frame-shift deletions and insertions (9/10 families) were mostly found in the DNA-BD or Prot-BD. The mutations found in our three PPS families were missense and located in the DNA-BD, now known as one of the most common PPS hot spots, p.Arg84Cys (Kondo et al., 2002; Richardson et al., 2006; de Lima et al., 2009).

**De novo mutations**

Of the mutations that were found in 16 families, 5 were de novo (Studies II and IV), which indicates that these incidents of the IRF6 gene are common, with a frequency of
~31 per cent of the affected individuals in our studies. MALDI-TOF genotyping of SNPs (Jurinke et al., 2002; see Fig. 11) was used to determine the relatedness of each of these family trios and was chosen for its ability to genotype several SNPs in a single well (multiplexing). We used SNPs known to be sufficiently polymorphic in the population. We used 20 SNPs in Study II (2 families) and 40 SNPs in Study IV (3 families) per individual, of which 19 to 20 SNPs and 31 to 35 SNPs, respectively, were informative.

**VWS/PPS mutation and genotype/allele at the promoter, SNP rs642961**
Association analysis of our entire VWS/PPS sample set revealed the A allele from SNP rs642961 in the promoter of IRF6 to be a risk allele in all but the Finnish families and when transmitted to an affected child, in a large majority of the cases this occurred on the same chromosome as the detected IRF6 mutation (80 per cent). This finding is to our knowledge new and requires further investigation of the consequences (the severity) for the patients’ phenotype, which the A allele of rs642961 may have for additive or modifier effects on the detected IRF6 coding mutation.

**VWS/PPS cases with no detected IRF6 mutation**
No IRF6 mutation was observed in eight of the 24 VWS (33 per cent) families tested. Sometimes a mutation can go undetected despite a relevant phenotype and family history. De Lima et al. found that 32 per cent went undetected in their VWS cohort of 307 families, as in our families from Studies II and IV. This observation could have several explanations. The mutation could be located in the promoter region of the gene or in uncharacterized intronic regulatory sequences located outside the regions screened in our study. Alternatively, involvement of a locus other than of IRF6 could explain part of the negative findings. A second VWS locus has been identified in a region of 1q34 in a large Finnish VWS family (Koillinen et al., 2001). However, linkage to this locus was tested and excluded in one of our Finnish families, namely VWS-17 (Koillinen et al., 2001). The range of phenotypes for VWS and PPS indicates the likely contributions of random events and genetic modifiers for IRF6-related disorders (Sertie et al., 1999).

**Variable phenotypical characteristics in VWS/PPS individuals**
In PPS-2 and VWS-10, the proband’s mother and maternal grandmother, respectively, displayed the same IRF6 mutation as the proband, but their only clinical feature of
VWS was hypodontia. In VWS-23, the mutation detected in the proband was identified in two other individuals of the family in whom toe-syndactily was the only observable clinical symptom (Fig. 17). This illustrates the condition’s variable characteristics, probably due to a low penetrance or a variable expressivity. The findings show that a VWS phenotype can be clinically missed in a pedigree if no other individual is found with a more severe phenotype than hypodontia or toe-syndactylies. This raises the question of whether this low penetrance or variable expressivity have to do with another gene interacting with the IRF6 or whether the relatives are mosaic carriers. Further investigations are needed to clarify this.

Figure 17. Pedigree of three VWS/PPS families, with variable phenotypical characteristics.

Another typical clinical finding is the mixed cleft-type that was detected in 37.5 per cent of the VWS families. This can be a valuable aid when diagnosing VWS and NSCL/P when no lip pits are present (31 per cent) (Studies II and IV) and also for deciding which families are adequate to test for IRF6 mutation.

IRF6 role in NSCL/P

Since VWS and NSCL/P have a phenotypical overlap (Houdayer et al., 2001) and because IRF6 is expressed in the midline of the normally fusing palate and is also involved in epithelial formation (Kondo et al., 2002; Richardson et al., 2006), we further investigated IRF6 allelic variations in NSCL/P families (Study III). A syndromic Mendelian disorder could be a shortcut for finding genes responsible for NSCL/P. These genes could contribute to the population of non-syndromic clefts, perhaps through variable penetrance or the action of different modifiers (Stanier & Moore, 2004).
In Study III, two non-coding polymorphisms of IRF6 showed association to the NSCLP phenotype, rs861019 in exon 2 and rs7552506 in intron 3. A total of seven polymorphic SNPs were found in this study (SNP 1–6 and SNP 8), but for SNP 7, rs2235371, all individuals were homozygous for the most common C allele. Zucchero et al. (2004) found a higher risk of recurrence of NSCL/P for carriers of the C allele (Val274) of rs2235371 and suggested that this SNP accounts for 12 per cent of the genetic contribution to CL/P and gives a threefold increased risk of recurrence in families who already have one affected child. As rs2235371 (SNP 7) was seen as monomorphic (C allele only) in our study, it is possible that the C-allele contributes to the NSCL/P phenotype. High LD between SNPs 5 and 6 in our study (rs2013162 and rs2235375) confirms previously published data (Fig. 15), (Blanton et al., 2005; Scapoli et al., 2005). In a genome-wide linkage study followed by a candidate gene association study and fine mapping, Marazita et al. (2009) found that SNP rs2013162 of IRF6 reached significant association with the CL±P phenotype (rs2013162 lies between rs642961 and rs2235371; which is SNP 5 in our study).

Many studies have concluded that the IRF6 gene is involved in the development of NSCL/P in a complex modifying/polygenic manner rather than as a monogenic/major disease locus (Houdayer et al., 2001; Zucchero et al., 2004; Blanton et al., 2005; Blanton et al., 2010; Ghassibe et al., 2005; Scapoli et al., 2005; Srchomthong et al., 2005; Park et al., 2007; Jugessur et al., 2008; Diercks et al., 2009). In a study of 357 candidate genes in two different NSCL/P cohorts from Norway and Denmark, Jugessur et al. (2008) chose genes known from earlier studies or syndromes to be related to orofacial clefts. The IRF6, the alcohol dehydrogenase 1C (ADH1C), and the fibroblast growth factor 12 (FGF12) genes were strongly associated with NSCL/P in both populations. They also found a difference between the two cohorts: a number of genes were significantly associated in one population but not in the other. FOXE1 was found to be significantly associated with NSCL±P in their Danish sample, but not in their larger Norwegian sample. On the other hand, MSXI was identified among the genes associated with NSCL±P in their Norwegian sample, whereas no such association was seen in their Danish sample. The population-specific susceptibility found by Jugessur et al. (2009) highlights the problem with lack of replication across diverse populations.

A recent study of the IRF6 gene in NSCL/P families highlighted rs642961 as associated with NSCL±P (Rahimov et al., 2008). The study showed an over-transmission of the A
allele of SNP rs642961, located in the promoter of the gene, to affected individuals, and also that the A-C haplotype of rs642961 and rs2235371 was associated with the CL and CLP phenotypes but not with CP. We investigated whether this also held in our Swedish and Finnish sample sets, and whether this promoter SNP of IRF6 could be a modifier in our VWS/PPS families (Study IV). Our tests of haplotype transmission in the family and trio sample sets consisting of 95 families gave results that differed from those of Rahimov et al. (2008). None of the findings in our total NSCL/P sample set were significant. However, in our sub-sample set of Swedish CP families we did find that the A-C haplotype for rs642961 and rs2235371 was significantly protective, while G-C was a haplotype risk. Taqman genotyping was our method of choice in this study because it can be performed on a few specific SNPs (see Fig. 12; Livak et al., 1995). Our lack of replication could be a consequence of our smaller sample size or of an ethnic/population difference; as Taqman is rather robust and reliable, the explanation is unlikely to be a lack of sensitivity of our genotyping method. Lately, others such as Blanton et al. (2010) have reported an ethnic difference (analysing these haplotypes of IRF6) in their Hispanic and non-Hispanic white (NHW) NCLP sample sets. When the two SNPs were considered as a haplotype (showing the SNPs from the positive strand, as in our study), the A-C and the G-C haplotypes were equally over-transmitted in the NHW group. In their Hispanic group, the A-C haplotype only was over-transmitted. This suggests that in their NHW material, it is not rs642961 that shows association with the disease but rs2235371. In a Honduran population, Larrabee et al. (2011) found the same association (to rs2235371) but no association between the rs642961 and NSCL±P, either in their case-control analysis or in the family-based association testing.

Complex traits are difficult to map since several genes are often involved and different combinations of the disease genes may result in similar phenotypes. Moreover, when a specific marker is shown to be connected to the disease, it is not always the exact risk SNP that has been found but instead an SNP in LD with the risk allele, as a haplotype.

From the point of view of embryological development it is clear that from such a complex and numerous series of events such as cell migration, apoptosis, merging and fusion between the developing processes, it is expected that many genes are involved.
To conclude, VWS/PPS can be a clinically under-diagnosed phenotype because of the variable expressivity and can also sometimes be classified, misleadingly, as NSCL/P. NSCL/P is a complex anomaly, where disturbed dental development is a frequent finding. NSCL/P is not single-gene dependent, unlike the Mendelian inherited VWS. Dividing our material into sub-phenotypes gave rather small groups but we do show a significant risk with a haplotype of IRF6 in the NSCP group, as well as an association for two SNPs of the IRF6 gene with NSCLP. Our results on NSCL/P emphasize the need for additional evaluation of the IRF6 gene and other genes/modifiers, to further clarify their roles in the development of the NSCL/P phenotype.

6 FUTURE PERSPECTIVES

The field of genetic research has changed rapidly during my period as a research student, especially through the enormous advances in the area of genome-wide association studies. Since NSCL/P is thought to be a multi-factorial disease, several genes can be involved in the development of this anomaly. Other genes, epigenetic and/or environmental factors causing a cleft, investigated in other studies, lie outside the scope of this project. In one of our ongoing projects, a genome-wide approach that includes seven NSCL/P families (20 affected individuals) with Affymetrix10K resolution (10 000 SNPs), as well as 50 affected individuals analysed with the Affymetrix 6.0 resolution (1 million SNPs) has been performed. Our results have not attained significance in any specific region and the heterogeneity of the affected individuals has made analysis statistically difficult. The project is continuing with a new approach, that is, whole exome sequencing and we hope to complete some of the new analyses in the spring of 2012.

Our goal is still to identify the genes involved in the development of NSCL/P in order to contribute knowledge for more accurate risk counselling and prevention programmes for families at risk. When markers for the disease have been discovered, the risk can easily be analysed from a blood sample. This is already feasible with the monogenic VWS anomaly, as we have shown in Studies II and IV. The ultimate goal of all ongoing research in the field is to provide, for each individual patient, more knowledge for the development of preventive therapies for this severe congenital anomaly.
7 POPULÄRVETENSKAPLIG SAMMANFATTNING

Min forskning rör barn som är födda med läpp-, käk- och/ eller gomspalt (LK/G). I mitt kliniska arbete behandlar jag dessa barn med tandreglering och käkortopedi i Stockholms Kraniofaciala team.


Patienter med popliteal pterygium syndrom (PPS) har samma symtom som patienter med VWS, men med allmän påverkan på kroppen såsom sammanväxningar mellan tårna (syndaktylier) och extra hud på benen. Vid såväl VWS som PPS förekommer mutationer på IRF6 genen.

Non-syndrom (NS) läpp-, käk- med eller utan gomspalt utan tillhörande missbildningar i andra organ, uppträder hos cirka 70 procent av patienterna. Arvsmönstret för läpp-, käk- med eller utan gomspalt är komplext där flera gener troligen samverkar och även miljöfaktorer inverkar. Förändringar på IRF6 genen har föreslagits vara en del av orsaken till uppkomsten av läpp-, käk- med eller utan gomspalt.

Målet med avhandlingen var att studera familjer och individer födda med non-syndrom LK/G och syndromen VWS och PPS; att registrera symtomen för sjukdomarna inklusive den dentala påverkan, att upptäcka mutationer på IRF6 genen i de syndroma fallna av VWS/ PPS och att undersöka om genen är orsaken till symtomen även i NSLK/G.

Studierna inkluderade finska och svenska LK/G familjer samt familjer med VWS/ PPS från andra länder.

I studie I undersöktes den dentala utvecklingen på 129 barn födda med NS läpp-, käk- med eller utan gomspalt från 5 till 10 år. Många av patienterna hade en avvikande tandutveckling med rotationer och atypisk tandform samt övertal eller undertal. Den laterala framtanden (incisiven) var ofta missformad vid diagnosen läppspalt medan laterala incisiven ofta saknades vid diagnosen LKG.
När den permanenta tanden i spaltområdet växer (erupterar) behöver ben tillföras till spaltområdet, vilket görs med en operation, så kallad bentransplantation, där kroppseget ben används. Tidpunkten för bentransplantationen beror på om den laterala incisiven finns. Vi kontrollerade om det kunde förutses om den laterala incisiven fanns och på vilken sida om spalten den erupterade. Inget signifikant mönster (p=0.15) kunde ses vid växlingen från primära till permanenta laterala incisiven i spaltområdet, även om de flesta av de befintliga laterala incisiverna erupterade distalt om spalten.

I **studie II** studerades 17 familjer med kliniska symtom på VWS/PPS. *IRF6* mutation återfanns i 59% av familjerna.

Vi fann att det förelåg en nymutation ("de novo"), ett sjukt barn med friska föräldrar, i ca 31% av familjerna (VWS/PPS familjerna från studie II och studie IV).

VWS/PPS familjer kan ha olika spaltdiagnoser (gomsplatt och läpp-, käk- med eller utan gomsplatt) inom samma familj, detta fann vi i 37.5% av VWS/PPS familjerna (från studie II och IV) men inom familjer med non syndrom LK/G förekom det i endast i 6% av familjerna (studie III och IV). Följaktligen är en blandad spalttyp inom familjen ett kännetecken för VWS/PPS. Vid enbart spalt (inga läppgropar) kan det kliniskt vara svårt att särskilja VWS/PPS från non-syndrom LK/G. Följaktligen om det inom familjen finns en blandad spalttyp kan man screena *IRF6* genen för mutationer, för att särskilja risken att överföra anlaget för VWS/PPS till sina barn gentemot den mindre risken vid NSLK/G.

Vi fann även att individer med VWS/PPS med en mutation på *IRF6* ofta hade den studerade riskmarkören (rs642961; studie IV) på *IRF6* genen. När riskmarkören överfördes till ett drabbat barn, överfördes den oftast från samma förälder (samma kromosom) som *IRF6* mutationen. Eventuellt kan detta påverka effekten av *IRF6* mutationen. Ytterligare studier behövs för att undersöka detta.

Då symtomen hos patienter med NSLK/G och VWS är likartade, samt att tidigare studier visat att *IRF6* genen är närvarande vid normal gomslutning, ville vi undersöka *IRF6* genen även på familjer med NSLK/G.

I **studie III** undersöktes *IRF6* genen hos 17 familjer med NSLK/G. Vi fann två markörer på *IRF6* genen som visade ett samband med spalt, men inga proteinbildande mutationer hittades. Markörerna (rs861019 och rs7552506) kan påverka risken för att utveckla spalt.
I studie IV undersökte 95 familjer med NSLK/G. Vi undersökte två markörer på IRF6 genen (rs642961 och rs2235371) som i tidigare studier visat sig ha ett samband med NS läpp-, käk- med eller utan gomspalt. Vi fann inget samband i våra familjer mellan markörerna och NS läpp-, käk- med eller utan gomspalt. Vi fann däremot i den grupp av vårt svenska familjematerial som har NS isolerade gomspalter, en signifikant risk vid en kombination av markörerna rs642961-rs2235371 (G-C). Detta var i motsats till vad tidigare studier funnit. Det kan eventuellt bero på en etnisk skillnad (som andra studier från olika länder också påvisat i år) eller på att antalet undersökta individer behöver vara större.


Det övergripande målet för oss och för all pågående forskning inom området, är att skapa mer kunskap för att kunna utveckla bra preventiva åtgärder och behandlingar för patienter med medfödd LK/G.
8 ACKNOWLEDGEMENTS

This thesis would not have been possible without the help and commitment of many persons. In particular I want to thank:

My principal supervisor and co-author of all the articles, Agneta L-A Karsten, for support and encouragement and for persuading me to start with and continue this research project.

My co-supervisor and co-author of three of the articles, Myriam Peyrard-Janvid, who has taught me all I know about genetics and never been bothered about my many questions. Thank you for all your advice.

All the other co-authors: Catharina Larsson for co-authoring and editing the first article, Ingergerd Fransson for teaching me laboratory work, Marco Zuchelli for statistical calculations and answering my questions, Nadia Alqadi for skilful measurements of the models, Ola Larsson for encouragement, Juha Kere for believing in this project, Hannele Koillinen, Jyri Hukki and Jorma Rautio for collaboration with the Finnish sample-set.

To all my research colleagues at the JKE group at Novum for good advice and support. Special thanks to: Cilla (for reading and commenting on this thesis), Virpi (we enjoyed the Ark at Globen), Tiina, Isabel, Kristiina (good luck with the baby), Hong, Per, Helena, Mauro, Shintaro and Lovisa. Thanks also to the people at MAF and B.E.A, especially those of you who helped with the analyses: Kristina, Astrid, Päivi and Marika. To the former PhD students: Hanna, Sara, Christina, Heidi, Erik and Anna. The IT section has helped me a lot when I had to exchange my computer: thank you Anders, Johan and Erik.

Elisabeth Berg, at Lime, for statistical help and Carl-Magnus Forsberg, for good advice.

My colleagues at the Craniofacial Team in Stockholm, Solna, especially Anita Fridell and Katrin Stabel-Svensson.

My colleagues at the Department of Dental Medicine, Division of Orthodontics, Huddinge especially Jan Huggare. Thank you for your support.

My colleagues at Folktandvården Sörmland, especially at Tandregleringen Eskilstuna for taking care of my orthodontic patients while I was writing this thesis and also for encouraging me on the research pathway. The former and present manager at Tandregleringen Eskilstuna, Margareta Sörensen and Torsten Frost, and also FoU
Folktandvården Sörmland, especially Lars Hjalmarsson, for giving me time to write my thesis.

My fantastic family: my supportive husband, Claes, and my children, Fredrik and Amanda, for your love and patience.

This work was supported by the Swedish Research Council ("Vetenskapsrådet" 2007-255-47325-41), the Swedish Dental Society, Sigrid Juselius Foundation, Karolinska Institutet, Stockholm County Council (K2002), Kvinnliga tandläkarklubben and FoU “Folktandvården Sörmland”, Sweden.
9 REFERENCES


Diercks GR, Karnezis TT, Kent DT, Flores C, Su GH, Lee JH, Haddad J, Jr. (2009) The association between interferon regulatory factor 6 (IRF6) and
nonsyndromic cleft lip with or without cleft palate in a Honduran population. *Laryngoscope* 119: 1759-1764


Falconer (1965) The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet* 29: 51-71


clefts: single-marker and haplotype-based analyses in a population-based case-control study of facial clefts in Norway. *Genet Epidemiol* 32: 413-424


Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip. *Nat Genet* 40: 1341-1347


