Genetic aspects of Cleft Lip and Palate

PhD thesis

Linda P. Jakobsen

Copenhagen

2006

Cover: Wendy Plovmand

The cover picture is a graphic modification of a scanning electron micrograph kindly provided by Dr Kathleen K. Sulik (from http://www.med.unc.edu/embryo_images/) of the craniofacial region in a mouse embryo corresponding to the developmental stage in the human embryo during the 5th week of gestation.

Following the defence, the thesis will be available on the website for "Landsforeningen Læbe-Ganespalte" (the Danish Support Group for Patients with Cleft, Lip and Palate), http://www.llg.dk/.

Table of contents

Acknowledgements

This thesis is based on work performed at Wilhelm Johannsen Centre for Functional Genome Research (WJC), Institute of Medical Biochemistry and Genetics (IMBG), Panum Institute, Copenhagen and the Department of Plastic and Reconstructive Surgery and Burns, Rigshospitalet, Copenhagen in the period 2003–2006.

I am privileged to have had splendid and skilful colleagues and supervisors, helping, guiding and motivating me all the way through the current research. It has been an exciting journey, which hopefully will continue. In particular, I would like to thank my supervisors, Niels Tommerup at the WJC, IMBG and Mary Knudsen at the Department of Plastic and Reconstructive Surgery and Burns, Rigshospitalet and my colleagues at the WJC, IMBG, the Department of Plastic and Reconstructive Surgery and Burns, Rigshospitalet, and the Copenhagen Cleft Lip and Palate Centre, who have taught me a lot as we discuss topics related to cleft lip and palate on a daily basis: Hans Eiberg, Lars Hansen, Klaus Kjær, Lars Allan Larsen, Claus Hansen, Zeynep Tümer, Sakari Kauppinen, Mikael Andersen and Kirsten Mølsted.
Also, I would like to thank my colleagues at WJC, with whom I have had the daily contact, as it would not have been so much fun without them: Erling Lauridsen, Evete Mawlad, Rikke Møller, Mette Gilling, Asli Silahtaroglu, Iben Bache, Elisabete Engenheiro, Kirsten Svenstrup, Suzanne Lindquist, Janni Vestergaard, Mads Bak, Mette Christensen and many more.
I have received skilful help in part of the laboratory work from Karen Henriksen, Anita Niebuhr, Minna Becker, Ingrid Kjær, Elisabeth Larsen, Lillian Rasmussen and Anne Mette Mikkelsen (WJC and IMBG), and Karen Girke (WJC) and Vivi Lisfeldt, Department of Plastic and Reconstructive Surgery and Burns, Rigshospitalet, have been helpful finding patient addresses and medical records.
Moreover, I would like to thank my other colleagues who made this work possible: Rehannah Borup, Lisa Leth Maroun, Steen B. Christensen, Karl Erik Jensen at Rigshospitalet, Reinhard Ullmann at the Max Planck Institute for Molecular Genetics, Berlin, Kasper Lage at the Technical University of Denmark and Kjeld Møllgård at the Institute of Medical Anatomy, Panum Institute, Copenhagen.
The work was supported by grants from the Augustinus Foundation, the Vanfoere Foundation, Doctor Johan Boserup and Lise Boserup's Foundation, Aase and Ejnar Danielsen's Foundation, Jacob Madsen and wife Olga Madsen's Foundation, the Research Foundation of the Queen Louise Children's Hospital, WJC (which was established by the Danish National Research Foundation), and Rigshospitalet.

Most importantly, I am privileged to have a wonderful husband and family who support me. They are the centre of my life and I am grateful to be exploring life together with them.

Linda P. Jakobsen,
Copenhagen, November 2006

The Ph.D. thesis is based on the following four papers:

1) *The genetic basis of the Pierre Robin Sequence* (Paper I, published:
Jakobsen LP, Knudsen MA, Lespinasse J, Garcia Ayuso C, Ramos C, Fryns JP, Bugge M, Tommerup N. The genetic basis of the Pierre Robin Sequence. Cleft Palate Craniofac J. 2006;43(2):155-159.

2) *Pierre Robin Sequence may be caused by SOX9 dysregulation* (Paper II, submitted following revision to the Journal of Medical Genetics).

3) *Expression analyses in cleft lip and palate tissue point to a possible role for the immune system in cleft lip and palate* (Paper III, in preparation. Will be submitted to the Journal of Medical Genetics).

4) *Suggestive linkage to a neighboring region of IRF6 in a CL/P multiplex family*
(Paper IV, in preparation. Will be submitted to the American Journal of Medical Genetics).

Abbreviations

CGH, Comparative Genome Hybridisation

CL/P, cleft lip and/or palate

CL(P), cleft lip with or without cleft palate

CL, isolated cleft lip

CP, isolated cleft palate

DNA, deoxyribonucleic acid

FISH, Flourescence In Situ Hybridisation

HCDB, Human Cytogenetics Database

IMBG, Institute of Medical Biochemistry and Genetics

LD, linkage disequilibrium

LOD, logarithm ($\log_{10}$) of the odds

Mb, megabase (10^6)

MCNdb, Mendelian Cytogenetics Network Database

MGI, Mouse Genome Informatics

miRNA, microRNA

NSCLP, non-syndromic CL/P

OMIM, Online Mendelian Inheritance in Man

PCR, polymerase chain reaction

PRS, Pierre Robin Sequence

QPCR, quantitative PCR

SNP, single nucleotide polymorphism

TDT, transmission disequilibrium test

WJC, Wilhelm Johannsen Centre for Functional Genome Research

Electronic-Database Information

These are the URLs referred to in the thesis (the URLs referred to in the manuscripts are listed in the manuscripts):

Craniofacial and Oral Gene Expression Network (COGENE), http://hg.wustl.edu/cogene/

EUROCRAN, http://www.eurocran.org/

Mouse Genome Informatics (MGI), The Jackson Laboratory, http://www.informatics.jax.org/

Mendelian Cytogenetics Network Database (MCNdb), http://www.mcndb.org/

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/omim/

Pubmed, http://www.ncbi.nlm.nih.gov/entrez/query.fcgi/

UCSC Genome Bioinformatics site, March 2006 assembly, http://www.genome.ucsc.edu/

Preface and aims of the study

Cleft lip and/or palate (CL/P) is a common congenital malformation which affects approximately 2 per 1000 newborns worldwide and places a huge burden on the families affected (Murray, 2002; Jugessur and Murray, 2005). Although a certain amount of insight into this craniofacial malformation has been achieved, we are still far from having a complete understanding and even further from being able to transform our knowledge into useful clinical measures, such as diagnosing and advising families. The foremost purpose of CL/P research is to gain insight into the underlying processes leading to CL/P. This knowledge may enable us to provide a form of intervention that may decrease the risk of a fetus developing CL/P or decrease the extent of the malformation.
As will be described in "***Introduction***", several approaches serve as eligible means to elucidate aspects of the aetiology of CL/P, and a selection of these approaches were used in the work presented here.

The specific aims in this Ph.D. thesis were:

1) **To investigate the aetiology of the Pierre Robin Sequence** (PRS), a syndromic form of cleft palate, consisting of cleft palate, micrognathia, and early neonatal respiratory difficulties due to glossoptosis. Although it is well known that PRS is a heterogeneous group and has a partly genetic aetiology, because it forms part of many Mendelian syndromes, the aetiology of isolated PRS is largely unknown.
2) **To study gene expression in lip and palate tissue from patients with different types of CL/P.** Epidemiological data point to the existence of different aetiologies in the three subgroups, cleft palate, cleft lip and palate, and cleft lip. If these differences are reflected in tissue expression profiles, expression analysis may convey useful information on differences in the molecular processes between the CL/P subgroups. Expression studies in lip and palate tissue from patients with CL/P have not been carried out before.
3) **To identify non-syndromic CL/P multiplex families (two or more affected family members) and perform genetic marker analysis, in order to find new CL/P loci or confirm/reject already known loci.** Since the aetiology of CL/P is complex, it may be useful to focus on either CL/P subgroups (such as PRS mentioned above), or large CL/P multiplex families where a single gene or a few genes play a major role. Moreover, the recent refinements of the genetic markers and array-based methods enable detailed genomewide scans. A large cohort of Danish CL/P multiplex families was collected in the 1980s, and a follow up on these families identified one large family eligible for a genomewide scan.

Introduction

This section provides an overview of the various study types applied in the field of genetic research in CL/P. This is not meant to be an exhaustive overview, since many new methods, especially technical facilities and bioinformatic tools, are constantly emerging.
We will start in Denmark first, because one could say that cleft lip and palate research was "born" here!

Cleft Lip and Palate in a short Danish historical perspective

The causes of CL/P have been investigated extensively for many years. In Denmark particularly, we have a long tradition of thorough and comprehensive CL/P registration, treatment and research. Dr Poul Fogh-Andersen initiated this tradition by defending his thesis "Inheritance of Harelip and Cleft Palate" in 1942. His registration of CL/P cases provided evidence for a genetic basis of CL/P, and he also showed that isolated cleft palate (CP) and cleft lip with or without cleft palate CL(P) are distinct aetiological subgroups, since they do not co-segregate in affected families. Professor Kaare Christensen continued this tradition by contributing with many important high-quality studies to the field. He initiated the Danish Cleft Lip and Palate Registry, which includes 99.9% of all people born with a cleft in Denmark since the 1950s (Christensen, 1999). Using this registry and data from the Danish Twin Registry, he provided evidence that genetic factors are the major contributors to the aetiology of CL/P (Christensen and Fogh-Andersen, 1993; Christensen and Mitchell, 1996), a view shared by most experts in the field of CL/P. Dr Christensen has also made a substantial contribution in the fields of environmental factors and gene-environment interactions in CL/P, and he added the issue of prognosis by providing evidence that CL/P patients have a significantly increased lifetime mortality from several different types of diseases (Christensen et al., 2004). Recently, Dr Camilla Bille provided important information on the risk of cancer in people with CL/P. The risk is not increased overall, but some cancer forms may have increased occurrence in patients with CL/P (Bille et al., 2005).
Last but not least, Kirsten Mølsted and Professor Sven Kreiborg have contributed tremendously with many important cephalometric studies on craniofacial growth in CL/P patients.

Aetiology; Genes versus Environment

Although genetic factors are the major contributors to the aetiology of CL/P, the environment plays a role too. This was pointed out by Warkany (1943), who found a higher incidence of congenital malformations, including cleft palate, in the offspring of female rats deprived of riboflavin. A lot of research into environmental factors and their influence on CL/P has been carried out, and to date no single environmental factor has been identified as posing a major risk for CL/P. However, alcohol intake, anti-epileptic medication, and smoking during pregnancy have all been shown to increase the risk of CL/P and other malformations in humans (Shaw and Lammer, 1999; Little et al., 2004; Artama et al., 2005). Only weak or ambiguous results have been found for folic acid and other nutrients, maternal disease and stress during pregnancy, chemical exposures and corticosteroids (Hayes in Cleft Lip and Palate: from Origin to Treatment, 2002).

Evidence of that genetic factors play a major role in the aetiology of CL/P comes from epidemiological observations. In larger cohorts of people with CL/P, approximately 20% have other relatives with CL/P, and an increased prevalence is observed among first and second degree relatives (Christensen and Mitchell, 1996). Moreover, the CL/P concordance rate in monozygotic twins (60%) is considerable higher than the CL/P concordance rate in dizygotic twins and siblings (5-10%) (Farrall and Holder, 1992; Christensen and Fogh-Andersen, 1993; Christensen and Mitchell, 1996). Syndromic CL/P cases also indicate a genetic aetiology, because more than 400 known syndromes include orofacial clefting, and many of these follow classic Mendelian inheritance patterns.

It is evident that CL/P does not have a simple monogenic basis because the segregation patterns do not fit the classical Mendelian inheritance patterns. Segregation analyses of CL/P point to polygenic or multifactorial inheritance, with each locus only providing a minor contribution to the risk (Farrall and Holder, 1992; Mitchell and Christensen, 1996). From the segregation of CL/P in multiplex families (families with two or more members with CL/P), it has been estimated that the most likely number of involved loci is between 2 and 14 (Schliekelman and Slatkin, 2002). Obviously, the polygenic inheritance pattern complicates the unravelling of the CL/P aetiology, and the fact that gene-environment interactions and maternal genotypes may play a role as well (Gaspar et al., 2004; Lammer et al., 2005; Shaw et al., 2005) complicate the dissection of this complex disease even further.

But genetic research in complex diseases, such as CL/P, has experienced some successes recently, and the hope that unravelling this frequent and partly disabling malformation may in fact offer improvement to patients with CL/P in the future, spurs the continuing investigation of the genetic aetiology in CL/P.

Cleft Lip and Palate Genetics

To study the complex disease of CL/P, it is necessary to use a variety of different approaches. Animal studies, linkage, association and cytogenetic studies are widely used approaches to identifying the genes involved in the development of CL/P. The study of syndromic CL/P forms has recently also had a major impact in this field. No particular approach is superior to any other; each has its strengths and weaknesses. The genes known to be involved in CL/P, or associated with CL/P are listed in the ***Appendix***.

The information in the ***Appendix*** are from various types of studies (animal studies, linkage and association studies, CL/P syndromic forms, and cytogenetic studies), and provide an extensive, but not exhaustive overview of CL/P candidate genes, as new data is generated almost on a weekly basis. The data on genetically altered mouse models may be insufficient, as many mouse mutants displaying orofacial clefts are generated in laboratories all over the world and data may not always reach publicly available databases like the Mouse Genome Informatics (MGI) from the Jackson Laboratory. An overview of the various approaches used to gain insight into the aetiology of CL/P is presented below.

Animal studies

The mouse model provides an excellent opportunity to study CL/P. It is the preferred mammalian animal model, because it has a short reproduction cycle and a known genomic sequence, very close to the human genomic sequence (Thyagarajan et al., 2003). Many genetically manipulated mouse models (e.g., knock-out, knock-in, gene-trapped, spontaneous mutations and chemically induced mutations) display CL/P, although cleft palate (CP) is the cleft type most often encountered. An overview of the mouse models showing features of CL/P is given in the ***Appendix***.

Many CL/P mouse models have counterparts in syndromic CL/P forms found in humans. For instance, mutations in the gene *COL11A1* cause Marshall syndrome (MIM#154780) or Stickler syndrome type II (MIM#604841) in humans, and a similar phenotype is found in mice, supporting the use of mouse models in CL/P research. On the other hand, known CL/P genes in humans do not always cause CL/P in mice. Heterozygous mutations in *IRF6,* known to cause a syndromic form of CL/P in humans (Kondo et al., 2002), do not result in CL/P in mice (Ingraham et al., 2006), perhaps because of functional redundancy with other genes (Thyagarajan et al., 2003).

The list of genes involved in CL/P in mice is long, and will probably continue to increase. While it is questionable whether all the information on mouse models can be transferred to humans, there is no current evidence that the aetiology of CL/P in humans is less complicated than in mice.

Linkage studies and linkage disequilibrium tests

In the initial investigation of a disease with a genetic background, the disease-causing genes are unknown. In these circumstances, linkage studies offer a method of identifying the locus harbouring the disease-causing gene. Several genes involved in the classical single gene disorders, such as Huntington disease (MIM#143100) and cystic fibrosis (MIM#219700), have been identified through linkage studies in large

multiplex families. When a specific marker allele (or a set of marker alleles, i.e., haplotypes) is transmitted together with the disease of interest in a pedigree, the markers are linked to the disease gene. A statistical calculation, the LOD score (logarithm ($\log_{10}$) of the odds for linkage versus no linkage), is used as a measurement of linkage, which is normally considered significant when the LOD score is >3, (i.e., the odds of linkage vs. no linkage is 1000:1). Marker alleles are DNA polymorphisms (most often microsatellites or single nucleotide polymorphisms (SNPs)) which are scattered throughout our genome. Due to polymorphism, the markers/haplotypes found in a family are specific for this individual family. The marker/haplotypes do not represent the actual disease gene, but may be linked to the disease-causing gene. From each generation to the next, there is a chance of a crossing over of the homologue chromosomes (during the meiosis), causing recombination events between the marker and disease-locus. The more recombination events occurred, the more narrow will the detected linkage interval become. So linkage studies require large multiplex families, preferably with one disease locus, since significant linkage will only be detected in the case of one, or at the most a few, major disease loci. These criteria are extremely difficult to meet in complex diseases, due to the lack of large multiplex families and the involvement of several loci in the same disease (locus heterogeneity). Even when the criteria have been met, the disease locus identified usually extends over many megabases (Mbs) and may contain hundreds of genes. Consequently, a major sequencing effort is usually needed to identify the disease-causing mutation, as in the presented linkage study (Paper IV). Moreover, an assumed inheritance pattern has to be applied in the linkage calculations (parametric tests), and this is difficult, as complex diseases do not apply to any of the classical inheritance patterns.

These factors contribute to the lack of significant LOD scores in linkage studies of complex diseases. Consequently, LOD scores of 2 or more are generally accepted as *suggestive* of linkage (Lander and Kruglyak, 1995; Altmüller et al., 2001). Altmüller and colleagues (2001) suggested increasing the sample homogeneity in genome scans in order to overcome the problem of non-significant linkage results. This approach was tried in Paper IV, by studying a single CL/P multiplex family at the expense of the sample size. The genome scan yielded a LOD score of 2.73, *suggestive* of linkage, and was validated with microsatellite markers to a 6.5 Mb interval neighbouring, but *not* including, the gene *IRF6* on 1q32. Whether this interval in fact harbours genes involved in CL/P separately, or genes/non-coding regions regulating *IRF6*, is presently unknown, but sequencing of four potential candidate genes (*SOX13, FMOD, OPTC* and *IKBKE*) and functional RNAs (miRNAs) in the interval did not reveal any mutations.

A review of the literature on CL/P linkage studies produced multiple loci of suggestive linkage, but only very few loci with significant LODs: a meta-analysis of 13 genomewide linkage studies (Marazita et al., 2004) and a recent linkage study in two large Indian families (Radhakrishna et al., 2006) (see ***Appendix*** for CL/P loci identified in linkage studies). As genomewide searches increase the chances of finding spurious positive results, the threshold for significant LODs is increased in proportion to the number of markers used, and thus a LOD score of ~3.3 is usually considered as the threshold for linkage in genome scans

(Lander and Kruglyak, 1995). In the meta-analysis, seven genome scans were summarised because they were analysed with the same markers, and showed a significant linkage (summed LODs ≥ 3.2) to six loci on five chromosomes (1p12-13, 6p23, 6q23-25, 9q21, 14q21-24, and 15q15). Furthermore, according to the meta-analysis calculations of all the 13 genome-wide scans, 10 additional loci showed genome-wide significant linkage (1q32, 2q32-35, 3p25, 7p12, 8p21, 8q23, 12p11, 17q21, 18q21 and 20q13). The study in the two Indian families identified linkage to 13q33.1-34 (LOD score of 4.45). These CL/P loci are highly important and will help guide the future search for CL/P candidate genes. One way of applying these results could be to identify genes located in these linkage intervals, and in which mutations cause CL/P in mice or syndromic CL/P in humans. Candidate genes chosen from the ***Appendix*** based on the above criteria would then be: *COL11A1* (1p12-13); *COL11A2, EDN1, TFAP2A* and *FOXF2* (6p23); *GJA1, TCF21* and *PEX7* (6q23-25); *ROR2, PTCH* and *FOXE1* (9q21); *BMP4, TGFB3* and *POMT2* (14q21-24). No genes at 15q15 are presently known to be involved in CL/P in mouse models or CL/P syndromes. Some of these genes have already been shown to contribute to the CL/P aetiology (Table 1).

A related, but distinct type of linkage study is the linkage disequilibrium test (LD). The LD study is a family-based association test, and a hybrid between an association and a linkage study. It is a non-parametric test, as the assumed inheritance pattern does not have to be specified. An LD study often used is the transmission disequilibrium test (TDT), which determines whether a given allele (or marker) from a heterozygous parent is transmitted to an affected child more often (or seldom) than the random 50%. When transmission deviates from the 50% (and the allele and the disease are linked), LD is present and the allele may be disease-causing or in close linkage with a disease-causing mutation. The advantage of LD is the opportunity of studying a large number of smaller affected families (e.g., triads) together, obtaining more recombination events. Consequently, the TDT may detect linkage disequilibrium in a smaller interval than the classical linkage studies. The TDT is relatively powerful and also has the advantage of not being hampered by population stratification, as the case-control association test may be (see *Association studies*), because the parents serve as controls and they share 50% of the genetic background with their child (Cardon and Palmer, 2003).

Association studies

Association studies determine whether alleles occur together with a specific phenotype more often (or seldom) than in a control group. A specific allele has to be suspected as disease-causing (or protective) and therefore tested in either a case-control study or a family-based association design (see *Linkage studies and linkage disequilibrium tests*). Clues as to candidate genes typically originate from basic biological studies such as animal models. Association studies provide a powerful tool for identifying alleles of minor importance, but may also involve a risk of finding false associations. A false association may be detected if the tested variant is not the actual disease-causing variant, but closely located to the disease-causing variant. False associations may also occur if the case and control groups show differences in the studied

allele frequencies, not causally related to the studied disease, i.e., population stratification (Cardon and Palmer, 2003). Because of the risk of population stratification, family-based association designs (such as the TDT) have become increasingly popular. Case-control association studies have provided evidence that variants in the genes *MTHFR, ARNT, TGFA, GAD1, MSX1, RARA, TGFB3* and anonymous markers in the region 4q31 are associated with the CL/P phenotype (***Appendix***).

Syndromic CL/P

Many syndromes with phenotypes that include CL/P are now known. A search in OMIM (as of October 2006) revealed 467 and 258 hits searching for "cleft palate" and "cleft lip" respectively. In some of the CL/P syndromes, the genes involved have been identified (***Appendix***, CL/P syndromes) and the list is constantly increasing, primarily due to improved sequencing facilities (Vieira et al., 2005). The dissection of CL/P syndromic forms has revolutionized the field of CL/P genetics, by demonstrating that genes involved in syndromic forms of CL/P may also play a role in the group of non-syndromic CL/P (NSCLP) patients (Stanier and Moore, 2004). The classical example of this is the interferon regulatory factor 6 (*IRF6*), which was identified as the causative gene in Van der Woude syndrome (VWS, MIM#119300) and popliteal pterygium syndrome (PPS, MIM# 119500) (Kondo et al., 2002). Subsequent screening of a large cohort of patients with NSCLP showed that common polymorphisms in *IRF6* contribute to about 12% of the genetic aetiology in NSCLP (Zucchero et al., 2004). Likewise, similar stories can be told for several other genes: *PVRL1* (CLPED1, MIM#225060) (Suzuki et al., 2000; Sozen et al., 2001; Avila et al., 2006), *MSX1* (Witkop syndrome, MIM#189500 and #608874) (van den Boogaard et al., 2000; Jezewski et al., 2003), *TP73L* (ADULT syndrome, MIM#103285; AEC, MIM#106260; LMS, MIM#603543; EEC3, MIM#604292; SHFM4, MIM#605289) (Celli et al., 1999; Leoyklang et al., 2006) and *TBX22* (CPX, MIM#303400) (Braybrook et al., 2001; Marcano et al., 2004).

The fact that syndromic CL/P and NSCLP may share some of the genetic aetiology is intriguing; one could hypothesise that the dissection of all the CL/P syndromes would lead to complete understanding of the complex genetic aetiology in NSCLP. However, it is probably not as simple as that, because not only the coding regions, but also the non-coding regions in the genome (which make up about 98% of the total genome) may contribute to NSCLP. As suggested in Paper II, dysregulation of the transcription factor *SOX9* (and perhaps the gene *KCNJ2*) may be involved in the aetiology in at least a subgroup of the patients with Pierre Robin Sequence (PRS) (cleft palate, micrognathia and early neonatal respiratory difficulties). It is probably not mutations in the *SOX9* coding region which cause isolated PRS, but rather genomic alterations in the non-coding regulatory regions (mutations, polymorphisms or chromatin conformation) surrounding *SOX9*, or other genes acting in signalling pathways with *SOX9*. It is a major current challenge to gain insight into the role of these non-coding and potentially regulatory regions, to understand the regulation of genes. Many computer-based prediction programs have been launched, most on the assumption that the non-coding regulatory regions have conserved sequences or structure (Woolfe et al.,

2005; Mattick and Makunin, 2006). But despite these challenges (relating to the non-coding and potentially regulatory regions), the syndromic approach appears attractive.

Table 1 shows the genes involved in the aetiology of NSCLP. Mutations in the coding regions or association of specific polymorphisms in and around these genes have been identified in patients with CL/P. The table may be viewed as a preliminary list of genes relevant to screen for in NSCLP, in order to understand the genetic basis of the cleft in each individual person. To be able to make this knowledge clinical useful, we need to know a lot more.

Recent technical progress has made it possible to perform high-throughput sequencing of candidate genes in large patient cohorts. Not only genes identified in syndromic CL/P, but also candidates selected on the basis of animal studies, linkage and cytogenetic studies have now been sequenced (Vieira et al., 2005). This may increase the number of genes known to contribute to CL/P substantially, but we still have important work to do in the field of biological causality. We need to know which genetic mutations are both necessary *and* sufficient to cause CL/P to diagnose each individual patient. Identifying much more than three alleles acting together and cause CL/P may well turn out to be very complicated. In other words, the biological causality is obscure, and since it has been estimated that up to 14 loci (and therefore genes) contribute to the aetiology of CL/P (Schliekelman and Slatkin, 2002), this task may in fact turn out to be impossible.

One approach could be to focus on the key players in NSCLP. Maybe the future will show that NSCLP is caused by genetic alterations in genes acting in a few common signalling pathways (as *SOX9* dysregulation may cause PRS, which was suggested in Paper II). This may lead us to focus on dysregulation of pathways instead of on each specific gene. An attractive aspect of the "pathway-approach" is that it may be possible to regulate the pathway "back to normal".

The number of genes listed in the ***Appendix*** will probably increase in the coming years. Perhaps this will position us to select and sequence some candidate genes for each individual patient with CL/P. Initially, this screening may be limited to cases where there is evidence of genes playing a major role in the aetiology, i.e. multiplex families or patients with syndromic CL/P; but in the future, screening and genetic counselling may well include all NSCLP patients.

Gene	Comments on the genetic contribution to NSCLP	Authors
IRF6	Common polymorphisms in *IRF6* contribute to ~12% of the genetic aetiology in NSCLP, but no specific disease-causing mutations have been identified	Zucchero et al., 2004
MSX1	*MSX1* mutations and rare variants are found in ~2% of cases with NSCLP	Lidral et al., 1998; Jezewski et al., 2003 and Tongkobpetch et al., 2006
PAX9	1/128 NSCLP had a mutation in *PAX9*	Ichikawa et al., 2006
PTCH	Three rare variants of *PTCH* were associated with NSCLP	Mansilla et al., 2006
PVRL1	Heterozygosity of the nonsense mutation W185X is associated with NSCLP. Rare and common variants of *PVRL1* are associated with NSCLP	Sozen et al., 2001, Scapoli et al., 2006 and Avila et al., 2006
PVR and *PVRL2*	Association of NSCLP and a *PVR* marker. Perhaps rare variants in *PVR* and *PVRL2* contribute in causing NSCLP	Warrington et al. 2006
RYK	1/355 NSCLP patients had a missense mutation in *RYK*	Watanabe et al., 2006
SKI	A polymorphism (257C>G) in *SKI* exon 1 was associated with decreased risk of NSCLP	Lu et al., 2005
TBX22	Up to 4% of NSCLP have coding mutations in *TBX22*	Marcano et al., 2004
TP73L	1/100 NSCLP patients had a non-synonymous change in *TP73L* predicted to have a damaging function	Leoyklang et al., 2006
TGFB3	Association of NSCLP and *TGFB3*. 2/93 patients with NSCLP had rare variants in *TGFB3*	Lidral et al., 1998
FOXE1, GLI2, JAG2, LHX8, MSX1, MSX2, SATB2, SKI, SPRY2 and *TBX10*	Mutations in these genes may provide 5-6% of the genetic contribution in NSCLP	Vieira et al., 2005; Ichikawa et al., 2006

Table 1 A list of genes relevant to screen for in non-syndromic CL/P (NSCLP). Mutations or common variants in these genes contribute to the aetiology of NSCLP to some extent. The comments on the genetic contribution are from the papers referred to.

Cytogenetic studies

Studying chromosomal aberrations in patients with the disease phenotype of interest is a useful method of obtaining ideas for candidate genes and loci. Recent refinements in cytogenetic techniques, such as Comparative Genome Hybridisation (CGH) and Flourescence In Situ Hybridisation (FISH), have increased the chances of detecting chromosomal aberrations. Highlighting the relevance of the chromosomal approach is the recent work performed by Alkuraya and colleagues (2006), who found evidence for the gene *SUMO1* to play a role in palate development. The study identified a patient with NSCLP and an abnormal karyotype 46,XX,t(2;8)(q33.1;q24.3), where *SUMO1* on 2q was interrupted by the breakpoint. Subsequently, a *Sumo1* mouse mutant was generated which displayed cleft palate or oblique facial cleft in 4 out of 46 mice, compared with none of the wild type mice. Likewise, although we did not generate a mouse model, the identification of a translocation t(2;17)(q23.3;q24.3) in a patient with PRS and subsequent genetic studies in PRS patients prompted us to suggest that dysregulation of the transcription factor *SOX9,* and perhaps *KCNJ2*, may be involved in the aetiology of PRS.

A structured compilation of large genetic datasets facilitates the dissection of genetic diseases. The Wilhelm Johannsen Centre for Functional Genome Research (WJC) has established a comprehensive cytogenetic database, the Mendelian Cytogenetics Network Database (MCNdb), from which information on chromosomal aberrations was retrieved to identify potential candidate genes in the Pierre Robin sequence (Paper I).

Searching for "cleft" in MCNdb (November 2006) identified 70 cases with cleft lip and/or palate (including midline clefts, but omitting rare facial clefts). Most frequently, the chromosomes involved are 1, 2, 6, 7, 9 and 17 with ten or more cases having breakpoints involving these chromosomes. In the ***Appendix*** the chromosomal aberrations involved in patients with CL/P from the MCNdb are listed (most often the cases were syndromic CL/P forms or CL/P with associated malformations). The chromosomal aberrations are listed according to the loci involved (e.g., t(4;11) is listed at chromosomes 4 and 11).

Because the genome shows hotspots for rearrangements, some loci will be over- or underrepresented in cytogenetic databases (Bailey and Eichler, 2006) and consequently, important loci for CL/P will not necessarily present as the loci most frequently involved. Another issue causing difficulties with the cytogenetic approach is that more and more breakpoints are mapped to non-coding regions, potentially representing regulatory regions. These non-coding regulatory regions may be located up to ~1 Mb away from the gene (as showed in Paper II, concerning the translocation > 1Mb away from *SOX9*) and even within intronic regions of neighbouring genes (Kleinjan and van Heyningen, 2005), making it a troublesome task to unravel which gene caused the disease. Despite the inherent limitations of the cytogenetic approach, it serves as important primary steps in the positional cloning of a disease gene.

Expression studies

This is a growing field, because array-based methods are constantly being refined and increasing in capacity. Expression data is useful for studying gene regulation and signalling pathways and thus may convey information on the dynamic aspects of the processes underlying CL/P.

The Craniofacial and Oral Gene Expression Network (COGENE), which is a publicly available database of gene expression in the human embryonic craniofacial region, serves as an important tool for validating CL/P genes suggested on the basis of other studies. Beside COGENE only a few expression studies concerning CL/P have been published, and only in mice (Brown et al., 2003; Mukhopadhyay et al., 2004). These studies reported changes in expression profiles during the closure of the palate in murine embryos.

In the presented Paper III, expression analyses were performed in lip and palate tissue from patients with CL/P in order to study differences in expression patterns between the CL/P subgroups, because epidemiological data indicates different aetiologies in the subgroups. Most convincingly, Affymetrix Genechip analyses identified osteopontin (*SPP1*) and several other genes related to the immune response as being up-regulated in palate tissue from patients with cleft lip and palate compared to isolated cleft palate. Interestingly, some of these genes, osteopontin (*SPP1),* chemokine receptor 4 *(CXCR4)* and serglycin *(PRG1*) were validated by immunohistochemical staining in human fetal palate, because these genes were expressed in the palatal shelves during the time of palatal fusion.

Expression analysis should be interpreted with caution though, and at least validated using several other approaches, but they offer ways to gain more insight into the complex signalling pathways of genes.

Own work

The following part presents the findings made during the project period, and includes a description of the background of the study, material and methods, results, and a critical discussion of the work.

Paper I

Reviewing the genetic basis of the Pierre Robin Sequence

Background

The Pierre Robin Sequence (PRS) is a heterogeneous group, because it forms part of the phenotypic features in many syndromes (Cohen, 1999). PRS consists of cleft palate, micrognathia and respiratory difficulties in the early neonatal period caused by glossoptosis (MIM#261800). Even though the aetiology in some syndromic PRS forms is known, such as Stickler syndrome (MIM#108300, #604841, #184840), the aetiology in non-syndromic or isolated PRS is not fully elucidated. As explained in "*Syndromic CL/P*", dissecting the aetiology of PRS may be useful, because the genetic basis of some syndromic CL/P forms may also be applied to non-syndromic CL/P. Likewise, PRS may represent an extreme isolated cleft palate (CP) phenotype, because CP also has micrognathia to a certain degree (Eriksen et al., 2006). To gain insight into the genetic basis of PRS, we searched the literature and a cytogenetic database (MCNdb).

Material and Methods

We searched Pubmed for "pierre robin and genetics" and MCNdb for "robin" and "pierre robin". Our findings were subsequently compared with data obtained by Brewer and colleagues (1998, 1999), who had searched the Human Cytogenetics Database and identified duplications and deletions significantly ($p<.05$) associated with cleft palate and micrognathia.

Results

Loci 2q24.1-q33.3, 4q32-qter, 11q21-q23.1 and 17q21-q24.3 were identified more than once, and were therefore suggested as potential important loci for PRS. The genes *GAD1* (2q31.1), *PVRL1* (11q23.3) and the *SOX9* gene (17q24.3) are located in these regions and were suggested as being important in PRS, because they have been involved in or associated with CL/P in humans. Unfortunately, an error has occurred in the paper, as it is stated that *PVRL1* mouse mutants have CL/P, which is not true.

Discussion

The presented approach may serve as useful guidelines where to look for PRS candidate genes, although it does not exclude other regions from being important in PRS. We focused on the most frequently involved loci, but loci encountered less frequently in PRS may be of importance too, as our genome has hotspots for rearrangements (Bailey and Eichler, 2006). Thus, some chromosomal aberrations in these hotspot regions may be over-represented, despite the regions may not be more important than other less frequently represented regions, in terms of harbouring disease genes.

Moreover, pointing to a single disease gene in each locus may be too simplistic. Recently, it has been shown that the expression of a range of genes in the surroundings of a chromosomal aberration changes (Merla et al., 2006), suggesting that a phenotype conveyed by a chromosomal aberration may be the result

of dysregulation of many genes. This may point to, that the suggested loci for PRS harbour several genes important in PRS. Recent studies support this, as additional genes located in the loci suggested in the present study, may play a role in PRS and CL/P: *SATB2* (Vieira et al., 2005) and *SUMO1* (Alkuraya et al., 2006) in 2q24.1-q33.3 and *WNT9B* (Juriloff et al., 2006), *RARA* (Peanchitlertkajorn et al., 2003) and *KCNJ2* (Andelfinger et al., 2003) in 17q21-q24.3.

Paper I is presented on page 51.

Paper II

Genetic and cytogenetic studies in the Pierre Robin Sequence

Background

As mentioned in the Paper I, it is useful to study the genetic basis of the PRS, because the aetiology in PRS may also be valid in non-syndromic CL/P (NSCLP). Since little is known about the aetiology of non-syndromic or isolated PRS, we studied a group of ten patients presenting with isolated PRS.

Material and methods

For a period of one year, from 1st March 2003 until 28th February 2004, all the patients who were admitted to the Department for Plastic and Reconstructive Surgery, at Rigshospitalet, for primary repair of the orofacial cleft were included in the study (parental informed consent was obtained). Each child had a blood sample drawn during the operation, and in addition some of the patients donated lip and palate tissue samples (see Paper III). Since surgical treatment of CL/P is centralised in Denmark, all patients born in Denmark, Greenland and the Faroe Islands are operated at the Department for Plastic and Reconstructive Surgery. During this one year period, the department admitted 148 patients for primary repair. The families of 121 children accepted to participate. Of the 121 patients, 5 patients had non-syndromic PRS. During the hospitalisation of their child, the parents also donated blood samples, so DNA samples from triads (the child with the cleft and the parents) were obtained. Some of this DNA was also included in the European multi-centre study, EUROCRAN (work package 2), where DNA from approximately 1000 European CL/P triads have been collected for future large scale association and sequencing studies.

To obtain material from additional PRS patients, we also approached non-syndromic PRS patients born in the eastern part of Denmark in the period 1988-1991. Five out of seven patients agreed to participate.

In total, DNA samples were obtained from ten patients with non-syndromic PRS. The study was approved by the local scientific ethics committee (KF 01-168/02 and KF 01-185/00). The patients were examined thoroughly using chromosome analyses, fluorescence in situ hybridisation, southern blotting, comparative genome hybridisation, sequencing and quantitative PCR, and skeletal X-ray surveys.

Results

One of the ten PRS patients had a balanced translocation 46,XX,t(2;17)(q23.3;q24.3) involving a non-coding and potentially regulating region 1.13 Mb upstream of the transcription factor *SOX9* and 0.8 Mb downstream of the gene *KCNJ2*. Moreover, by QPCR we found reduced *SOX9* and *KCNJ2* expression in lymphoblastoid cell lines from the PRS patients compared to a group of normal controls.

In addition, a microdeletion was detected by comparative genome hybridisation (CGH), in one of the patients with PRS, involving the gene *ZNF804B*.

Discussion

Mutations in *SOX9* and balanced translocations up to 950 kb upstream of *SOX9* result in Campomelic Dysplasia (MIM#114290), a rare skeletal (and often lethal) dysplasia in which features of PRS are included. *SOX9* had not been related to non-syndromic PRS earlier, but while we were mapping our case with the translocation, two interesting articles on position effects of *SOX9* in mild campomelic dysplasia/ mild skeletal dysplasia were published (Hill-Harfe et al., 2005 and Velagaleti et al., 2005), suggesting *SOX9* as a candidate gene for PRS as well. The patients in these papers were more severely affected than our translocation patient, and had breakpoints closer to *SOX9*.

Based on the translocation patient and the reduced expression of *SOX9* in PRS patients, the study suggested a role for *SOX9* dysregulation in PRS. Although no position effect of the gene *KCNJ2* (which encodes a potassium channel) has been reported before, *KCNJ2* may also be of interest in the aetiology of PRS. Mutations in *KCNJ2* in humans cause Andersens syndrome (MIM #170390) which include cleft palate, *Kcnj2* knockout mice display cleft palate and skeletal dysplasia (Zaritsky et al., 2000), and we found *KCNJ2* expression to be reduced in the PRS patients compared to a group of normal controls.

To further elucidate the role of *SOX9* and *KCNJ2* dysregulation as causative factors in PRS, and perhaps even isolated cleft palate (CP), it is important to sequence *SOX9* and *KCNJ2* in a larger cohort of patients with PRS and CP. Also, we may have to look for linkage disequilibrium of polymorphic variants, in the coding and non-coding regions of *SOX9* and *KCNJ2*. Peanchitlertkajorn and colleagues (2003) found significant association between NSCLP and marker D17S1301, which is located less than 3 Mb 3´of *SOX9*, perhaps pointing to a role for *SOX9* in NSCLP as well.

Future studies should also focus on the regulation of *SOX9* and *KCNJ2,* which may be exerted by neighbouring non-coding regions and genes involved in signalling pathways. *SOX9* signalling pathways may involve collagen encoding genes, *RUNX2, WNTs* and *TGFBs* (Dong et al., 2006). Melkoniemi and colleagues (2003) identified mutations in collagen encoding genes in isolated PRS patients, supporting that PRS may be caused by dysregulation of the *SOX9* signalling pathway in general.

Interestingly, in this study, a microdeletion was detected by CGH, in one of the patients with PRS, involving the gene *ZNF804B*, which encodes a C2H2-type zinc finger protein. One of our future projects is to investigate the parents of this patient by CGH, and sequence *ZNF804B* (and genes in signalling pathways) in a cohort of patients with PRS and NSCLP, to determine if this gene contributes to the aetiology of PRS and CL/P.

Paper II is presented on page 57.

Paper III

Gene expression in cleft lip and palate tissue

Background

Array-based gene expression analyses have been possible for some years now, and it provides the possibility of applying new strategies in CL/P research. Gene expression in cleft lip and palate tissue from CL/P patients has not been carried out before, and this approach may potentially convey useful information on genes and families of genes interacting and playing a role in the different subgroups of the CL/P population.

Epidemiological studies point to different aetiologies in cleft lip and palate subgroups, so we compared expression profiles in CL/P subgroups, as this enabled us to identify distinct expression profiles in the subgroups.

Although the presented expression data must be interpreted with caution, as they may not reflect expression profiles during embryonic development, we chose to perform this study and analyse the data subsequently by immunohistochemistry in sections of human embryonic palate.

Material and methods

During all the cleft lip and palate operations, small amounts of tissue are resected, both from the lip and soft palate, in order to obtain raw surfaces and optimal healing potentials. During the collection of blood samples (as described in *Material and methods* of Paper II), this resected tissue was also collected from a subgroup of the patients, when informed consent had been obtained. The study was approved by the local scientific ethics committee (KF 11-002/04).

Affymetrix GeneChip analysis was performed on 6 lip samples (3 from isolated cleft lip (CL) and 3 from combined cleft lip and palate (CLP)) and on 6 palate samples (3 from isolated cleft palate (CP) and 3 from combined cleft lip and palate (CLP)). Comparisons of the expression profiles in the two CL/P subgroups for each type of tissue (lip and palate tissue) were then made.

The results were also analysed by quantitative PCR, immunohistochemistry on craniofacial tissue sections from human embryos and results were compared to gene expression profiles from the craniofacial region in human fetuses (COGENE).

Results

Several genes showed significantly differential expression patterns in the lip and the palate tissue, but only results from the palate tissue were validated by QPCR. Most notably, osteopontin (*SPP1*) and some immune related factors had a significantly higher expression in the palate tissue from patients with CLP than in patients with CP. This finding was supported by immunohistochemical staining of sections of human embryonic palate which showed that the genes *SPP1,* chemokine receptor 4 (*CXCR4)* and serglycin *(PRG1)* were present in the fusing palate.

Discussion

The validation of the selected genes *SPP1, CXCR4* and *PRG1* in sections of human embryonic palate supported the usefulness of our approach.

In agreement with the epidemiological data in the field, the expression analysis indicated that molecular processes involved in palatal clefting differ, depending on the type of cleft, CLP or CP.

Moreover, the analyses in palate tissue from patients with isolated cleft palate and combined cleft lip and palate, suggested that genes otherwise related to the immune response, may also play a role in the development of the palate. A few already identified CL/P genes, such as *IRF6* (Zucchero et al., 2004) and *PVRL1* (Suzuki et al., 2000) may play a role in the immune system. Our study highlights this link between CL/P and immunology as it points to that some immune-related genes may play a role in craniofacial development.

Because we identified expression levels, the genes identified as differentially expressed may not be the direct cause of CL/P. More likely, dysregulation of signalling pathways may have conveyed the CL/P phenotype (in agreement with the discussion of the *SOX9* dysregulation in Paper II). Thus, sequencing of the genes identified as differentially expressed may not reveal many mutations, although it may perhaps reveal some.

This study may serve as an initial approach towards gaining insight into the different molecular processes underlying the development of the palate in CP and CLP. Future challenges involve the studying of signalling pathways and interactions of the genes identified. This was tried in the present study by searching for protein interactions, and as the bioinformatics tools will improve, this approach will become more feasible. Also, we need to know more about the role of the immune system in craniofacial development in general. This field is very complicated, and necessitates the use several approaches.

The continuous refinement of array-based techniques and bioinformatics tools may facilitate progress in these areas.

Paper III is presented on page 77.

Paper IV

A linkage study in a CL/P multiplex family

Background

Although linkage studies have some inherent limitations as pointed out in the section on "*Linkage studies and linkage disequilibrium tests,"* they are also extremely valuable in dissecting genetic diseases, because they point to specific genomic regions where to focus our attention. At the same time, regions not involved in the disease of interest (negative LOD scores) are also identified.

In the 1980s, the early days of linkage analyses, Hans Eiberg and colleagues approached Danish multiplex CL/P families and collected blood samples in order to perform linkage analysis in CL/P multiplex families. DNA markers were extremely limited at that time, but their study of 58 multiplex families showed significant linkage (LOD score of 3.66) to *F13A1* on the short arm of chromosome 6 (Eiberg et al., 1987). Although the CL/P causing gene(s) in this locus (6p23-p25) have not been determined yet (but many genes have been suggested, see ***Appendix***), this locus has been verified in several subsequent CL/P linkage studies.

We performed a 25-year follow-up on these families, as it was possible that a new generation had been born in these families, and because linkage markers and methods have been refined during this period. Our initial idea was to perform a genomewide scan in the total group of families, but this turned out to be too costly, so we focused on the largest CL/P family available.

Material and methods

A total of 48 multiplex families were approached and most of the members of these families consented to give information on family and medical history. Blood samples were obtained from members where samples were not obtained in the 1980s and from people born after 1980 (but not from children younger than 12 years old) to improve the power of the pedigrees. DNA samples were obtained from 347 people, of whom 119 were affected with some type of CL/P (this includes the samples collected in the 1980s). One family was identified as large enough to yield significant linkage and was thus chosen for a genome scan. The study was approved by the local scientific ethics committee (KF 11-022/03).

Results

The genome scan (10K SNP) *suggested* linkage (LOD score of 2.73) to a region neighbouring, but not including, the CL/P gene, *IRF6*. The linkage interval was confirmed by microsatelite markers. In the most severely affected family member (with bilateral CL/P), we sequenced the translated region of selected genes and a group of functional RNAs, microRNAs, in the linkage interval and found no mutations in *SOX13, FMOD, OPTC, IKBKE* and microRNAs mir 29b-2, 29c and 135b. Moreover, no mutations were found in *IRF6* (although not located in the interval, it was important to rule out a disease-causing mutation in this gene) and CGH was normal in this person.

Discussion

It is generally recommended that the threshold for linkage is a LOD score of ~3.3 in genome scans (Lander and Kruglyak, 1995). The presented linkage study only yielded *suggestive* linkage, and in accordance with this it should be interpreted with caution. On the other hand, significant linkage results in complex diseases are rare findings, and in reality, any chromosomal loci with LOD of more than 2 is potentially interesting (Altmüller et al., 2001). It was surprising to find that the gene *IRF6*, in which mutations cause the two allelic CL/P syndromes, Van der Woude syndrome and popliteal pterygium syndrome (Kondo et al., 2002), was excluded from the linkage interval at 1q32. Since we did not identify any coding region mutations in the chosen candidate genes, the possibility of non-coding regulatory regions exerting an effect on *IRF6* must be considered, which is supported by the findings of long linkage disequilibrium blocks from 40 kb 5´to 100 kb 3´of *IRF6* (Zucchero et al., 2004).

We may have to sequence the total 6.5 Mb interval before we find the CL/P causing mutation, and if it is common variants (or combination of variants) which cause CL/P, it may be even more complicated. However, this interval might be narrowed further by testing markers in the other 47 multiplex families available.

Paper IV is presented on page 101.

Conclusion and Perspectives

The aim of this thesis was to investigate aspects of the genetic aetiology of CL/P, and different approaches were successfully adopted in order to achieve this.
In conclusion, the thesis offers new insights into the aetiology of PRS and CL/P, by pointing to candidate genes and signalling pathways potentially involved.
The main findings in relation to the specific aims were achieved as follows:

1) The cytogenetic and genetic studies in patients with the Pierre Robin Sequence (PRS) suggested a possible involvement of *SOX9* dysregulation in PRS. *SOX9* haploinsufficiency is known to cause campomelic dysplasia, a rare skeletal dysplasia in which features of PRS are part of the phenotype, but no studies have pointed to a role of *SOX9* in isolated PRS before. Moreover, the expression studies pointed to a possible role for *KCNJ2* in PRS as well. This was supported by our review of genetic data on PRS, which indicated that the chromosomal locus 17q21-q24.3 might harbour a gene or genes involved in PRS.

2) Comparison of expression profiles in palate tissue from patients with isolated cleft palate (CP) and combined cleft lip and palate (CLP) showed osteopontin (*SPP1*) and genes known to be involved in the immune response as differentially expressed. This finding was supported by immunohistochemical staining of sections of human embryonic palate, indicating that the genes *SPP1,* chemokine receptor 4 (*CXCR4)* and serglycin *(PRG1)* were present in the fusing palate. This suggests a role for these genes in the development of cleft palate, and that these genes may even play a different role in CP and CLP. Moreover, it was hypothesised that genes related to the immune response may play a role in craniofacial development.

3) A genomewide scan in one large CL/P multiplex family suggested linkage to a region neighbouring *IRF6* (1q32), but not including *IRF6*. The role of the gene *IRF6* in NSCLP is well-documented from association studies, but despite extensive search, no specific NSCLP causing mutations have been identified. Although the present linkage study only obtained suggestive linkage (LOD score of 2.73), it is interesting that *IRF6* was excluded from the linkage region.
This could point to the presence of another CL/P gene or cis-acting regulatory elements within the linkage interval, regulating *IRF6*.

The insights obtained, raises a number of future challenges. Genes with a potential role in CL/P are being suggested at a regular basis, and the list of genes in Table 1, will increase. But before we are able to diagnose each individual patient, we need to know more. To understand the role of the identified genes and related pathways in CL/P, it is necessary to sequence these genes in a large cohort of patients with CL/P. As many common variants will be identified, it is also necessary to sequence a large cohort of normal

controls, to be able to determine which variants are the disease-causing variants (as pointed out by Vieira et al., 2005). Currently, this demands a massive sequencing effort, but perhaps in a few years techniques will have developed, so that sequencing of larger genomic regions can be done at an acceptable cost.
Moreover, the findings in Paper II-IV all highlight the importance of gene regulation in the CL/P aetiology. Further studies are needed to elucidate this issue. We will use publicly available bioinformatic tools and the protein interaction database used in Paper III to study gene regulation. As knowledge on signalling interactions are increasing and refining, these tools will eventually become more and more useful.
As previously discussed, gene regulation may also be exerted by non-coding elements and functional, non-coding RNAs. In the future the role of one group of these non-coding RNAs, the microRNAs, will be studied in the context of CL/P aetiology, as they control gene regulation on the translational level. The first approach for this will be to analyse global microRNA expression in lip and palate tissue from cleft lip and palate patients (tissue samples as in Paper III) by a microRNA array. This approach may shed light on whether microRNAs are involved in the aetiology of the CL/P subgroups, and may provide more information on gene regulation in CL/P.

Moreover, this thesis highlights the importance of a thorough characterisation of the CL/P patients studied. Genetic dissection of CL/P syndromic forms or subgroups is useful and worth pursuing, as evidenced in Paper II studying the PRS and in Paper III, studying CL/P subgroups, but this genetic dissection is only possible, when the patients are diagnosed and categorised correctly. In order to ensure a correct diagnosis, genetic and clinical departments have to be in close contact and collaborate. Recognising the need for a precise diagnosis of the patients with CL/P, we have planned to carry out a prospective study on associated malformations in patients with CL/P, from which DNA will also be obtained. The precise diagnosis of patients may facilitate more focused genetic studies in clinically well-examined CL/P subgroups.
These are important future challenges in genetic research of CL/P, and they will be facilitated by the high throughput sequencing and array-based methods that develops and improves constantly. These approaches may eventually lead to a more dynamic understanding of the molecular processes leading to CL/P, which will enable us to improve the diagnosis and counselling of patients and families affected by CL/P and to develop rational intervention that may decrease the risk of developing CL/P, or decrease the extent of the malformation.

Finally, the genetic aspects of CL/P studied in this thesis may constitute an important piece in the complex puzzle underlying this common craniofacial malformation. If this puzzle is ultimatively solved, the applied approaches may serve as template for the dissection of other congenital malformations and the wide range of other diseases with a complex background.

Summary

The thesis is based on work performed at Wilhelm Johannsen Centre for Functional Genome Research and the Department of Plastic and Reconstructive Surgery and Burns, Rigshospitalet in the period 2003–2006. It comprises four studies.

The general aim of this thesis was to study genetic aspects of cleft lip and palate, which is a common and complex congenital malformation caused by environmental and genetic factors. This was achieved by studying in detail the Pierre Robin Sequence (PRS), which is a subgroup of the cleft lip and palate population. Moreover, the genetic aspects were studied by performing gene expression profiles in lip and palate tissue from patients with different types of cleft lip and/or palate (CL/P), and by carrying out a linkage study in a large CL/P multiplex family.

It was found, that part of the aetiology in PRS might be caused by dysregulation of the gene *SOX9,* and perhaps also the gene *KCNJ2*.

Expression of genes in lip and palate tissue suggests that the gene osteopontin (*SPP1*), and other genes normally related to the immune response may have a function in the development of the palate, a function that may differ according to the cleft palate type.

Finally, the genomewide linkage study in the CL/P multiplex family suggested linkage to a chromosomal region (1q32) close to, but not including a well-known CL/P gene, *IRF6*. It was suggested that this region may harbour genes or non-coding elements, regulating *IRF6*, although genes with a distinct function in CL/P development are a possibility too.

Together, the thesis point to several new candidate genes and signalling pathways potentially involved in the aetiology of PRS and CL/P.

Dansk resumé (Danish summary)

Denne afhandling er baseret på arbejde udført på Wilhelm Johannsen Centret for Funktionel Genomforskning og Klinik for Plastikkirurgi & Brandsårsbehandling, Rigshospitalet i perioden 2003-2006. Afhandlingen består af fire studier.

Det overordnede formål med afhandlingen var at belyse genetiske aspekter ved læbe-ganespalte, som er en hyppig medfødt misdannelse, der er forårsaget af både miljømæssige og genetiske faktorer. Dette mål blev opfyldt ved i detaljer at undersøge patienter med Pierre Robin Sekvens (PRS), som er en delgruppe af den generelle gruppe af patienter med læbe-ganespalte.

Desuden blev der udarbejdet gen ekspressions profiler i læbe- og ganevæv fra patienter med forskellige former for læbe-ganespalte, og en genomscanning og koblingsanalyse af en familie med flere medlemmer med læbe-ganespalte (multiplex familie) blev udført.

Det blev vist, at en del af årsagerne til PRS muligvis kan forklares ved ændret regulering af genet *SOX9*, og at genet *KCNJ2* måske også spiller en rolle for udviklingen af PRS.

Gen ekspressions studiet i læbe- og gane væv tydede på, at genet osteopontin (*SPP1*) og flere andre gener, hvor nogle af disse normalt er involveret i immunresponset, kan spille en rolle for udviklingen af ganespalte, samt at disse gener kan have forskellige funktioner i forskellige typer af ganespalte.

Endelig tydede genom-scanningen af multiplex familien på mulig kobling til et område (1q32) tæt på et velkendt læbe-ganespalte gen, *IRF6*, dog uden at inkludere *IRF6* i intervallet. Dette kan tyde på, at koblingsintervallet indeholder gener eller ikke-kodende elementer, der regulerer *IRF6*. Det kan dog også betyde, at koblingsintervallet indeholder gener med selvstændig betydning for udviklingen af læbe-ganespalte.

Samlet set, peger afhandlingen på flere nye kandidatgener og signaleringsveje, som kan være involveret i årsagerne til PRS og læbe-ganespalte.

Appendix – An overview of CL/P candidate genes and loci (see page 45 for further information)

Chromosome	Animal studies	Linkage-, LD- and association studies	CL/P syndromes	Chromosomal aberrations
1				
				46,XY,t(1;5)(p11;q23)
	Procollagen, type XI, alpha 1 (Col11a1, 1p21.1)	1p12-13 (COL11A1)	Marshall and Stickler type II (COL11A1, 1p21.1)	
				46,XX,t(1;16)(p22;q12.1)
				46,XY,t(1;11)(p22;p13)
	Heparan sulfate 2-O-sulfotransferase 1 (Hs2st1, 1p22.3)			
				46,XY,t(1;5)(p31;q14)
	LIM homeobox protein 8 (Lhx8, 1p31.1)			
				46,XY,ins(1)(p31.2;q23q42)
		1p32		
			Desmosterolosis (DHCR24, 1p32.3)	
	Endothelin 2 (Edn 2, 1p34)			
				46,XX,t(1;7)(p35;q36)
	Polyhomeotic-like 2 (Phc 2, 1p35.1)			
		1p36 (MTHFR, PAX7)		
	Stratifin (Sfn, 1p36.11)			
	Perlecan (Hspg2, 1p36.12)			
	Catenin beta interacting protein 1 (Ctnnbip1, 1p36.22)			
				46,XY,inv(1)(p36.3q42)
				46,XY,t(1;16)(q11;q12)
		1q21 (ARNT)		
				46,XY,ins(1)(p31.2;q23q42)
				46,XX,t(1;12)(q31;p21)
		1q32 (IRF6)	Van der Woude syndrome and popliteal pterygium syndrome (IRF6, 1q32.2)	
	Calcium channel, voltage-dependent, L type, alpha 1S subunit (Cacna1s, 1q32.1)			
	Transforming growth factor, beta 2 (Tgfb2, 1q41)			
				46,XX,t(1;3)(q42;q12),t(3;1)(q22;q42),ins(4;3)(q31;q12q22)
				46,XY,ins(1)(p31.2;q23q42)
2				
				46,XY,inv(2)(p12q12)
		2p13 (TGFA, VAX2)		
			Holoprosencephaly-type 2 (SIX3, 2p21)	
	Son of sevenless homolog 1 (Sos1,			

	2p22.1)			
		2p25		
	SRY-box containing gene 11 (Sox11, 2p25.2)			
				46,XY,t(2;20)(q11;p11)
				46,XY,inv(2)(p12q12)
	GLI-Kruppel family member GLI2 (Gli2, 2q14.2)	GLI2 (2q14.2)		
	Inhibin beta-B (Inhbb, 2q14.2)			
				46,XX,t(2;18)(q21;p11)
		ZFHX1B (2q22.3)	Mowat-Wilson syndrome (ZFHX1B, 2q22.3)	
	Activin receptor IIA (Acvr2A, 2q22.3-q23.1)			
				46,,XX,t(2;10)(q23;p13)
				46,XY,t(2;4)(q23;q33)
			Nemaline myopathy-2 (NEB, 2q23.3)	
				46,XX,t(2;17)(q31;q25)
	Glutamate decarboxylase 1 isoform (GAD1, 2q31.1)	2q31 (GAD1, ZNF533)		
	Distal-less homeobox 1 and 2 (Dlx1/2, 2q31.1)			
	Integrin alpha V (Itgav, 2q32.1)			
				46,XY,t(2;6)(q33;p25)
	Small ubiquitin-like modifier 1 isoform a (Sumo1, 2q33.1)			
	SATB family member 2 (Satb2, 2q33.1)	2q32-35 (SATB2, WNT6, WNT10A, DLX2, TNS1)		
				46,XY,t(2;17)(q35;q23)
				46,XX,t(2;14)(q36;q21)
			Waardenburg syndrome type I (PAX3, 2q36.1)	
		2q36.3 (COL4A3, COL4A4)		
		2q37 (SP100, MLPH, HDAC4)		
3				
			Larsen syndrome atelosteogenesis, spondylocarpotarsal synostosis syndrome (FLNB, 3p14.3)	
		3p21.2		
			Hyaluronidase deficiency (HYAL1, 3p21.31)	
	Transforming growth factor, beta receptor II (Tgfbr2, 3p24.1)		Marfan, Loeys-Dietz syndrome (TGFBR2, 3p24.1)	
		3p25 (WNT7A, RARB, THRB, TGFBR2)		
				46,XX,t(1;3)(q42;q12),t(3;1)(q22;q42),ins(4;3)(q31;q12q22)
				46,X,t(X;3;8)(q13.1;q13.1;q24.3)
				46,XX,t(3;18;8)(q13.1;q22.2;q21.2)

				46,XY,t(3;10)(q13.2;q21.2)
	Receptor-like tyrosine kinase (Ryk, 3q22.1)			
			Seckel syndrome (ATR, 3q23)	
	Short stature homeobox 2 (Shox2, 3q25.32)		Cornelia de Lange syndrome (SHOX2, 3q25.32)	
		3q26		
	Transformation related protein 63 (Trp63, 3q28)		TP73L (3q28) mutations cause acro-dermato-ungual- lacrimal-tooth syndrome (ADULT syndrome), ankyloblepharon- ectodermal defects-cleft lip/palate (AEC syndrome), ectrodactyly, ectodermal dysplasia, Limb-mammary syndrome (LMS), Rapp-Hodgkin syndrome and cleft lip/palate syndrome 3 (EEC3) and split-hand/foot malformation 4 (SHFM4)	
				46,XY,t(3;14)(q29;q24)
	Discs, large homolog 1 (Dlg1, 3q29)			
4				
				46,XY,t(4;20)(q;p)
				46,XX,inv(4)(p14q27)
			Ellis-van Creveld syndrome (EVC, 4p16.1)	
	Homeo box, msh-like 1 (Msx1, 4p16.2)	4p16 (MSX1)	Wolf-Hirschhorn syndrome, Witkop syndrome, autosomal dominant hypodontia with or without orofacial clefting (MSX1, 4p16.2)	
				46,XX,der(4),t(4;7)(p16.3;p22)
				46,XY,inv(4)(p16q12-13?)
	Platelet-derived growth factor receptor alpha (Pdgfra, 4q12)			
				46,XY,t(4;7)(q21;q36)
			Fraser syndrome (FRAS1, 4q21.1-q21.21)	
				46,XX,t(4;12)(q22;q24.5)
	Paired-like homeodomain transcription factor 2 (Pitx2, 4q25)		Rieger Syndrome Type 1 (PITX2, 4q25)	
				46,XX,t(4;17)(q21;q23)
		4q21-31 (LEF1, FGF2, BMPR1B, SMAD1)		
				46,XX,t(1;3)(q42;q12),t(3;1)(q22;q42),ins(4;3)(q31;q12q22)
				46,XX,inv(4)(p14q27)
		Anonymous markers (D4S175 and D4S192) at 4q31.1-q31.2		
	Endothelin receptor type A (Ednra, 4q31.23)			
	Platelet-derived growth factor, C polypeptide (Pdgfc, 4q32.1)			
				46,XY,t(2;4)(q23;q33)
				46,XY,del(4)(q33;ter)

5				
	Fibroblast growth factor 10 (Fgf10, 5p12)			
				46,XY,t(5;19)(p13;q10)
			Cornelia de Lange syndrome (NIPBL, 5p13.2)	
	Methionine synthase reductase (Mtrr, 5p15.31)			
		5q11		
				46,XY,t(1;5)(p31;q14)
	Ephrin A5 (Efna5, 5q21.3)			
				46,XY,t(5;9)(q22;p13)
				46,XY,t(1;5)(p11;q23)
	Paired-like homeodomain transcription factor 1 (Pitx1, 5q31.1)			
			Neonatal osseous dysplasia /diastrophic dysplasia (DTDST, 5q33.1)	
			Treacher Collins syndrome (TCOF1, 5q33.1)	
	Homeo box, msh-like 2 (Msx2, 5q35.2)		Craniosynostosis type 2, Parietal Foramina 1 (MSX2, 5q35.2)	
6				
		6p12		
	Runt related transcription factor 2 (Runx2, 6p12.3)		Cleidocranial dysplasia (RUNX2, 6p12.3)	
	Vascular endothelial growth factor A (Vegfa, 6p21.1)			
				46,XY,inv(6)(p21.3q22)t,(6;17)(q22 ;q23)
	Procollagen, type XI, alpha 2 (Col11a2, 6p21.32)		Otospondylomegaepiphyseal dysplasia and Stickler type III (COL11A2, 6p21.32)	
		6p23-25 (F13A1, TFAP2A, BMP6, EDN1, COL11A2, OFC1, C6orf105)		
				46,XX,t(6;9)(p24;p13)
	Endothelin 1 (Edn1, 6p24.1)			
				46,X,del(6)(p24.3),del(6)(p25p25)
	Transcription factor AP-2, alpha (Tcfap2a, 6p24.3)			
				46,XY,t(2;6)(q33;p25)
				46,XX,t(6;7)(p25;q31)
	Forkhead box F2 (Foxf2, 6p25.3)			
				46,XY,inv(6)(p21.3q22)t,(6;17)(q22 ;q23)
			Oculodentodigital dysplasia (GJA1, 6q22.31)	
	Transcription factor 21 (Tcf21, 6q23.2)			
			Rhizomelic chondrodysplasia	

			punctata type 1 (PEX7, 6q23.3)	
		6q23-25		
				46,XX,t(6;17)(q27;q25,1)
7				
	Epidermal growth factor receptor (Egfr, 7p11.2)			
		7p12 (EGFR)		
				46,XX,t(7;16)(p13;q22)
	GLI-Kruppel family member GLI3 (Gli3, 7p14.1)		Greig cephalopolysyndactyly, Pallister-Hall and acrocallosal syndrome, preaxial polydactyly type IV, postaxial polydactyly types A1 and B (GLI3, 7p14.1)	
	Inhibin beta-A (Inhba, 7p14.1)			
		7p15		
	Homeo box A1/A2 (Hoxa1/2, 7p15.2)			
	Integrin beta 8 (Itgb8, 7p15.3)			
		7p21		
	Mesenchyme homeobox 2 (Meox2, 7p21.1)			
	Twist gene homolog 1 (Twist1, 7p21.1)		Saethre-Chotzen Syndrome (TWIST1, 7p21.1)	
				46,XX,der(4),t(4;7)(p16.3;p22)
				46,XX,t(6;7)(p25;q31)
				46,XX,t(7;20)(q32.1;q13.2)
				46,XX,dir ins(7)(q21.3;q22q31.1)
				46,XX,t(1;7)(p35;q36)
				46,XY,t(4;7)(q21;q36)
				46,XY,t(7;11)(q36;q13.3)
				46,XX,del(7)(q36)
	Sonic hedgehog (Shh) (7q36.3)		Holoprosencephaly-3 (SHH, 7q36.3)	
8				
	MYST histone acetyltransferase 3 (Myst3, 8p11.21)			
		8p21 (FGFR1, EGR3, PPP3CC, FZD3, NRG1)		
			Kallmann syndrome (KAL2) and craniosynostosis (FGFR1, 8p12)	
			Roberts syndrome (ESCO2, 8p21.1)	
				46,XY,t(8;11)(q11;q22)
			Waardenburg syndrome type II (SNAI, 8q11.21)	
	Chromodomain helicase DNA binding protein 7 (Chd7, 8q12.2)		CHARGE syndrome (CHD7, 8q12.2)	
	Eyes absent 1 homolog (Eya1, 8q13.3)		Branchio-oto-renal Syndrome (EYA1, 8q13.3)	
	Musculin (Msc, 8q13.3)			
				46,XY,t(8;12)(q22;q21)
				46,XX,t(3;18;8)(q13.1;q22.2;q21.2)
			Nijmegen breakage syndrome (NBS1, 8q21.3)	
	Odd-skipped			

	related 2 (Osr2, 8q22.2)			
		8q23 (FZD6)		
				46,X,t(X;3;8)(q13.1;q13.1;q24.3)
	RecQ protein-like 4 (Recql4, 8q24.3)		Rothmund-Thomson and Rapadilino syndrome (RECQL4, 8q24.3)	
9				
				46,XX,t(6;9)(p24;p13)
				46,XY,t(5;9)(q22;p13)
		9q21 (PTCH, ROR2, TGFBR1, ZNF189, FOXE1, clf2 homology)		
	Receptor tyrosine kinase-like orphan receptor 2 (Ror2, 9q22.31)		Robinow syndrome (ROR2, 9q22.31)	
	Patched (Ptch, 9q22.32)		Gorlin syndrome, basal naevus syndrome (PTCH, 9q22.32)	
	Forkhead box E1 (thyroid transcription factor 2) (Foxe1, 9q22.33)		Congenital hypothyroidism and cleft palate with thyroid dysgenesis, Bamforth-Lazarus syndrome (FOXE1, 9q22.33)	
	Transforming growth factor, beta receptor I (Tgfbr1, 9q22.33)		Loeys-Dietz syndrome (TGFBR1, 9q22.33)	
			Walker-Warburg syndrome (FCMD, 9q31.2)	
	RAD23b homolog (Rad23b, 9q31.2)			
			Nail-patella syndrome (LMX1B, 9q33.3)	
			Walker-Warburg syndrome (POMT1, 9q34.13)	
10				
				46,XY,inv(10)(p11.2q21.2)
	Zinc finger homeobox 1a (Zfhx1a, 10p11.22)			
	Polycomb group ring finger 4 (Pcgf4, 10p12.31)			
				46,,XX,t(2;10)(q23;p13)
				46,XY,t(10;17)(p15;q21)
				46,XY,rev ish enh(10q11q11)
				46,XY,inv(10)(q11.2q25.2)
				46,XY,t(3;10)(q13.2;q21.2)
				46,XY,inv(10)(p11.2q21.2)
			Goldberg-Shprintzen syndrome (KIAA1279, 10q21.3)	
	3'-phosphoadenosine 5'-phosphosulfate synthase 2 (Papss2, 10q23.2)			
	Bone morphogenetic protein receptor, type IA (Bmpr1a, 10q23.2)			
				46,XY,inv(10)(q11.2q25.2)
				46,XY,t(10;14)(q24;q32.3)
				46,XY,t(10;12)(q24.1;p13.2)
	Conserved helix-loop-helix			

	ubiquitous kinase (Chuk, 10q24.2)			
			Split hand/foot malformation 3 (FBXW4, 10q24.32)	
		10q25		
	Fibroblast growth factor receptor 2 (Fgfr2, 10q26.13)		Crouzon, Pfeiffer, Apert, Jackson-Weiss, Beare-Stevenson cutis gyrata and Saethre-Chotzen syndrome (FGFR2, 10q26.13)	
11				
		11p11		
	Aristaless 4 (Alx4, 11p11.2)	11p11.2 (ALX4)		
				46,XY,t(1;11)(p22;p13)
		11p12-q14 (FGF3, FGF4, Folate receptor cluster, CD44)		
	Cyclin-dependent kinase inhibitor 1C (P57) (Cdkn1c, 11p15.4)			
	Integrin linked kinase (Ilk, 11p15.4)			
	T-box 10 (Tbx10, 11q13.2)			
				46,XY,t(7;11)(q36;q13.3)
	7-dehydrocholesterol reductase (Dhcr7, 11q13.4)		Smith-Lemli-Opitz Syndrome (DHCR7, 11q13.4)	
				46,XY,t(8;11)(q11;q22)
	Sterol-C5-desaturase (Sc5dl, 11q23.3)		Lathosterolosis (SC5DL, 11q23.3)	
			Margarita Island ectodermal dysplasia, Zlotogora-Ogur syndrome, and a cleft lip/palate-ectodermal dysplasia (PVRL1, 11q23.3)	
		11q24.1 (ZNF202)		
				46,XX,t(11;18)(q25;q21.2)
12				
				46,XX,t(12;20)(p10;q10)
		12p11		
	SRY-box containing gene 5 (Sox5, 12p12.1)			
		12q13		
				46,XY,t(10;12)(q24.1;p13.2)
	Polyhomeotic-like 1 (Phc1, 12p13.31)			
				46,XX,t(1;12)(q31;p21)
	Keratin 5 (Krt5, 12q13.13			
	Procollagen, type II, alpha 1 (Col2a1, 12q13.11)		Kniest dysplasia, Spondyloepiphyseal Dysplasia Congenita and Stickler type I (COL2A1, 12q13.11)	
				46,XY,dup(12)(q21;q22)
				46,XY,t(8;12)(q22;q21)
	Cartilage homeo protein 1 (Cart1, 12q21.31)			
	Apoptotic peptidase			

	activating factor 1 (Apaf1, 12q23.1)			
			LEOPARD and Noonan syndrome (PTPN11, 12q24.13)	
				46,XX,t(4;12)(q22;q24.5)
13				
	Fibroblast growth factor 9 (Fgf9, 13q12.11)			
			Fraser syndrome (FREM2, 13q13.3)	
				46,XY,t(13;16)(q14;q24)
	Intraflagellar transport 88 homolog (Ift88, 13q22.11)			
				46,XY,t(13;16)(q32;q13)
			Holoprosencephaly-type 5 (ZIC2, 13q32.3)	
		13q33.1-34 (TFDP1, ING1, COL4A1)		
14				
	Forkhead box G1 (Foxg1, 14q12)			
				46,XY,t(14;17)(q13;q24)
	Paired box gene 9 (Pax9, 14q13.3)		Hypodontia and oligodontia (PAX9, 14q13.3)	
				46,XX,t(2;14)(q36;q21)
		14q21-24 (TGFB3, BMP4)		
	Bone morphogenetic protein 4 (Bmp4, 14q22.2)			
				46,XY,t(3;14)(q29;q24)
				46,XY,t(14;17)(q24;q23)
	Transforming growth factor, beta 3 (Tgfb3, 14q24.3)			
			Walker-Warburg syndrome (POMT2, 14q24.3)	
		14q32		
				46,XY,t(10;14)(q24;q32.3)
	Jagged 2 (Jag2, 14q32.33)	14q32.33 (JAG2)		
15				
		15q12 (GABRB3)		
	Gamma-aminobutyric acid receptor, subunit beta 3 (Gabrb3, 15q12)			
		15q15 (ZNF291, NRG4)		
			Marfan syndrome, isolated ectopia lentis, autosomal dominant Weill-Marchesani syndrome, MASS syndrome, and Shprintzen-Goldberg craniosynostosis syndrome (FBN1, 15q21.1)	
				46,XX,t(15;20)(q25;p12)
		15q26		
	Aggrecan 1 (Agc1, 15q26.1)			
16				
				46,XY,inv(16)(p13q24)
		16p13.3 (MMP25)		

	Creb binding protein (Crebbp, 16p13.3)		Rubinstein-Taybi syndrome (CREBBP, 16p13.3)	
				46,X,t(X;16)(q21.3;q11.2)
				46,XY,t(1;16)(q11;q12)
				46,XX,t(1;16)(p22;q12.1)
				46,XY,t(13;16)(q32;q13)
				46,XX,t(7;16)(p13;q22)
		16q24		46,XY,inv(16)(p13q24)
				46,XY,t(13;16)(q14;q24)
	Forkhead box C2 (Foxc2, 16q24.1)		Lymphoedema-distichiasis syndrome (FOXC2, 16q24.1)	
17				
		17p11		
		17p13		
	Tumor protein p53 (P53, 17p13.1)			
	V-crk sarcoma virus CT10 oncogene homolog (Crk, 17p13.3)			
	Dph1 homolog (Dph1, 17p13.3)			
	Max binding protein (Mnt, 17p13.3)			
			Miller-Dieker lissencephaly syndrome (PAFAH1B1, 17p13.3)	
	Polycomb group ring finger 2 (Pcgf2, 17q12)			
				46,XY,t(10;17)(p15;q21)
		17q21 (RARA, WNT9B, *clf1* homology, CRHR1, MAPT)		
	Wingless-type 3 and 9 (Wnt 3/9, 17q21.32) *clf1* homology		Autosomal recessive tetra-amelia (WNT3, 17q21.32)	
				46,XY,t(2;17)(q35;q23)
				46,XX,t(4;17)(q21;q23)
				46,XY,inv(6)(p21.3q22)t,(6;17)(q22;q23)
				46,XY,t(14;17)(q24;q23)
			Small patella syndrome (TBX4, 17q23.2)	
				46,XY,t(14;17)(q13;q24)
	Potassium inwardly-rectifying channel, subfamily J, member 2 (Kcnj2, 17q24.3)		Andersen syndrome (KCNJ2, 17q24.3)	
	SRY-box containing gene 9 (Sox9, 17q24.3)		Campomelic dysplasia (SOX9, 17q24.3)	
				46,XX,t(2;17)(q31;q25)
	Phosphatidylserine receptor (Ptdsr, 17q25.2)			
			Hereditary neuralgic amyotrophy (SEPT9, 17q25.2-q25.3)	
18				
				46,XX,t(2;18)(q21;p11)

			Holoprosencephaly-4 (TGIF, 18p11.31)	
		18q21 (SMAD2)		
				46,XX,t(11;18)(q25;q21.2)
	Retina and anterior neural fold homeobox (Rax, 18q21.32)			
				46,XX,t(3;18;8)(q13.1;q22.2;q21.2)
				46,XX,t(3;18;8)(q13.1;q22.2;q21.2)
19				
				46,XY,t(5;19)(p13;q10)
		19q13 (BCL3, PVRL2, TGFB1, APOC2)		
20				
				46,XY,t(4;20)(q;p)
				46,XY,t(2;20)(q11;p11)
				46,XX,t(15;20)(q25;p12)
				46,XX,t(12;20)(p10;q10)
				46,XX,t(7;20)(q32.1;q13.2)
		20q13 (EDN3, BMP7)		
22				
	T-box 1 (Tbx1, 22q11.21)		DiGeorge/Velocardiofacial syndrome (TBX1, 22q11.21)	
	Meningioma 1 (Mn1, 22q12.1)			
X			Oculofaciocardiodental syndrome (BCOR, Xp11.4)	
	Calcium/calmodulin-dependent serine protein kinase (Cask, Xp11.4)			
			Siderius type X-linked mental retardation (PHF8, Xp11.22)	
			Renpenning syndrome (PQBP1, Xp11.23)	
				46,X,inv(X)(p21q24)
			Snyder-Robinson syndrome (SMS, Xp22.11)	
			Opitz syndrome (MID1, Xp22.2)	
	Oral-facial-digital syndrome type 1 gene homolog (Ofd1, Xp22.2)		Oral-facial-digital syndrome type I (OFD1, Xp22.2)	
		Xcen		
				46,Y,del(X)(q13::q21.3)
				46,X,t(X;3;8)(q13.1;q13.1;q24.3)
			Craniofrontonasal syndrome (EFNB1, Xq13.1)	
				46,X,t(X;16)(q21.3;q11.2)
			X-linked cleft palate and ankyloglossia (TBX22, Xq21.3)	
				46,X,inv(X)(p21q24)
				46,XY,rev ish enh(Xq26q26)
			Simpson-Golabi-Behmel overgrowth syndrome (GPC3, Xq26.2)	
			Otopalatodigital syndrome type I/II, periventricular nodular heterotopia I/IV, frontometaphyseal dysplasia and Melnick-Needles syndrome (FLNA, Xq28)	

Appendix

Cleft lip and/or palate (CL/P) candidate loci and genes based on evidence from animal-, linkage-, association studies, CL/P syndromes and chromosomal aberrations.

Animal studies: Information from Mouse Genome Informatics (MGI), The Jackson Laboratory, September 2006. MGI phenotypes were searched for "cleft palate". Some extra mouse mutants have been added to the list: *Tgfbr1* (Dudas et al., 2006), *Sumo1* (Alkuraya et al., 2006), *Crk* (Park et al., 2006), *Meox2* (Jin and Ding, 2006), *Satb2* (Britanova et al., 2006), *P53* (Baatuot et al., 2002), *Bmp4* and *Bmpr1a* (Liu et al., 2005). The mouse models have different genetic backgrounds, not specified in the table and they were of different genetically altered categories. The human chromosome positions for the genes are listed in brackets.

Linkage and linkage disequilibrium (LD) studies: searched for "linkage and cleft palate" in Pubmed (LOD scores > 2 and significance levels <.05 have been included). For reasons of clarity only positive findings have been listed.

Association studies (case-control): searched for "association and cleft palate" in Pubmed. The genes listed in brackets are either proposed by the authors of the linkage studies or specifically studied in LD or case-control association studies. For reasons of clarity only positive findings have been listed.

CL/P syndromes are syndromes including features of CL/P and with a known or assumed genetic defect (from Online Mendelian Inheritance in Man (OMIM), August 2006).

Gene positions are from UCSC Genome Bioinformatics site, March 2006 assembly.

Chromosomal aberrations are from Mendelian Cytogenetics Network Database (MCNdb), November 2006. A total of 70 cases (syndromic and non-syndromic CL/P, including midline clefts, but omitting rare facial clefts) with chromosomal aberrations were encountered in MCNdb, searching "cleft". For a complete description of the phenotypes, see MCNdb. Aberrations are listed according to the breakpoints (e.g., t(4;11) are listed twice, at chromosomes 4 and 11, respectively). Del, deletion; dup, duplication; t, translocation; ins, insertion; inv, inversion; der, derivative; rev ish enh, reverse in situ hybridization enhancement.

References

Alkuraya FS, Saadi I, Lund JJ, Turbe-Doan A, Morton CC, Maas RL. SUMO1 haploinsufficiency leads to cleft lip and palate. Science. 2006;313:1751

Altmüller J, Palmer LJ, Fischer G, Scherb H, Wjst M. Genomewide scans of complex human diseases: true linkage is hard to find. Am J Hum Genet. 2001;69:936-950

Artama M, Auvinen A, Raudaskoski T, Isojarvi I, Isojarvi J. Antiepileptic drug use of women with epilepsy and congenital malformations in offspring. Neurology. 2005;64:1874-1878

Avila JR, Jezewski PA, Vieira AR, Orioli IM, Castilla EE, Christensen K, Daack-Hirsch S, Romitti PA, Murray JC. PVRL1 variants contribute to non-syndromic cleft lip and palate in multiple populations. Am J Med Genet. 2006;140A:2562-2570

Baatout S, Jacquet P, Michaux A, Buset J, Vankerkom J, Derradji H, Yan J, von Suchodoletz H, de Saint-Georges L, Desaintes C, Mergeay M. Developmental abnormalities induced by X-irradiation in p53 deficient mice. In Vivo. 2002;16:215-221

Bailey JA, Eichler EE. Primate segmental duplications: crucibles of evolution, diversity and disease. Nat Rev Genet. 2006;7:552-564

Bille C, Winther JF, Bautz A, Murray JC, Olsen J, Christensen K. Cancer risk in persons with oral cleft-a population-based study of 8,093 cases. Am J Epidemiol. 2005;161:1047-1055

Braybrook C, Doudney K, Marcano ACB, Arnason A, Bjornsson A, Patton MA, Goodfellow PJ, Moore GE, Stanier P. The T-box transcription factor gene TBX22 is mutated in X-linked cleft palate and ankyloglossia. Nature Genet. 2001;29:179-183

Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal deletion map of human malformations. Am J Hum Genet. 1998;63:1153-1159

Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal duplication map of malformations: Regions of suspected haplo- and triplolethality- and tolerance of segmental aneuploidy- in humans. Am J Hum Genet. 1999;64:1702-1708

Britanova O, Depew MJ, Schwark M, Thomas BL, Miletich I, Sharpe P, Tarabykin V. Satb2 haploinsufficiency phenocopies 2q32-q33 deletions, whereas loss suggests a fundamental role in the coordination of jaw development. Am J Hum Genet. 2006;79:668-678

Brown NL, Knott L, Halligan E, Yarram SJ, Mansell JP, Sandy JR. Microarray analysis of murine palatogenesis: temporal expression of genes during normal palate development. Dev Growth Differ. 2003;45:153-165

Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet. 2003;361:598-604

Celli J, Duijf P, Hamel BC, Bamshad M, Kramer B, Smits AP, Newbury-Ecob R, Hennekam RC, Van Buggenhout G, van Haeringen A, Woods CG, van Essen AJ, de Waal R, Vriend G, Haber DA, Yang A, McKeon F, Brunner HG, van Bokhoven H. Heterozygous germline mutations in the p53 homolog p63 are the cause of EEC syndrome. Cell. 1999;99:143-153

Christensen K, Fogh-Andersen P. Cleft lip (+/- cleft palate) in Danish twins, 1970-1990. Am J Med Genet. 1993;47:910-916

Christensen K, Mitchell LE. Familial recurrence-pattern analysis of nonsyndromic isolated cleft palate - a Danish Registry study. Am J Hum Genet. 1996;58:182-190

Christensen K. The 20th century Danish facial cleft population--epidemiological and genetic-epidemiological studies. Cleft Palate Craniofac J. 1999;36:96-104

Christensen K, Juel K, Herskind AM, Murray JC. Long term follow up study of survival associated with cleft lip and palate at birth. BMJ. 2004;328:1405

Cohen MM Jr. Robin sequences and complexes: causal heterogeneity and pathogenetic/phenotypic variability. Am J Med Genet. 1999;84:311-315

Dong YF, Soung do Y, Schwarz EM, O'Keefe RJ, Drissi H. Wnt induction of chondrocyte hypertrophy through the Runx2 transcription factor. J Cell Physiol. 2006;208:77-86.

Dudas M, Kim J, Li WY, Nagy A, Larsson J, Karlsson S, Chai Y, Kaartinen V. Epithelial and ectomesenchymal role of the type I TGF-beta receptor ALK5 during facial morphogenesis and palatal fusion. Dev Biol. 2006;296:298-314

Eiberg H, Bixler D, Nielsen LS, Conneally PM, Mohr J. Suggestion of linkage of a major locus for nonsyndromic orofacial cleft with F13A and tentative assignment to chromosome 6. Clin Genet. 1987;32:129-132

Eriksen J, Hermann NV, Darvann TA, Kreiborg S. Early postnatal development of the mandible in children with isolated cleft palate and children with nonsyndromic Robin sequence. Cleft Palate Craniofac J. 2006;43:160-167

Farrall M, Holder S. Familial recurrence-pattern analysis of cleft lip with or without cleft palate. Am J Hum Genet. 1992;50:270-277

Fogh-Andersen P. Inheritance of Harelip and Cleft Palate, Arnold Busck, Copenhagen, 1942

Gaspar DA, Matioli SR, de Cassia Pavanello R, Araujo BC, Alonso N, Wyszynski D, Passos-Bueno MR. Maternal MTHFR interacts with the offspring's BCL3 genotypes, but not with TGFA, in increasing risk to nonsyndromic cleft lip with or without cleft palate. Eur J Hum Genet. 2004;12:521-526

Hayes C. Environmental risk factors and oral clefts, p. 159-169, in Cleft Lip and Palate: from Origin to Treatment. Edited by Wyszinski DF, Oxford University Press, 2002

Hill-Harfe KL, Kaplan L, Stalker HJ, Zori RT, Pop R, Scherer G, Wallace MR. Fine mapping of chromosome 17 translocation breakpoints > or = 900 Kb upstream of SOX9 in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. Am J Hum Genet. 2005;76:663-671

Ichikawa E, Watanabe A, Nakano Y, Akita S, Hirano A, Kinoshita A, Kondo S, Kishino T, Uchiyama T, Niikawa N, Yoshiura K. PAX9 and TGFB3 are linked to susceptibility to nonsyndromic cleft lip with or without cleft palate in the Japanese: population-based and family-based candidate gene analyses. J Hum Genet. 2006;51:38-46

Ingraham CR, Kinoshita A, Kondo S, Yang B, Sajan S, Trout KJ, Malik MI, Dunnwald M, Goudy SL, Lovett M, Murray JC, Schutte BC. Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6). Nat Genet. 2006;38:1335-1340

Jin JZ, Ding J. Analysis of Meox-2 mutant mice reveals a novel postfusion-based cleft palate. Dev Dyn. 2006;235:539-546

Jezewski PA, Vieira AR, Nishimura C, Ludwig B, Johnson M, O'Brien SE, Daack-Hirsch S, Schultz RE, Weber A, Nepomucena B, Romitti PA, Christensen K, Orioli IM, Castilla EE, Machida J, Natsume N, Murray JC. Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. J Med Genet. 2003;40:399-407

Jugessur A, Murray JC. Orofacial clefting: recent insights into a complex trait. Curr Opin Genet Dev. 2005;15:270-278

Juriloff DM, Harris MJ, McMahon AP, Carroll TJ, Lidral AC. Wnt9b is the mutated gene involved in multifactorial nonsyndromic cleft lip with or without cleft palate in A/WySn mice, as confirmed by a genetic complementation test. Birth Defects Res A Clin Mol Teratol. 2006;76:574-579

Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet. 2005;76:8-32

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, de Lima RL, Daack-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, Lammer EJ, Aylsworth AS, Ardinger HH, Lidral AC, Pober BR, Moreno L, Arcos-Burgos M, Valencia C, Houdayer C, Bahuau M, Moretti-Ferreira D, Richieri-Costa A, Dixon MJ, Murray JC. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. Nat Genet. 2002;32:285-289

Lammer EJ, Shaw GM, Iovannisci DM, Finnell RH. Maternal smoking, genetic variation of glutathione s-transferases, and risk for orofacial clefts. Epidemiology. 2005;16:698-701

Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet. 1995;11:241-247

Leoyklang P, Siriwan P, Shotelersuk V. A mutation of the p63 gene in non-syndromic cleft lip. J Med Genet. 2006;43:e28

Lidral AC, Romitti PA, Basart AM, Doetschman T, Leysens NJ, Daack-Hirsch S, Semina EV, Johnson LR, Machida J, Burds A, Parnell TJ, Rubenstein JL, Murray JC. Association of MSX1 and TGFB3 with nonsyndromic clefting in humans. Am J Hum Genet. 1998;63:557-568

Little J, Cardy A, Munger RG. Tobacco smoking and oral clefts: a meta-analysis. Bull World Health Organ. 2004;82:213-218

Liu W, Sun X, Braut A, Mishina Y, Behringer RR, Mina M, Martin JF. Distinct functions for Bmp signaling in lip and palate fusion in mice. Development. 2005;132:1453-1461

Lu W, Volcik K, Zhu H, Wen S, Shaw GM, Lammer EJ, Finnell RH. Genetic variation in the proto-oncogene SKI and risk for orofacial clefting. Mol Genet Metab. 2005;86:412-416

Mansilla MA, Cooper ME, Goldstein T, Castilla EE, Lopez Camelo JS, Marazita ML, Murray JC. Contributions of PTCH gene variants to isolated cleft lip and palate. Cleft Palate Craniofac J. 2006;43:21-29

Marazita ML, Murray JC, Lidral AC, Arcos-Burgos M, Cooper ME, Goldstein T, Maher BS, Daack-Hirsch S, Schultz R, Mansilla MA, Field LL, Liu YE, Prescott N, Malcolm S, Winter R, Ray A, Moreno L, Valencia C, Neiswanger K, Wyszynski DF, Bailey-Wilson JE, Albacha-Hejazi H, Beaty TH, McIntosh I, Hetmanski JB, Tuncbilek G, Edwards M, Harkin L, Scott R, Roddick LG. Meta-

analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35. Am J Hum Genet. 2004;75:161-173

Marcano ACB, Doudney K, Braybrook C, Squires R, Patton MA, Lees MM, Richieri-Costa A, Lidral AC, Murray JC, Moore GE, Stanier P. TBX22 mutations are a frequent cause of cleft palate. J Med Genet. 2004;41:68-74

Mattick JS, Makunin IV. Non-coding RNA. Hum Mol Genet. 2006;15 Spec No 1:R17-29

Melkoniemi M, Koillinen H, Mannikko M, Warman ML, Pihlajamaa T, Kaariainen H, Rautio J, Hukki J, Stofko JA, Cisneros GJ, Krakow D, Cohn DH, Kere J, Ala-Kokko L. Collagen XI sequence variations in nonsyndromic cleft palate, Robin sequence and micrognathia. Eur J Hum Genet. 2003;11:265-270

Merla G, Howald C, Henrichsen CN, Lyle R, Wyss C, Zabot MT, Antonarakis SE, Reymond A. Submicroscopic deletion in patients with Williams-Beuren syndrome influences expression levels of the nonhemizygous flanking genes. Am J Hum Genet 2006;79:332-341

Mitchell LE, Christensen K. Analysis of the recurrence patterns for nonsyndromic cleft lip with or without cleft palate in the families of 3,073 Danish probands. Am J Med Genet. 1996;61:371-376

Mukhopadhyay P, Greene RM, Zacharias W, Weinrich MC, Singh S, Young WW Jr, Pisano MM. Developmental gene expression profiling of mammalian, fetal orofacial tissue. Birth Defects Res A Clin Mol Teratol. 2004;70:912-926.

Murray JC. Gene/environment causes of cleft lip and/or palate. Clin Genet. 2002;61:248-256

Park TJ, Boyd K, Curran T. Cardiovascular and craniofacial defects in Crk-null mice. Mol Cell Biol. 2006;26:6272-6282

Peanchitlertkajorn S, Cooper ME, Liu YE, Field LL, Marazita ML. Chromosome 17: gene mapping studies of cleft lip with or without cleft palate in Chinese families. Cleft Palate Craniofac J. 2003;40:71-79

Radhakrishna U, Ratnamala U, Gaines M, Beiraghi S, Hutchings D, Golla J, Husain SA, Gambhir PS, Sheth JJ, Sheth FJ, Chetan GK, Naveed M, Solanki JV, Patel UC, Master DC, Memon R, Antonarakis GS, Antonarakis SE, Nath SK. Genomewide scan for nonsyndromic cleft lip and palate in multigenerational Indian families reveals significant evidence of linkage at 13q33.1-34. Am J Hum Genet. 2006;79:580-585

Scapoli L, Palmieri A, Martinelli M, Vaccari C, Marchesini J, Pezzetti F, Baciliero U, Padula E, Carinci P, Carinci F. Study of the PVRL1 gene in Italian nonsyndromic cleft lip patients with or without cleft palate. Ann Hum Genet. 2006;70:410-413

Schliekelman P, Slatkin M. Multiplex relative risk and estimation of the number of loci underlying an inherited disease. Am J Hum Genet. 2002;71:1369-1385

Shaw GM, Lammer EJ. Maternal periconceptional alcohol consumption and risk for orofacial clefts. J Pediatr. 1999;134:298-303

Shaw GM, Iovannisci DM, Yang W, Finnell RH, Carmichael SL, Cheng S, Lammer EJ. Endothelial nitric oxide synthase (NOS3) genetic variants, maternal smoking, vitamin use, and risk of human orofacial clefts. Am J Epidemiol. 2005;162:1207-1214

Sozen MA, Suzuki, K.; Tolarova, M. M.; Bustos, T.; Fernandez Iglesias, J. E.; Spritz, R. A. Mutation of PVRL1 is associated with

sporadic, non-syndromic cleft lip/palate in northern Venezuela. Nature Genet. 2001;29:141-142

Stanier P, Moore GE. Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. Hum Mol Genet. 2004;13 Spec No 1:R73-81

Suzuki K, Hu D, Bustos T, Zlotogora J, Richieri-Costa A, Helms JA, Spritz RA. Mutations of PVRL1, encoding a cell-cell adhesion molecule/herpesvirus receptor, in cleft lip/palate-ectodermal dysplasia. Nat Genet. 2000;25:427-430

Thyagarajan T, Totey S, Danton MJ, Kulkarni AB. Genetically altered mouse models: the good, the bad, and the ugly. Crit Rev Oral Biol Med. 2003;14:154-174

Tongkobpetch S, Siriwan P, Shotelersuk V. MSX1 mutations contribute to nonsyndromic cleft lip in a Thai population. J Hum Genet. 2006;51:671-676

van den Boogaard M-JH, Dorland M, Beemer FA, Ploos van Amstel HK. MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. (Letter) Nature Genet. 2000;24:342-343

Velagaleti GV, Bien-Willner GA, Northup JK, Lockhart LH, Hawkins JC, Jalal SM, Withers M, Lupski JR, Stankiewicz P. Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of SOX9 in two patients with campomelic dysplasia. Am J Hum Genet. 2005;76:652-662

Vieira AR, Avila JR, Daack-Hirsch S, Dragan E, Felix TM, Rahimov F, Harrington J, Schultz RR, Watanabe Y, Johnson M, Fang J, O'Brien SE, Orioli IM, Castilla EE, Fitzpatrick DR, Jiang R, Marazita ML, Murray JC. Medical sequencing of candidate genes for nonsyndromic cleft lip and palate. PLoS Genet. 2005;1:e64

Warkany J, Nelson RC, Schraffenberger E. Congenital malformations induced in rats by maternal nutritional deficiency. IV. Cleft Palate. Am J Dis Child 1943:65:882–894

Warrington A, Vieira AR, Christensen K, Orioli IM, Castilla EE, Romitti PA, Murray JC. Genetic evidence for the role of loci at 19q13 in cleft lip and palate. J Med Genet. 2006;43:e26

Watanabe A, Akita S, Tin NT, Natsume N, Nakano Y, Niikawa N, Uchiyama T, Yoshiura K. A mutation in RYK is a genetic factor for nonsyndromic cleft lip and palate. Cleft Palate Craniofac J. 2006;43:310-316

Woolfe A, Goodson M, Goode DK, Snell P, McEwen GK, Vavouri T, Smith SF, North P, Callaway H, Kelly K, Walter K, Abnizova I, Gilks W, Edwards YJ, Cooke JE, Elgar G. Highly conserved non-coding sequences are associated with vertebrate development. PLoS Biol. 2005;3:e7

Zaritsky JJ, Eckman DM, Wellman GC, Nelson MT, Schwarz TL. Targeted disruption of Kir2.1 and Kir2.2 genes reveals the essential role of the inwardly rectifying K(+) current in K(+)-mediated vasodilation. Circ Res. 2000;87:160-166

Zucchero TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S, Schutte BC, Marazita ML, Murray JC. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. N Engl J Med. 2004;351:769-780

Paper I

The Genetic Basis of the Pierre Robin Sequence

Linda P. Jakobsen, M.D., Mary A. Knudsen, M.D., Ph.D., James Lespinasse, M.D., Ph.D., Carmen García Ayuso, M.D., Carmen Ramos, M.D., Jean-Pierre Fryns, Prof.M.D., Merete Bugge, M.D., Ph.D., Niels Tommerup, Prof.M.D., Ph.D.

Objective: **The Pierre Robin Sequence (PRS) is subgroup of the cleft palate population. As with the etiology of cleft lip or palate, the etiology of PRS is generally unknown. Some factors are suggestive of a genetic basis for PRS. The purpose of this study was to compare genetic information on PRS available in the literature and in a cytogenetic database to facilitate focused genetic studies of PRS.**

Design: **After searching Medline for "pierre robin and genetics," the Mendelian Cytogenetics Network database for "robin" and "pierre robin," and two reviews from the Human Cytogenetics Database for "cleft palate" and "micrognathia," a comparison of the data and a search in Online Mendelian Inheritance in Man (OMIM) Gene Map was performed to identify relevant candidate genes.**

Results: **The findings revealed consistency to a certain degree to loci 2q24.1-33.3, 4q32-qter, 11q21-23.1, and 17q21-24.3. A search in the OMIM Gene Map provided many candidate genes for PRS in these regions. The GAD67 on 2q31, the PVRL1 on 11q23-q24, and the SOX9 gene on 17q24.3-q25.1 are suggested to be of particular importance.**

Conclusion: **Candidate loci and a few potential candidate genes for PRS are proposed from the present study. This may enable researchers to focus their effort in the studies of PRS.**

KEY WORDS: *chromosomes, craniofacial anomalies, etiology, genetics, Pierre Robin Sequence*

The Pierre Robin Sequence (PRS) is a clinical entity characterized by cleft palate and micrognathia that results in glossoptosis, which was first described by Robin (1923) as the tongue tending to obstruct the airway and causing feeding and respiratory difficulties during the early postnatal period. Even though the clinical entity is well defined, the pathogenesis of PRS is debated. Formerly, it was named Pierre Robin syndrome, anomalad, or complex, but because it is regarded as a series of events (i.e., the micrognathia resulting in glossoptosis, which prevents the palatal shelves to fuse) it is now referred to as a sequence (Cocke, 1966; Latham, 1966; Poswillo, 1966; Ricks et al., 2002). Other experts regard it as a primary growth deficiency resulting in cleft palate and micrognathia (Rintala et al., 1984; Edwards and Newall, 1985) or causally heterogeneous (Cohen, 1999). Furthermore, it is unclear whether the respiratory and feeding difficulties are a result of the anatomy or of an intrinsic dysfunction within the pharynx (Cohen, 1999; Renault et al., 2000).

A high incidence of twinning in series of PRS is noted: 9% versus 1% in the general population. The twins are most often discordant for PRS, also when monozygotic (Holder-Espinasse et al., 2001). This could imply that PRS appears as a result of the twinning process. Moreover, the discordance in twins has raised a hypothesis of mechanical constraint *in utero*. However, no other evidence of constraint has been noted (Poswillo, 1966; Holder-Espinasse et al., 2001). Mitchell and Christensen (1997) evaluated family history in Danish twins and found no evidence of cleft lip and palate (CL/P) resulting from the twinning process.

The overall incidence of PRS is low: 1 per 8500 to 14,000 births in the general population (Bush and Williams, 1983;

Dr. Jakobsen is with the Clinic for Plastic and Reconstructive Surgery and Burns Unit, University Hospital of Copenhagen, Copenhagen, Denmark, and Wilhelm Johannsen Centre for Functional Genome Research, Institute of Medical Biochemistry and Genetics, Panum Institute 24.4., Copenhagen, Denmark. Dr. Knudsen is with the Clinic for Plastic and Reconstructive Surgery and Burns Unit, University Hospital of Copenhagen, Copenhagen, Denmark. Dr. Lespinasse is with Cytogenetic Laboratory, General Hospital, Chambery, France. Drs. García Ayuso and Ramos are with the Department of Genetics, Fundación Jimenez-Diaz, Madrid, Spain. Dr. Fryns is with the Genetics Department, Center for Human Genetics, University Hospital of Leuven, Belgium. Drs. Bugge and Tommerup are with Wilhelm Johannsen Centre for Functional Genome Research, Institute of Medical Biochemistry and Genetics, Panum Institute 24.4., Copenhagen, Denmark.

Submitted January 2005; Accepted June 2005.

Address correspondence to: Linda P. Jakobsen, Wilhelm Johannsen Centre for Functional Genome Research, Institute of Medical Biochemistry and Genetics, Panum Institute 24.4., Blegdamsvej 3, 2200 Copenhagen N, Denmark. E-mail linda@imbg.ku.dk.

Printzlau and Andersen, 2004). The prevalence of PRS is the same in both sexes (Printzlau and Andersen, 2004).

Some factors point to a genetic etiology of PRS. Patients with PRS often have other family members with cleft lip or palate (13.0%–27.7%) (Bixler and Christian, 1971; Williams et al., 1981; Marques et al., 1998; Holder-Espinasse et al., 2001), and PRS is often present in other syndromes such as Stickler syndrome, Velocardiofacial syndrome, Marshall syndrome, Treacher Collins syndrome, Catel-Mancke syndrome, Kabuki syndrome, Nager syndrome, teratogene syndromes, and many more (Sheffield et al., 1987; Shprintzen, 1988; Marques et al., 1998; Cohen, 1999; Holder-Espinasse et al., 2001). The most common (12.0%–16.3%) underlying syndrome is Stickler syndrome (Sheffield et al., 1987; Marques et al., 1998; Printzlau and Andersen, 2004). When PRS is not associated with other malformations, it is referred to as nonsyndromic PRS.

To be able to identify candidate genes for PRS, a search of Medline, the Mendelian Cytogenetics Network database (MCNdb), and two reviews of the Human Cytogenetics Database (HCDB) (Brewer et al., 1998, 1999) was performed.

Searching for PRS Genes

Examining CL/P subgroups may provide us with knowledge that can be extended to the general CL/P population. An example is Zucchero et al. (2004), who found the IRF6 gene causing Van der Woude syndrome. Subsequently, they tested for linkage and transmission disequilibrium in a total of 1968 CL/P families and found that a polymorphism in IRF6 is responsible for 12% of the genetic contribution in CL/P.

Searching Medline (http://www4.ncbi.nlm.nih.gov/entrez/query.fcgi) for ''pierre robin and genetics'' revealed 20 articles that included genetic and clinical information on the cases. Deletions; duplications; translocations; and mutations in chromosomes 1 to 6, 10 to 13, and 16 to 18 were involved. Regions in chromosome 2 (2q24.1-33.3), chromosome 4 (4q32-qter), chromosome 11 (11q21-q23.1), and chromosome 17 (17q21-q24.3) were involved between three and six times; the other chromosomes were involved only once (Table 1). One of the translocations (t(16;17)(p13;q21)) has been included in the table even although it was the proband's father who had the translocation. The father is clinically normal, but the proband has the karyotype 46,XX and presents clinically as a male (Petrus et al., 1981). This inclusion was done to avoid leaving out any potentially important information.

One of the authors (N.T.) has initiated and now coordinates the MCNdb, where information on chromosomal aberrations and diseases or malformations is gathered. More than 300 cytogenetic laboratories cooperate and contribute to MCNdb. A search in MCNdb (http://www.mcndb.org/) on ''robin'' and ''pierre robin'' revealed six cases (Table 2). One case was already published (Jamshidi et al., 2004) and therefore was omitted from MCNdb to avoid cases occurring from both the MCNdb and the Medline searches. Chromosome 2 (2q32 and 2q33) and 17 (17q21 and 17q24) were involved twice; other chromosomal regions were involved only once.

Brewer et al. (1998, 1999) reviewed the HCDB in the search for chromosome regions associated with malformations. The HCDB is a commercially available computerized catalog of postnatally ascertained cytologically detectable human chromosomal aberrations. Through statistical analysis, the authors found cleft palate to be significantly associated ($p < .05$) with deletions in 2q32, 4p16-13, and 4q31-35 (Brewer et al., 1998) and duplications in 3p24-23, 3p26, 3q23-25, 7q22-32, 8q21, 10p15-11, 14q11-21, 16p12-13, and 22q12-13 (Brewer et al., 1999).

Micrognathia was found to be significantly associated ($p < .05$) with deletions in 4p16-14, 4q31-35, 6q25-27, and 11q23 (Brewer et al., 1998) and duplications in 10q24 and 18q12-23 (Brewer et al., 1999). No cytogenetic information on PRS is available in the work of Brewer et al. (1998, 1999).

Despite the limited material and that the number of PRS cases may be an underestimate because they were not registered if described as having, for instance, ''cleft palate and small chin'' in Medline or in the MCNdb, the findings from Medline, MCNdb, and data from Brewer et al. (1998, 1999) point to some interesting loci for PRS at 2q24.1-33.3, 4q32-qter, 11q21-23.1, and 17q21-24.3.

A search in these regions in the Online Mendelian Inheritance of Man (OMIM) Gene Map (http://www.ncbi.nlm.nih.gov/omim/) revealed a large number of potential candidate genes. Only a few of the hundreds of genes in these regions have been shown to be involved in causing cleft palate in transgene mice (GAD67, IDH1, ITGAV, DLX2, PDGFC, PVRL1, SOX9) and even fewer in humans (GAD67, PVRL1, and SOX9). GAD67 (or GAD1) is the gene coding for glutamate decarboxylase, an enzyme that converts glutamic acid to gamma-amino butyric acid (GABA). GABA is the major inhibitory neurotransmitter in the vertebral central nervous system. GAD67 is mainly found in the brain but also in the kidney. On the basis of the frequency distribution of a GAD67 haplotype, Kanno et al. (2004) suggested that GAD67 on chromosome 2q31 is involved in the pathogenesis of CL/P in the Japanese population. The GAD67 gene is not necessarily the correct or the only gene in this chromosomal region causing cleft palate. FitzPatrick et al. (2003) found cytogenetic evidence of the SATB2 gene as the cleft palate gene on 2q32-q33. This gene is involved in transcriptional control and is expressed in the developing palate. Despite these facts, no mutations in the SATB2 gene were found in 70 patients with cleft palate.

PVRL1 is a member of the Ig superfamily and a receptor for Herpes viruses. Sozen et al. (2001) found that a heterozygote mutation in the gene coding for PVRL1 on chromosome 11q23-q24 is a moderate genetic risk factor for nonsyndromic CL/P in northern Venezuela.

Besides GAD67 and PVRL1, the transcription factor SOX9 could be an interesting gene. Mutations in the SOX9 gene and approximately 1 Mb upstream the SOX9 gene, in an apparently gene-empty region, cause campomelic dysplasia (a rare skeletal dysplasia) and autosomal sex reversal. The gene-empty region is believed to be a regulatory SOX9 control region.

TABLE 1 Medline Search Results for "Pierre Robin and Genetics"*

Case No.	*Genetic Findings*	*Selected Clinical Findings*	*Authors*
Deletions			
1	Deletion 2q33.1–33.3 *de novo*	PRS, PMR, hypotonia, thin skin and hair, abnormal dentition	Vogels et al. (1997)
2	Deletion 2q32.3–q33.2 mat.	PRS and minor facial abnormalities. The mother had an unbalanced translocation and was unaffected.	Houdayer et al. (2001)
3	Deletion 4q25-q27 (unknown origin)	PRS, PMR, pulmonary stenosis, contractures, scoliosis, clinodactyly	Chudley et al. (1988)
4	Deletion 4q31-qter *de novo*	PRS, PMR, depressed nasal bridge, clinodactyly	Davis et al. (1981)
5	Deletion 4q32-4qter (unknown origin)	PRS and PMR	Fryns et al. (1981)
6	Deletion 4q33-qter *de novo*	PRS, short left forearm, hip dysplasia	Menko et al. (1992)
7	Deletion 4q32.1-qter *de novo*	PRS, cardiac and pulmonary malformations, death in infancy	Rethore et al. (1979)
8	Deletion 4q33 *de novo*	PRS, clinodactyly, PFO	Strehle et al. (2001)
9a	Deletion 11q21-q23 *de novo*	PRS, PMR, renal and cardiac malformations	De Lonlay-Debeney et al. (1998)
10	Deletion 16q12.1-q13 *de novo*	PRS, PMR, hypertelorism, dysplastic hips, fusion of the third to sixth vertebrae of the neck	Schuffenhauer et al. (1992)
Duplications			
11	Duplication 1(q12-q25) *de novo*	PRS, hydrocephalus, PDA, pulmonary hypoplasia, cryptorchidism, equinovarus	Chen et al. (1994)
12	Duplication 1(q23.1-q31.1) *de novo*	PRS, brachycephaly, flexion of fingers in both hands and camptodactyly	Aboura et al. (2002)
Translocations			
9b	Unbalanced translocation der(22)t(11;22) (q23;q11.2) mat.	PRS, renal and cardiac malformations, perinatal death. The mother had a balanced translocation and had had several spontaneous abortions	De Lonlay-Debeney et al. (1998)
13	Unbalanced translocation der(4)t(4;11) (q35;q23.1) mat.	PRS, dislocated hips, ASD, agenesis of corpus callosum and of one kidney, death in infancy	Francke et al. (1977)
14	Unbalanced translocation der(10),t(3;10) (q23;q26.3) *de novo*	PRS, hypertelorism, cryptorchidism, clinodactyly, hammer toes, death in infancy	Kleczkowska et al. (1988)
15	Unbalanced translocation der(5),t(5;11) (p15;q23) pat.	PRS, PMR, ASD and VSD, rockerbottom feet. The father had a balanced translocation and was unaffected.	Wallerstein et al. (1992)
16	Balanced translocation t(16;17)(p13;q21) pat.	The father had a translocation and was unaffected. The proband had a normal karyotype but was XX male (sex reversal) and had PRS, PMR, and scoliosis.	Petrus et al. (1981)
17	Balanced translocation t(13,17) (q22.1;q23.3)	PRS, skeletal defects and multiple miscarriages in a five-generation family (seven affected); PrS cosegregates with the translocation	Stalker et al. (2001)
18	Balanced translocation t(2;17) (q24.1;q24.3)	PRS in a three-generation family (six affected); PRS cosegregates with the translocation	Jamshidi et al. (2004)
Mutations			
19	Mutations in COL2A1(12q13.11-13.2), COL11A2 (6p21.3), and COL11A1 (1p21)	Mutation screening of collagene types in 23 patients with PRS. Mutations were found in seven patients.	Melkoniemi et al. (2003)
Isochromosomes			
20	46, XY,i(18q)	PRS, pulmonary hypoplasia, cryptorchidism, flexion deformities of fingers	Wiswell and Edwards (1986)

* Only articles with genetic and clinical information were included. Case 9 is listed twice because the authors report two cases (one with a deletion and one with a translocation). PMR = psychomotor retardation; PFO = patent foramen ovale; PDA = patent ductus arteriosus; ASD = atrial septal defects; VSD = ventricular septal defects. Explanation of the unbalanced translocations: der(n) is the chromosome missing material; the other chromosome has partial trisomy.

Patients with breakpoints outside the coding region are generally less severely affected than are patients with mutations in the SOX9 coding region (Wagner et al., 1994; Pfeifer et al., 1999). Pierre Robin Sequence is often part of the phenotype in campomelic dysplasia. Mansour et al. (2002) studied five patients with campomelic dysplasia and found cleft palate in two patients, micrognathia in five patients, and respiratory difficulties in four patients.

SOX9 has been shown to regulate collagen expression during cartilage and endochondral bone formation (Bell et al., 1997; Mori-Akiyama et al., 2003). Mutations in genes coding for collagen types have been found in patients with PRS (Table 1 and Melkoniemi et al., 2003), indicating a regulatory pathway involving SOX9 and genes coding for collagen, whose disruption can cause PRS. Furthermore, in a newly published paper, Velagaleti et al. (2005) suggest that PRS may result from dysregulation of SOX9.

CONCLUSION

Pierre Robin Sequence is a clinical well-characterized subgroup of CL/P. Some factors point to PRS as having at least

TABLE 2 MCNdb Search Results for "Robin" and "Pierre Robin"*

Case No.	*Genetic Findings*	*Selected Clinical Features*	*Contributors*
1	t(1;3)9q42;q12), t(3;1)(q22;q42), ins(4;3)(q31;q12q22)	PRS, Dandy-Walker malformation, agenesia of cerebellum and corpus callosum PMR, shunt ASD/VSD/PDA. Parental status unknown	Prof. J.-P. Fryns, Leuven
2	t(5;14)(q31;q12), inv(9)(p11;q12)X2 unknown	PRS, shunt ASD/VSD/PDA, pulmonic atresia, stillbirth. Parental status unknown	Dr. J. Lespinasse, Chambery
3	t(2;11)(q33;p14) *de novo*	PRS, PMR, microstoma, arachno- and clinodactyly. Parents unaffected	Dr. P.M. Ellis, Edinburgh†
4	t(2;17)(q32;q24) pat.	PRS. The father also had a translocation and PRS	Dr. R. Ayuso, Madrid
5	t(7;18)(q31.3;q22.2) mat.	PRS, death in infancy. Parental status unknown	Prof. T. Haaf, Mainz
6	t(7;17)9p11.1;q21) *de novo*	PRS, stillbirth, parents unaffected	Dr. C. Lundsteen, Copenhagen

* PMR = psychomotor retardation; ASD = atrial septal defects; VSD = ventricular septal defects; PDA = patent ductus arteriosus.
† Case 3 has been published as: Brewer CM, Leek JP, Green AJ, Holloway S, Bonthron DT, Markham AF, FitzPatrick DR. A locus for isolated cleft palate, located on human chromosome 2q32. *Am J Hum Genet*. 1999; 65:387–396.

partially a genetic basis. A comparison among cases in the literature and cases in cytogenetic databases revealed consistency to a certain degree to loci 2q24.1-33.3, 4q32-qter, 11q21-23.1, and 17q21-24.3. No particular candidate genes can be identified for certain from the present study, but GAD67 on 2q31, PVRL1 on 11q23-q24, and the SOX9 gene on 17q24.3-q25.1 are suggested to be important.

Future studies involving cytogenetic analyses and mutation analyses of candidate genes in nonsyndromic PRS are needed. Efforts should focus on the genomic regions and genes mentioned above. Furthermore, mutation analyses of the genes in populations with nonsyndromic CL/P may provide us with information on the genetic contribution of these genes in the general CL/P population.

Acknowledgments. We thank Dr. Lundsteen, Dr. Ellis, and Professor Haaf for making this work possible by contributing with some of the cases. This work was partly supported by the Augustinus Foundation, the Vanfoere Foundation, Doctor Johan Boserup and Lise Boserups Foundation, Aase and Ejnar Danielsens Foundation, Jacob Madsen and Olga Madsens Foundation, and Queen Louise Childhospital Research Foundation. Wilhelm Johannsen Centre for Functional Genome Research is established by the Danish National Research Foundation.

REFERENCES

Aboura A, Coulomb-LHermine A, Audibert F, Capron F, Frydman R, Tachdjian G. De novo interstitial direct duplication 1(q23.1 q31.1) in a fetus with Pierre Robin sequence and camptodactyly. *Am J Med Genet*. 2002;108:153–159.

Bell DM, Leung KKH, Wheatley SC, Ng LJ, Zhou S, Ling KW, Sham MH, Koopman P, Tam PPL, Cheah KSE. SOX9 directly regulates the type-II collagen gene. *Nat Genet*. 1997,16.174–178.

Bixler D, Christian JC. Pierre Robin syndrome occurring in two related sibships. *Birth Defects Orig Artic Ser*. 1971:7:67–71.

Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal deletion map of human malformations. *Am J Hum Genet*. 1998;63: 1153–1159.

Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal duplication map of malformations: regions of suspected haplo- and triplolethality—and tolerance of segmental aneuploidy—in humans. *Am J Hum Genet*. 1999;64:1702–1708.

Bush PG, Williams AJ. Incidence of the Robin anomalad (Pierre Robin syndrome). *Br J Plast Surg*. 1983;36:434–437.

Chen H, Kusyk CJ, Tuck-Muller CM, Martincz JE, Dorand RD, Wertelecki W. Confirmation of proximal 1q duplication using fluorescence in situ hybridization. *Am J Med Genet*. 1994;50:28–31.

Chudley AE, Verma MR, Ray M, Riordan D. Interstitial deletion of the long arm of chromosome 4. *Am J Med Genet*. 1988;31:549–551

Cocke WJ Jr. Experimental production of micrognathia and glossoptosis associated with cleft palate (Pierre Robin Syndrome). *Plast Reconstr Surg*. 1966; 38:395–403.

Cohen MM Jr. Robin sequences and complexes: causal heterogeneity and pathogenetic/phenotypic variability. *Am J Med Genet*. 1999;84:311–315.

Davis JM, Clarren SK, Salk DJ. Brief clinical report: the del(4) (q31) syndrome—a recognizable disorder with atypical Robin malformation sequence. *Am J Med Genet*. 1981;9:113–117.

De Lonlay-Debeney P, de Blois M-C, Bonnet D, Amiel J, Abadie V, Picq M, Lyonnet S, Sidi D, Munnich A, Vekemans M, Cormier-Daire V. Ebstein anomaly associated with rearrangements of chromosomal region 11q. *Am J Med Genet*. 1998;80:157–159.

Edwards JRC, Newall DR. The Pierre Robin Syndrome reassessed in the light of recent research. *Br J Plast Surg*. 1985;38:339–342.

FitzPatrick DR, Carr IM, McLaren L, Leek JP, Wightman P, Williamson K, Gautier P, McGill N, Hayward C, Firth H, Markham AF, Fantes JA, Bonthron DT. Identification of SATB2 as the cleft palate gene on 2q32-q33. *Hum Mol Genet*. 2003;12:2491–2501.

Francke U, Weber F, Sparkes RS, Mattson PD, Mann J. Duplication 11 (q21 to 23–>qter) syndrome. *Birth Defects Orig Artic Ser*. 1977;13:167–186.

Fryns JP, Timmermans J, Hoedemaekers J, Emmery L, Van den Berghe H. Pierre Robin anomalad, moderate mental retardation and distal 4q deletion. *Ann Genet*. 1981;24:187–188.

Holder-Espinasse M, Abadie V, Cormier-Daire V, Beyler C, Manach Y, Munnich A, Lyonnet S, Couly G, Amiel J. Pierre Robin sequence: a series of 117 consecutive cases. *J Pediatr*. 2001;139:588–590.

Houdayer C, Portnoi M-F, Vialard F, Soupre V, Crumière C, Taillemite J-L, Couderc R, Vasquez M-P, Bahuau M. Pierre Robin sequence and interstitial deletion 2q32.3-q33.2. *Am J Med Genet*. 2001;102:219–226.

Jamshidi N, Macciocca I, Dargaville PA, Thomas P, Kilpatrick N, McKinlay Gardner RJ, Farlie PG. Isolated Robin sequence associated with a balanced t(2;17) chromosomal translocation. *J Med Genet*. 2004;41:e1.

Kanno K, Suzuki Y, Yamada A, Aoki Y, Kure S, Matsubara Y. Association between nonsyndromic cleft lip with or without cleft palate and the glutamic acid decarboxylase 67 gene in the Japanese population. *Am J Med Genet*. 2004;127A:11–16.

Kleczkowska A, Fryns JP, Moerman F, Martens M, Eggermont E, Jaeken J, Van den Berghe H. Trisomy 3q2 and Pierre Robin sequence in a boy with unbalanced 46, XY, der(10), t(3;10)(q23;q26.3) de novo karyotype. *Helv Paediatr Acta*. 1988;43:245–248.

Latham RA. The pathogenesis of cleft palate associated with the Pierre Robin syndrome. An analysis of a seventeen-week human foetus. *Br J Plast Surg*. 1966;19:205–214.

Mansour S, Offiah AC, McDowall S, Sim P, Tolmie J, Hall C. The phenotype of survivors of campomelic dysplasia. *J Med Genet*. 2002;39:597–602.

Marques IL, Barbieri MA, Bettiol H. Etiopathogenesis of isolated Robin Sequence. *Cleft Palate Craniofac J*. 1998;35:517–525.

Melkoniemi M, Koillinen H, Mannikko M, Warman ML, Pihlajamaa T, Kaa-

riainen H, Rautio J, Hukki J, Stofko JA, Cisneros GJ, Krakow D, Cohn DH, Kere J, Alabama-Kokko L. Collagen XI sequence variations in nonsyndromic cleft palate, Robin sequence and micrognathia. *Eur J Hum Genet*. 2003;11:265–270.

Menko FH, Madan K, Baart JA, Beukenhorst HL. Robin sequence and a deficiency of the left forearm in a girl with a deletion of chromosome 4q33-qter. *Am J Med Genet*. 1992;44:696–698.

Mitchell LE, Christensen K. Evaluation of family history data for Danish twins with nonsyndromic cleft lip with or without cleft palate. *Am J Med Genet*. 1991;72:120–121.

Mori-Akiyama Y, Akiyama H, Rowitch DH, de Crombrugghe B. Sox9 is required for determination of the chondrogenic cell lineage in the cranial neural crest. *Proc Natl Acad Sci U S A*. 2003;100:9360–9365.

Petrus M, Bourrouillou G, Dutau G, Colombies P, Rochiccioli P. Association of male XX with Pierre Robin syndrome in a child whose father has a balanced 46XY, t(16;17)(p13;q21) translocation. *J Genet Hum*. 1981;29:191–196.

Pfeifer D, Kist R, Dewar K, Devon K, Lander ES, Birren B, Korniszewski L, Back E, Scherer G. Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to SOX9: evidence for an extended control region. *Am J Hum Genet*. 1999;65:111–124.

Poswillo D. Observations of fetal posture and causal mechanisms of congenital deformity of palate, mandible and limbs. *J Dent Res*. 1966;45(suppl 3):584–596.

Printzlau A, Andersen M. Pierre Robin sequence in Denmark: a retrospective population-based epidemiological study. *Cleft Palate Craniofac J*. 2004;41:47–52.

Renault F, Flores-Guevara R, Soupre V, Vasquez M-P, Baudon J-J. Neurophysiological brainstem investigations in isolated Pierre Robin sequence. *Early Hum Dev*. 2000;58:141–152.

Rethore M-O, Couturier J, Mselati JC, Cochois B, Lavaud J, Lejeune J. De novo monosomy 4q32.1–>4qter in a newborn with multiple malformations. *Ann Genet*. 1979;22:214–216.

Ricks JE, Ryder VM, Bridgewater LC, Schaalje B, Seegmiller RE. Altered mandibular development precedes the time of palate closure in mice homozygous for disproportionate micromelia: an oral clefting model supporting the Pierre Robin sequence. *Teratology*. 2002;65:116–120.

Rintala A, Ranta R, Stegars T. On the pathogenesis of cleft palate in the Pierre Robin syndrome. *Scand J Plast Reconstr Surg*. 1984;18:237–240.

Robin P. Backward lowering of the root of the tongue causing respiratory disturbances. *Bull Acad Natl Med*. 1923;89:37–41.

Schuffenhauer S, Callen DF, Seidel H, Shen Y, Lederer G, Murken J. De novo interstitial deletion 16(q12.1q13) of paternal origin in a 10-year-old boy. *Clin Genet*. 1992;42:246–250.

Sheffield LJ, Reiss JA, Stuolun K, Gilding M. A genetic follow-up study of 64 patients with the Pierre Robin complex. *Am J Med Genet*. 1987;28:25–36.

Shprintzen RJ. Pierre Robin, micrognathia, and airway obstruction: the dependency of treatment on accurate diagnosis. *Int Anesthesiol Clin*. 1988;26:64–71.

Sozen MA, Suzuki K, Tolarova MM, Bustos T, Fernandez-Iglesias JE, Spritz RA. Mutation of PVRL1 is associated with sporadic, non-syndromic cleft lip/palate in northern Venezuela. *Nat Genet*. 2001;29:141–142.

Stalker HJ, Gray BA, Zori RT. Dominant transmission of a previously unidentified 13/17 translocation in a five-generation family with Robin cleft and other skeletal defects. *Am J Med Genet*. 2001;103:339–341.

Strehle EM, Ahmed OA, Hameed M, Russell A. The 4q- syndrome. *Genet Couns*. 2001;12:327–339.

Velagaleti GV, Bien-Willner GA, Northup JK, Lockhart LH, Hawkins JC, Jalal SM, Withers M, Lupski JR, Stankiewicz P. Position effects due to chromosome breakpoints that map ~900 Kb upstream and ~1.3 Mb downstream of SOX9 in two patients with Campomelic Dysplasia. *Am J Hum Genet*. 2005;76:652–662.

Vogels A, Haegeman J, Fryns JP. Pierre-Robin sequence and severe mental retardation with chaotic behavior associated with a small interstitial deletion in the long arm of chromosome 2 (del(2)(q331q333)). *Genet Couns*. 1997;8:249–252.

Wagner T, Wirth J, Meyer J, Zabel B, Held M, Zimmer J, Pasantes J, Dagna Bricarelli F, Keutel J, Hustert E, Wolf U, Tommerup N, Schempp W, Scherer G. Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell*. 1994;79:1111–1120.

Wallerstein R, Desposito F, Aviv H, Schenk M, Wallerstein DF. Partial trisomy 11q in a female infant with Robin sequence and congenital heart disease. *Cleft Palate Craniofac J*. 1992;29:11–19.

Williams AJ, Williams MA, Walker CA, Bush PG. The Robin anomalad (Pierre Robin syndrome)—a follow up study. *Arch Dis Child*. 1981;56:663–668.

Wiswell TE, Edwards RG. Presentation of the isochromosome trisomy 18 syndrome in an infant with the Robin anomalad. *Hawaii Med J*. 1986;45:126–127.

Zucchero TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S, Schutte BC, Marazita ML, Murray JC. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *N Engl J Med*. 2004;351:769–780.

Paper II

Electronic letter to Journal of Medical Genetics

Title: **Pierre Robin Sequence may be caused by dysregulation of *SOX9***
Running title: **Pierre Robin and *SOX9***

Authors:
Linda P. Jakobsen, M.D.[1, 2]
Reinhard Ullmann M.Sc., Ph.D.[3]
Steen B. Christensen, M.D., Ph.D.[4]
Karl Erik Jensen, M.D., Ph.D.[5]
Kirsten Mølsted, D.D.S., Ph.D.[6]
Karen F. Henriksen[1]
Claus Hansen, M.Sc.[1]
Mary A. Knudsen, M.D., Ph.D.[2]
Lars A. Larsen, M.Sc., Ph.D.[1]
Niels Tommerup, M.D., Ph.D.[1]
Zeynep Tümer, M.D., Ph.D.[1]

1) Wilhelm Johannsen Centre for Functional Genome Research, Department of Medical Biochemistry and Genetics, Panum Institute 24.4., Copenhagen, Denmark.
2) Department of Plastic- and Reconstructive Surgery and Burns Unit, Rigshospitalet, University Hospital of Copenhagen, Denmark.
3) Department of Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Berlin, Germany.
4) Department of Orthopaedic Surgery, Section of Paediatric Orthopaedics, Rigshospitalet, University Hospital of Copenhagen, Denmark.
5) Department of Diagnostic Radiology, Rigshospitalet, University Hospital of Copenhagen, Denmark.
6) Copenhagen Cleft Palate Centre, Denmark.

Wordcount: 2767
Keywords: Campomelic Dysplasia; cleft lip and palate; *KCNJ2*; Pierre Robin Sequence; *SOX9*

Recommended reviewers: Dr. Stankiewicz (pawels@bmc.tmc.edu) and Dr. Scherer (scherer@humangenetik.ukl.uni-freiburg.de).

Correspondance to:
Linda P. Jakobsen, Wilhelm Johannsen Centre for Functional Genome Research, Department of Medical Biochemistry and Genetics, Panum Institute 24.4., Blegdamsvej 3, 2200 Copenhagen N, Denmark.
Tel: +45 35327836
Fax: +45 35327845
Email: linda@imbg.ku.dk

Abstract

Background: The Pierre Robin Sequence (PRS), consisting of cleft palate, micrognathia and glossoptosis, can be seen as part of the phenotype in other Mendelian syndromes, for instance Campomelic Dysplasia (CD) which is caused by *SOX9* mutations, but the aetiology of non-syndromic PRS has not been unravelled yet.

Objective: To get more insight into the aetiology of PRS by studying patients with PRS using genetic and cytogenetic methods.

Methods: Ten unrelated patients with PRS were investigated by chromosome analyses and BAC arrays. A balanced translocation was found in one patient and the breakpoints were mapped with fluorescence in situ hybridisation and southern blot analysis. All patients were screened for mutations in the *SOX9* gene, and in five of the patients *SOX9* expression analysis was carried out by quantitative PCR.

Results: We identified an abnormal balanced karyotype 46,XX, t(2;17)(q23.3;q24.3) in one PRS patient and mapped the 17q breakpoint to 1.13 Mb upstream of the transcription factor *SOX9*. This is the most distantly located upstream breakpoint reported to date that conveys an abnormal phenotype by a position effect. A subsequent skeletal X-ray survey revealed very mild features of CD in this patient. Radiological examinations were carried out in three other PRS patients, and a mild vertebral dysplasia was observed in one of them. Furthermore, reduced *SOX9* mRNA expression was observed in PRS patients.

Conclusion: Our findings suggest that non-syndromic PRS may be caused by *SOX9* dysregulation. Thus, PRS may represent the mildest form of a continuum of skeletal dysplasias caused by dysregulation of *SOX9*, where CD is the most severe form.

Abbreviations: CD, Campomelic Dysplasia; array-CGH, Comparative Genome Hybridisation; FISH, fluorescence in situ hybridisation; MCNdb, Mendelian Cytogenetics Network database; miRNA, microRNA; PRS, Pierre Robin Sequence; QPCR, quantitative real-time PCR; SNP, single nucleotide polymorphism; UTR, untranslated region

Keywords: Campomelic Dysplasia; cleft lip and palate; *KCNJ2*; Pierre Robin Sequence; *SOX9*

Introduction

Cleft lip and/ or palate (CL/P) is a common congenital malformation affecting approximately 2 per 1000 newborns worldwide. The aetiology of CL/P is largely unknown, however recent studies focusing on syndromic forms of CL/P have identified specific genes that may also be involved in non-syndromic CL/P.[1] The Pierre Robin Sequence (PRS [MIM 261800]) is a clinically well defined subgroup of the CL/P population with an unknown aetiology. PRS is characterised by cleft palate, micrognathia and respiratory difficulties in the early neonatal period (caused by glossoptosis), and is often observed as a part of other Mendelian syndromes, such as Stickler syndrome, Velocardiofacial syndrome and Marshall syndrome.[2] PRS is also seen as a part of Campomelic Dysplasia (CD [MIM 114290]), a rare skeletal dysplasia, consisting of bowing of the long bones (campomelia), malformation of the pelvis and spine, 11 pairs of ribs, hypoplastic scapulae, club feet, micrognathia, cleft palate, and in some patients male-to-female sex reversal. When severely affected, the children die in the neonatal period due to respiratory problems resulting from developmental defects in the respiratory system.[3] CD is due to haploinsufficiency of the transcription factor *SOX9* on 17q24.3, either caused by intragenic mutations, deletions involving the entire gene, or chromosomal rearrangements up to ~950 kb upstream of the *SOX9* coding region.[4-6] The phenotypes of the breakpoint cases in CD, believed to be caused by disruption of regulatory regions, are generally less severe than the intragenic mutations.[6]

Description of patients with skeletal abnormalities, including PRS features, caused by breakpoints 900 kb and 932 kb upstream of the *SOX9* coding region[7, 8] and the finding of a familial translocation t(2;17) co-segregating with PRS[9] suggest *SOX9* as a candidate gene also for non-syndromic PRS. Animal studies support this hypothesis, as *SOX9* mutant mice present with cleft palate and hypoplasia of cartilage derived skeletal structures[10] and *SOX9* is expressed in the fetal mouse mandible.[11]

In order to gain more insight into the aetiology of PRS, we performed genetic and cytogenetic analysis in ten unrelated patients with non-syndromic PRS.

Methods

The patients included in this study were identified through hospital records and selected if they had non-syndromic PRS (cleft palate, micrognathia and respiratory difficulties in the early neonatal period due to glossoptosis). One patient had an additional history of malignant hyperthermia and bilateral inguinal hernias. The patients were between two and 16 years old.

Chromosome analysis was carried out with the G-banding technique using peripheral blood lymphocytes. Fine mapping of the breakpoints was performed using flourescence in situ hybridisation (FISH), with BAC and fosmid clones obtained from the Sanger Institute. Genomic DNA was isolated from blood according to the standard protocols.

For Southern blot analysis the genomic DNA was digested with the restriction enzymes *Eco*RI and *Msc*I. Two hybridisation probes (SB817 and SB520) were prepared by PCR using primer sets SB817U/L (5'-GTAAAGAAACGGTCGCAAATA-3' / 5'-CAATAGCAGAGCGCAGTAG-3') and SB520U/L

(TGGGACAGAAACATTACCTTG-3' / 5'-TGATTGGAGGAGAAAACGAC-3'). The SB817 probe hybridised to the 5136 bp *Msc*I and the 7172 bp *Eco*RI fragments, and SB520 detected the 7102 bp *Msc*I and 3132 bp *Eco*RI fragments.

Comparative Genome Hybridisation (array-CGH) was carried out using a sub-megabase resolution whole genome tiling path BAC array consisting of the human genome high resolution 32 k re-arrayed clone set (BACPAC Resources), the 1 Mb Sanger set (Wellcome Trust Sanger Institute) and a set of 390 sub-telomeric clones (assembled by members of the COST B19 initiative: Molecular Cytogenetics of solid tumours).[12]

SOX9 and *KCNJ2* expression in cultured lymphoblastoid cell lines were analysed by quantitative real-time RT-PCR (QPCR) using an Opticon DNA engine RT-PCR machine (MJ Research, Waltham, MA) using SYBR green kit (GE Healthcare, Waukesha, WI). Relative gene expression was calculated using dilution curves and *UBC*, *YWHAZ, GAPDH, G6PD* and *B2M* mRNA as reference genes. The experiments were run in duplicates and repeated twice. Genomic PCR and sequencing were performed according to standard protocols (see supplementary table 1 for primer sequences and PCR conditions, mRNA isolation, and cDNA preparation).

The study was approved by the local scientific ethics committee and written informed consent was obtained from the patients and from the parents of children who were younger than 12 years old.

Results and Discussion

We performed chromosome analyses in ten unrelated patients with features of PRS. All the cases had normal karyotypes except one where PRS co-segregated with a balanced translocation t(2;17)(q23.3;q24.3) in a father and a daughter.

The translocation carrier (fig 1A and 1B) was a 15-year-old female. The father also had PRS according to hospital records, but he declined to participate in further clinical and genetic studies. The patient had mild facial dysmorphism; flat face, broad nasal bridge, low set ears and low set hairline. She was born as the only child of a 33-year-old father, and a 41-year-old mother with type 2 diabetes. Although the intellectual and psychomotor development was in the lower end of the normal spectrum, she attended a standard school. She had received speech therapy and a pharyngeal flap operation had been performed in order to decrease hypernasal speech. No visual or hearing deficits were reported. Growth and sexual maturation were normal. The mother has a healthy daughter from an earlier relationship. No other family member has PRS, cleft palate or other malformations. Material from paternal grandparents was not available, but a paternal uncle and a maternal half-sister of the father had normal karyotypes. After the translocation was identified a skeletal survey of the patient was performed, and it showed micrognathia, hypoplastic scapulae and depressed iliac wings, features also found in CD (fig 2A-C).

Array-CGH was performed in seven of the nine non-translocation patients with PRS, and it showed a 0.4 Mb microdeletion at 7q21.13 (chromosome position 88,473,693-88,909,911, UCSC Genome Browser, March 2006 assembly) in one of the patients, which was confirmed by FISH. Parental DNA was not

available to investigate the origin of this deletion. There is no known copy-number variation within this region (Database of Genomic Variants and unpublished observations by one of the authors, RU) and there is only one gene, *ZNF804B*, which encodes a C2H2-type zinc finger protein. Deletions at 7q21-22 have been associated previously with PRS, but several other malformations such as ectrodactyly were part of the clinical phenotype.[13] The other six patients did not show any significant changes with array-CGH.

Translocation breakpoints of the t(2;17) patient were mapped using FISH. The chromosome 2q23.3 breakpoint was within the BAC clone RP11-373H2 (AC113610) at chromosome position 150,777,641-150,874,211 (UCSC Genome Browser). This breakpoint is located in a gene empty region, approximately 200 kb downstream of the *RND3* gene and 350 kb upstream of *FJL32955* encoding a hypothetical protein (LOC150596). *RND3* encodes a plasma membrane bound GTPase involved in regulation of cellular response and cytoskeletal dynamics. *RND3* could be involved in the aetiology in the present case, but no phenotype is associated with mutations or chromosomal rearrangements in the vicinity of *RND3*, and a position effect in the present case is speculative. Chromosome abnormalities involving 2q22 and 2q24 have been observed previously in patients with PRS. Jamshidi and colleagues (2004) reported a familial translocation t(2;17)(q24.1;q24.3) co-segregating with PRS,[9] and mutations of *ZFHX1B* at 2q22.3 cause Mowat-Wilson syndrome [MIM 235730], in which submucous cleft palate is a part of the phenotype, although this syndrome does not otherwise resemble CD or PRS. However, these regions are located far from the present breakpoint and do not have an apparent effect on the phenotype of the patient.

We mapped the 17q24.3 breakpoint within an approximately 15 kb overlapping region of the BAC clone RP11-7D6 (fig 4A) and two fosmid clones (G248P89352E7 and G248P86963D5). The breakpoint was further mapped with Southern blot hybridisation using two probes detecting an approximately 12 kb overlapping region within the overlapping region of the FISH clones (fig 4B and 4C). Using the hybridisation probe SB817 and *Msc*I digestion, we detected the 5.1 kb normal fragment and a slightly larger band representing the junction fragment in the patient. This probe also detected the junction fragment using the restriction enzyme *Eco*RI. Hybridisation probe SB520 showed only the normal band pattern in the patient (data not shown). This suggests that the 17q breakpoint is within an approximately 5 kb region defined by two *Msc*I recognition sites (fig 4C), 1.13 Mb upstream of the *SOX9* gene (chromosome position 66,485,382-66,491,517). The total gene empty region upstream of *SOX9* is approximately 2 Mb where the proximal flanking gene is *KCNJ2*, which is located ~800 kb from the 17q breakpoint. The gene empty region harbours several breakpoints found in CD patients (fig 3 shows a schematic overview of breakpoints located >900 kb upstream of *SOX9*).

We sequenced the coding region of *SOX9* in all the ten PRS patients and detected several nucleotide polymorphisms, but no pathogenic mutations. We also screened the 3´untranslated region (3´-UTR) of *SOX9* for sequence variations, as *in silico* predictions of microRNA (miRNA) target sites suggest that *SOX9* expression is regulated by miRNAs (miRBase Targets Version 2.0). miRNAs are short non-coding RNAs involved in down regulation of genes at the translational level by binding to specific target sites in the 3´-UTR of the mRNA.[14] No sequence variations in the predicted miRNA target sites were identified.

Furthermore, we searched for evolutionary conserved sequences, and hence potentially regulating regions (in human, chimpanzee, mouse, rat, dog, chicken, zebrafish and fugu, UCSC Genome Browser) and identified a conserved region at chromosome position 66,520,880-66,521,660. This region overlaps with the candidate regulatory element "SOX9cre1" (chromosome position 66,520,446-66,521,622, UCSC Genome Browser) proposed by Velagaleti and colleagues (2005)[8] (fig 3). Screening this conserved region in the ten PRS patients revealed two unreported consecutive single nucleotide polymorphisms (SNPs) at 66,520,950- 66,520,951 (A/T and G/A) in two patients. These SNPs were also detected in two out of 18 patients with isolated cleft palate and in 20 out of 95 normal controls. We therefore concluded, that the two consecutive SNPs were not associated to PRS.

Haploinsufficiency of *SOX9* as a cause of CD has been shown by deletions involving the entire gene in two unrelated CD patients,[4, 5] and we considered whether expression of *SOX9* was dysregulated in PRS patients. By QPCR, we detected a reduced *SOX9* expression in cultured lymphoblastoid cell lines from five non-translocation PRS patients compared to 11 controls (Supplementary table 2). Unexpectedly, the expression of *SOX9* in the translocation patient was higher than in the group of non-translocation PRS patients, although still in the lower end of the normal range. A possible explanation to this is that some upstream regulators (<1 Mb from *SOX9* in the present case) are still active in the translocation patient. Likewise, Wirth and colleagues (1996)[15] found the same expression levels in both *SOX9* alleles in a lymphoblastoid cell line from a t(13;17) translocation patient, where the 17q breakpoint was located more than 130 kb from *SOX9*. The expression data should be interpreted with caution, as expression in the patients may not reflect expression during embryonic development. Moreover, expression studies were carried out in lymphoblastoid cell lines, not in chondrocytes, which would have been preferred, but were unavailable.

As a consequence of the apparently reduced expression of *SOX9* in the PRS patients, we carried out X-ray examination (of thorax, scapulae, cervical spine, patellae and hands) in three PRS patients available, and in one patient we detected slightly reduced vertical height of the cervical vertebrae C3-C6 (fig 2D). This feature, although very mild, is a part of the phenotypic features of CD. Whether this finding is correlated to the reduced expression of *SOX9* is unknown.

The breakpoint located most upstream within the *SOX9* regulatory landscape reported so far, ~932 kb from the coding region was associated with hypoplastic scapulae and 11 pairs of ribs and Robin sequence (family F).[7] The 17q breakpoint identified in the present study was located 1.13 Mb upstream of *SOX9*, which is 200 kb further upstream than the breakpoint in family F, and the absence of missing ribs in the present patient point to a milder phenotype. Thus, the present case constitutes the mildest phenotype and the most distantly located breakpoint upstream of the *SOX9* gene reported to date. Whether the mild phenotypes may be categorised as CD or represent a distinct clinical entity is a subject of debate.[16, 17] Considering the growing reports of distantly located chromosomal rearrangements co-segregating with milder types of skeletal dysplasia, it is possible that CD, acampomelic CD, mild acampomelic CD and PRS may represent a continuum of phenotypes caused by dysregulation of *SOX9*.

Our sequencing results show that mutations in the *SOX9* coding regions, in *SOX9* microRNA target sites and in an evolutionary conserved upstream region are not a frequent cause of PRS. Identifying and sequencing other candidate *cis*-regulatory regions is highly relevant but complicated, time- and cost-consuming, as regulatory elements may reside even within neighbouring genes.[18] Since translocations upstream of *SOX9* are causing phenotypes consistent with CD, *cis*-regulatory elements have been sought for in the conserved neighbouring regions of *SOX9* and many have been suggested (reviewed by Hill-Harfe et al.).[7] A ~1.1 kb candidate regulatory element "SOX9cre1" was suggested using *in silico* analysis and analysis of chromosomal rearrangements (fig 3).[8] The breakpoint in the present case is located even more upstream than the SOX9cre1, why one or more of the *cis*-regulatory regions must be located either in our breakpoint region or even more upstream, or both. Another explanation of the effect of the translocation could be that altered organisation of the chromatin structure causes changes in the gene regulation (e.g., euchromatic to heterochromatic conformation).[18]

In the region upstream of *SOX9* several non-coding transcripts, but no protein coding transcripts have been identified.[7] The 17q breakpoint is located in the intronic region of a 226 kb spliced human EST (BM678241) which has been detected in fetal and adult eye.[19] BM678241 is located proximal to BC039327 (fig 3), a human mRNA exclusively expressed in testis and disrupted by the 17q breakpoint in family F. It was suggested that BC039327 was a functional RNA, with an unknown role, since there was no testis phenotype in the family F.[7] Likewise, the role of BM678241 in relation to our breakpoint is unclear, as our translocation carrier does not have a phenotype involving the eyes.

The closest proximal gene to the 17q breakpoint is the gene *KCNJ2,* which encodes a potassium channel*,* may be of interest in relation to the PRS phenotype. In humans, *KCNJ2* mutations have been detected in Andersens syndrome [MIM 170390], which presents with potassium sensitive periodic paralysis, variable cardiac arrhythmias and dysmorphic features overlapping with CD to some extent, including low set ears, micrognathia, cleft palate and scoliosis. Moreover, mutations in *Kcnj2* cause cleft palate in mice.[20] To our knowledge, CD with periodic paralysis and cardiac arrhythmias has not been reported, but PRS occurring with cardiac arrhythmias has been described previously[21] and a higher prevalence of congenital cardiac anomalies has been found in PRS patients (13.6%) compared to the general cleft lip and palate population (6.7%).[22] We have sequenced the translated region of *KCNJ2* in the ten PRS patients including the translocation patient, without finding any mutations. Furthermore, we investigated the *KCNJ2* expression in five PRS patients and the translocation patient using QPCR analyses (supplementary table 3). The results showed an overall reduction in the *KCNJ2* expression in the patients compared to the controls. The expression levels of *KCNJ2* were very low both in the controls and patients thus, two different *KCNJ2* primer sets were tested, and the results were in agreement (supplementary table 3).

Our data may suggest, that *KCNJ2* may play a role in the PRS phenotype together with *SOX9*, but further studies are needed to determine this. It has recently been reported though, that chromosomal

aberrations may change the expression of several neighbouring genes,[23] thus supporting a role for both *KCNJ2* and *SOX9* in PRS patients.

Considering gene regulation, it is also necessary to include genes involved in the *SOX9* signaling pathways as *SOX9* is one of several factors involved in chondrogenesis.[24] Mutations and sequence variations in genes encoding collagen subunits (*COL2A1*, *COL11A1* and *COL11A2*) were found in patients with non-syndromic PRS,[25] supporting that PRS may result from dysregulation of *SOX9* signaling pathways in general.

In summary, our data suggest that dysregulation of *SOX9* may be involved in PRS, evidenced by a familial translocation with a breakpoint located 1.13 Mb upstream of *SOX9*, by reduced expression of *SOX9* in non-translocated PRS patients, and by the detection of discrete radiologic signs of skeletal dysplasia in one of these patients. To determine the extent of the involvement of *SOX9* and *KCNJ2* in PRS, skeletal surveys and sequencing of candidate *cis*-regulatory regions and signaling partners in a larger group of patients are needed.

Acknowledgments

We are grateful to the patients and their families for their participation in the study. The study was approved by the Local Scientific Ethics Committee and informed consent was obtained from all the patients and/or their families. Written consent has been obtained from the patient and the mother for the publication of the photographs. We wish to thank David FitzPatrick and Judy Fantes for helpful suggestions and assistance in FISH mapping. We also thank Lillian Rasmussen, Anita Niebuhr, Elisabeth Larsen, Minna Becker and Marei Schubert for skilful technical assistance. DNA for the Array-CGH set was kindly provided by Pieter de Jong and the clones for the 1 Mb Sanger set were kindly provided by Nigel Carter. The work was partly supported by grants from the Augustinus Foundation, the Vanfoere Foundation, Doctor Johan Boserup and Lise Boserups Foundation, Aase and Ejnar Danielsens Foundation, Jacob Madsen and wife Olga Madsens Foundation and The Research Foundation of the Queen Louise Childrens Hospital. Wilhelm Johannsen Centre for Functional Genome Research was established by the Danish National Research Foundation.

Authors´affiliations

Linda P. Jakobsen, Wilhelm Johannsen Centre for Functional Genome Research, Department of Medical Biochemistry and Genetics, Panum Institute and Department of Plastic- and Reconstructive Surgery and Burns Unit, Rigshospitalet, University Hospital of Copenhagen, Denmark.

Reinhard Ullmann, Department of Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Berlin, Germany.

Steen B. Christensen, Department of Orthopaedic Surgery, Section of Paediatric Orthopaedics, Rigshospitalet, University Hospital of Copenhagen, Denmark.

Karl Erik Jensen, Department of Diagnostic Radiology, Rigshospitalet, University Hospital of Copenhagen, Denmark.

Mary A. Knudsen, Department of Plastic- and Reconstructive Surgery and Burns Unit, Rigshospitalet, University Hospital of Copenhagen, Denmark.

Kirsten Mølsted, Copenhagen Cleft Palate Centre, Denmark.

Zeynep Tümer, **Lars A. Larsen, Niels Tommerup, Claus Hansen and Karen F. Henriksen**, Wilhelm Johannsen Centre for Functional Genome Research, Department of Medical Biochemistry and Genetics, Panum Institute, Copenhagen, Denmark.

Conflict of interest: none declared.

Electronic-Database Information

The URLs for data presented herein are as follows:

Database of Genomic Variants, http://projects.tcag.ca/variation/

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/

UCSC Genome Bioinformatics site, March 2006 assembly, http://genome.ucsc.edu/

The Mendelian Cytogenetics Network Database (MCNdb), http://www.mcndb.org/

NCBI dbSNP Build 125, http://www.ncbi.nlm.nih.gov/SNP/

miRBase Targets Version 2.0, http://microrna.sanger.ac.uk/targets/v2/

Sanger Institute, http://sanger.ac.uk/targets/

Figure 1A

Figure 1B

Figure 1A-B Photographs of the patient with Pierre Robin Sequence and the balanced translocation t(2;17)(q23.3;q24.3). The patient has consented to have her pictures published in a scientific paper.

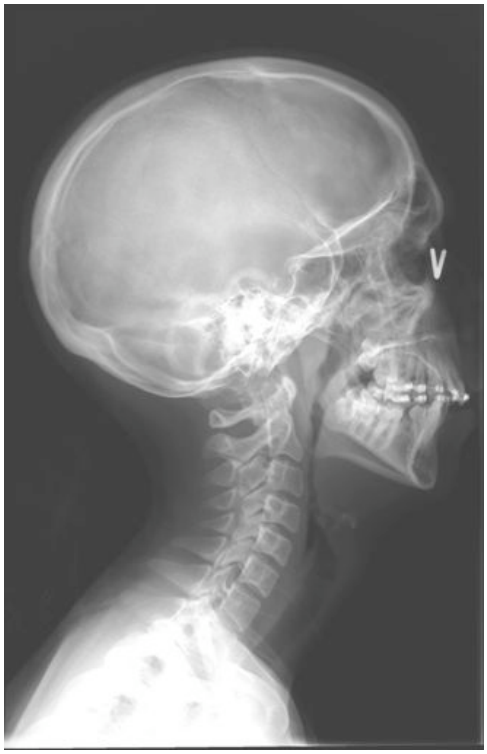

Figure 2A

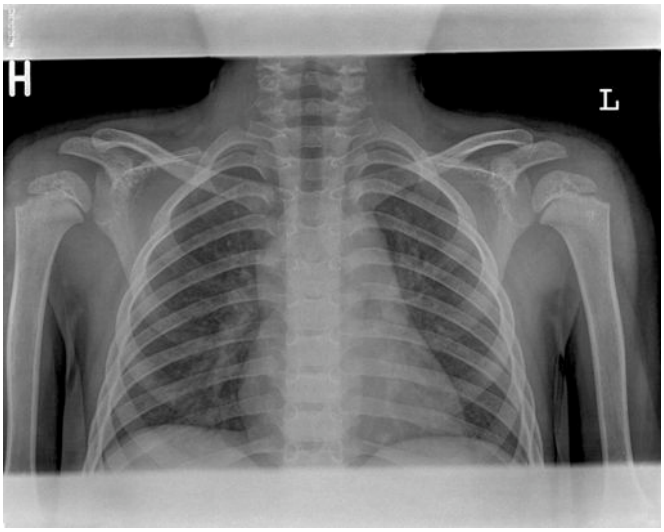

Figure 2B

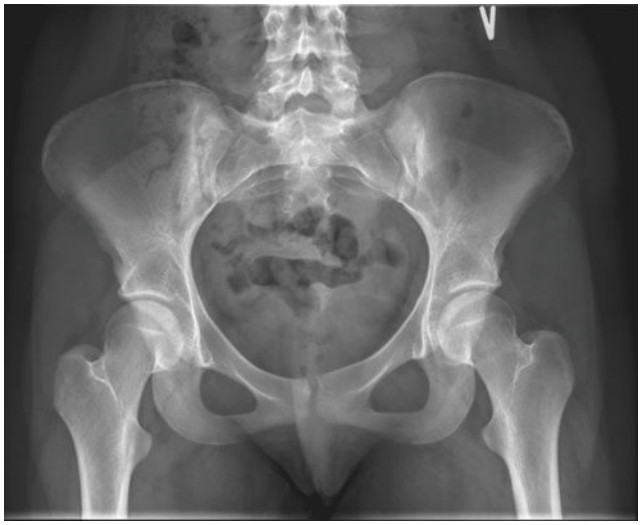

Figure 2C

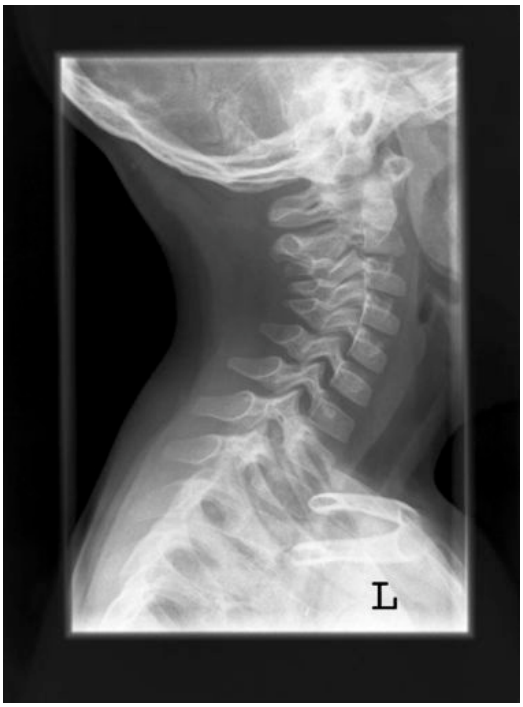

Figure 2D

Figure 2A-D X-ray examination of the translocation carrier showing **A**. micrognathia **B**. hypoplastic scapulae and **C**. reduced vertical height of the iliac bones. **D**. The cervical spine of one of the non-translocation PRS patients with reduced *SOX9* and *KCNJ2* expression, showing reduced height of the cervical vertebrae C3-C6.

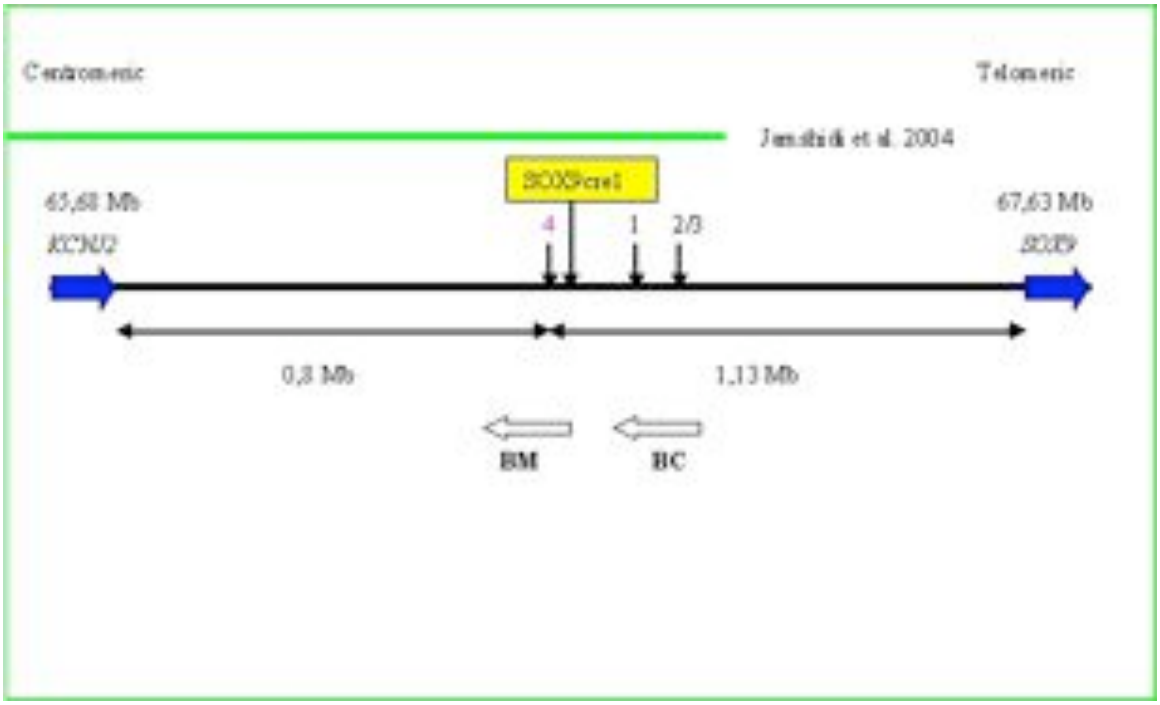

Figure 3 Schematic overview of the 2 Mb gene empty region upstream of *SOX9,* between *KCNJ2* and *SOX9*, showing known breakpoints more than 900 kb upstream of *SOX9*. Patients 1 and 2 were reported by Hill-Harfe et al. (2005)[7] and patient 3 by Velagaleti et al. (2005).[8] The patients have breakpoints in the following approximate chromosome positions: Patient 1 at 66.7 Mb, patients 2 and 3 at 66.73 Mb. Patient 4 is the translocation patient presented here with a breakpoint at 66.49 Mb. The green bar represents the breakpoint interval in the three-generation family with PRS reported by Jamshidi et al. (2004).[9] Velagaleti et al. (2005) [8] suggested a potential regulatory region, SOX9cre1, at 66.52 Mb. BM is the spliced human EST BM678241 and BC is the human mRNA BC039327.

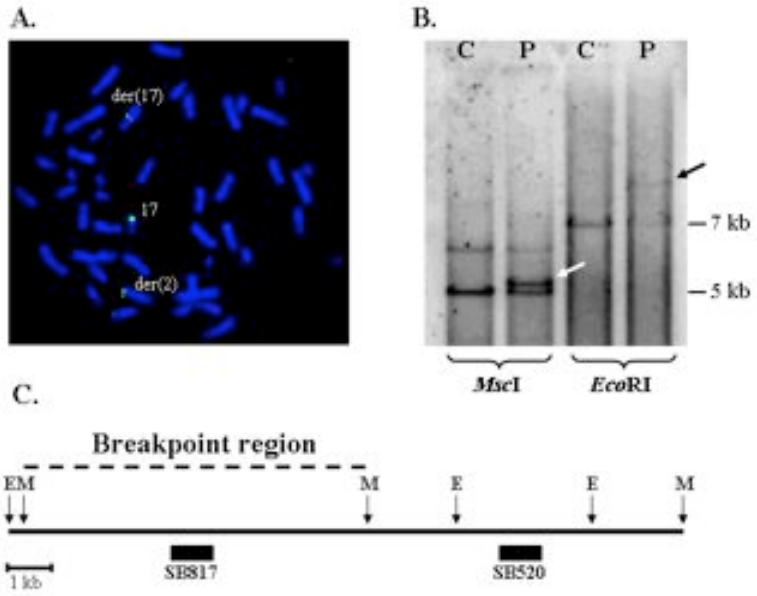

Figure 4 The breakpoint region at 17q. **A.** FISH analysis showing the breakpoint spanning BAC clone RP11-7D6. **B.** Southern blot hybridisation using probe SB817 and the restriction enzymes *Msc*I and *Eco*RI. With *Msc*I digests the 5 kb normal fragment and a larger band representing the junction fragment (white arrow) is detected in the patient. With *Eco*RI digests besides the 7 kb normal fragment the patient has a larger fragment (black arrow). **C.** The physical map of the breakpoint region. The 12 kb interval is within the 15 kb overlapping region of the spanning BAC (fig 4**A**) and the fosmid clones. The breakpoint region (dashed line) is mapped between the two *Msc*I restriction sites (chromosome position 66,485,382-66,491,517). The vertical arrows show the *Msc*I (M) and *Eco*RI (E) restriction sites. The filled boxes represent the two hybridisation probes SB817 and SB520. P, patient; C, control.

References

1. Jugessur A, Murray JC. Orofacial clefting: recent insights into a complex trait. *Curr Opin Genet Dev* 2005;**15**:270-278

2. Holder-Espinasse M, Abadie V, Cormier-Daire V, Beyler C, Manach Y, Munnich A, Lyonnet S, Couly G, Amiel J. Pierre Robin sequence: a series of 117 consecutive cases. *J Pediatr* 2001;**139**:588-590

3. Mansour S, Hall CM, Pembrey ME, Young ID. A clinical and genetic study of campomelic dysplasia. *J Med Genet* 1995;**32**:415-420

4. Pop R, Conz C, Lindenberg KS, Blesson S, Schmalenberger B, Briault S, Pfeifer D, Scherer G. Screening of the 1 Mb SOX9 5' control region by array CGH identifies a large deletion in a case of campomelic dysplasia with XY sex reversal. *J Med Genet* 2004;**41**:e47

5. Olney PN, Kean LS, Graham D, Elsas LJ, May KM. Campomelic syndrome and deletion of SOX9. *Am J Med Genet* 1999;**84**:20-24

6. Pfeifer D, Kist R, Dewar K, Devon K, Lander ES, Birren B, Korniszewski L, Back E, Scherer G. Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to SOX9: evidence for an extended control region. *Am J Hum Genet* 1999;**65**:111-124

7. Hill-Harfe KL, Kaplan L, Stalker HJ, Zori RT, Pop R, Scherer G, Wallace MR. Fine mapping of chromosome 17 translocation breakpoints > or = 900 Kb upstream of SOX9 in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. *Am J Hum Genet* 2005;**76**:663-671

8. Velagaleti GV, Bien-Willner GA, Northup JK, Lockhart LH, Hawkins JC, Jalal SM, Withers M, Lupski JR, Stankiewicz P. Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of SOX9 in two patients with campomelic dysplasia. *Am J Hum Genet* 2005;**76**:652-662

9. Jamshidi N, Macciocca I, Dargaville PA, Thomas P, Kilpatrick N, McKinlay Gardner RJ, Farlie PG. Isolated Robin sequence associated with a balanced t(2;17) chromosomal translocation. *J Med Genet* 2004;**41**:e1

10. Bi W, Huang W, Whitworth DJ, Deng JM, Zhang Z, Behringer RR, de Crombrugghe B. Haploinsufficiency of Sox9 results in defective cartilage primordia and premature skeletal mineralization. *Proc Natl Acad Sci U S A* 2001;**98**:6698-6703

11. Nie X. Sox9 mRNA expression in the developing palate and craniofacial muscles and skeletons. *Acta Odontol Scand* 2006;**64**:97-103

12. Erdogan F, Chen W, Kirchhoff M, Kalscheuer VM, Hultschig C, Müller I, Schulz R, Menzel C, Bryndorf T, Ropers H-H, Ullmann R. 2006. Impact of low copy repeats on the generation of balanced and unbalanced chromosomal aberrations in mental retardation. CGR 2006;**115:**247-253

13. Nunes ME, Pagon RA, Disteche CJ, Evans JP. A contiguous gene deletion syndrome at 7q21-q22 and implications for a relationship between isolated ectrodactyly and syndromic ectrodactyly. *Clin Dysmorphol* 1994;**3**:277-286

14. Berezikov E, Plasterk RH. Camels and zebrafish, viruses and cancer: a microRNA update. *Hum Mol Genet* 2005;**14**:R183-R190

15. Wirth J, Wagner T, Meyer J, Pfeiffer RA, Tietze HU, Schempp W, Scherer G. Translocation breakpoints in three patients with campomelic dysplasia and autosomal sex reversal map more than 130 kb from SOX9. *Hum Genet* 1996;**97**:186-193

16. Unger S. The mildest form of campomelic dysplasia. *Am J Med Genet A* 2005;**132**:113

17. Stalker HJ, Zori RT, Wallace M, Hill-Harfe KL, Kaplan L. Reply to Unger: the mildest form of campomelic dysplasia. *Am J Med Genet A* 2005;**132**;114-115

18. Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet* 2005;**76**:8-32

19. Bonaldo MF, Lennon G, Soares MB. Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res* 1996;**6**:791-806

20. Zaritsky JJ, Eckman DM, Wellman GC, Nelson MT, Schwarz TL. Targeted disruption of Kir2.1 and Kir2.2 genes reveals the essential role of the inwardly rectifying K(+) current in K(+)-mediated vasodilation. *Circ Res* 2000;**87**:160-166

21. Stoll C, Kieny JR, Dott B, Alembik Y, Finck S. Ventricular extrasystoles with syncopal episodes, perodactyly, and Robin in sequence in three generations: a new inherited MCA syndrome? *Am J Med Genet* 1992;**42**:480-486

22. Geis N, Seto B, Bartoshesky L, Lewis MB, Pashayan HM. The prevalence of congenital heart disease among the population of a metropolitan cleft lip and palate clinic. *Cleft Palate J* 1981;**18**:19-23

23. Merla G, Howald C, Henrichsen CN, Lyle R, Wyss C, Zabot MT, Antonarakis SE, Reymond A. Submicroscopic deletion in patients with Williams-Beuren syndrome influences expression levels of the nonhemizygous flanking genes. *Am J Hum Genet* 2006;**79**:332-341

24. de Crombrugghe B, Lefebvre V, Behringer RR, Bi W, Murakami S, Huang W. Transcriptional mechanisms of chondrocyte differentiation. *Matrix Biol* 2000;**19**:389-394

25. Melkoniemi M, Koillinen H, Mannikko M, Warman ML, Pihlajamaa T, Kaariainen H, Rautio J, Hukki J, Stofko JA, Cisneros GJ, Krakow D, Cohn DH, Kere J, Ala-Kokko L. Collagen XI sequence variations in nonsyndromic cleft palate, Robin sequence and micrognathia. *Eur J Hum Genet* 2003;**11**:265-270

Supplementary information

	Forward primer 5'-3'	Reverse primer 5'-3'	Annealing temp (°C)
SOX9			
Exon 1	GCTTCTCGCCTTTCCCGGCC	GGGGCAAATCAGCCCTGACCAG	62
Exon 2	CGACCTGACAGTTTGGCGGAT	GGTGTGCCAGGCGGGACG	60
Exon 3A	CCCGGAGGGTGCCTAAGACTA	AGCGTGGTCAGCGTGTGC	65
Exon 3B	CACCTACACGGGCAGCTACGG	GCTGTGGGTCTGCGGGATG	65
Exon 3C	CCCATCACCCGCTCAC	TAATGCGCTTGGATAGGTCAT	60
Mir-101 and Mir-145	AAATGGAGCAGCGAAATC	GGGCACACTGTTCAACTAAG	60
Mir-9 and Mir-124	GCTTTTCTTGCAACCAGAGTATTT	TGATAAAGCTTACCAAATGCTTCTCT	53
Mir-124 and Mir-124a	GAAAAACACCTTGAGCCTTA	GGAAAGCTCCAACAACCTA	60
Mir-124a and Mir-145	TTTGTATTCCTCACCCTAGA	CCCCTCCAGGTAGCC	56
Mir-9, Mir-30 and Mir-155	AACCTTTTGTTCTCTCCGTG	AATTTGCCAAATCATCCAAC	55
Upstream conserved region of *SOX9**	GGGAACTGTCAAAGGTGGATT	CATCCAGCCCCCTTATTTTT	53
KCNJ2			
Exon 1	CAAGGCTCCCAGAGACACCCATC	CTACAACATCTATGTGCCATCGGGTCA	65
***SOX9* mRNA (QPCR)**	GAGAGCGAGGAGGACAAGTTC	CCCGTTCTTCACCGACTT	60
***KCNJ2* (QPCR) A**	CACTGGATCTTACATGCCTCTGTACCC	AGACGATGCTGTAGCGGTTGGTT	65
***KCNJ2* (QPCR) B**	CACTGGATCTTACATGCCTCTG	TTTGAATGTTCGGTGAAGACAC	65
***UBC* (QPCR)**	ATTTGGGTCGCGGTTCTTG	TGCCTTGACATTCTCGATGGT	65
***YWHAZ* (QPCR)**	AGCGAGATCCAGGGACAGAGTCTCA	CAAGATGACCTACGGGCTCCTACAACA	66
***B2M* (QPCR)**	TGTGCTCGCGCTACTCTCTC	CTGAATGCTCCACTTTTTCAATTCT	60
***GAPDH* (QPCR)**	GGAAGGTGAAGGTCGGAGTCAA	GATCTCGCTCCTGGAAGATGGT	60
***G6PD* (QPCR)**	CCGGGACAACATCGCCTGCGTTATC	ACGGCTGCAAAAGTGGCGGTGGT	60

Table 1 Primer sequences and annealing temperatures. * Position 66,520,880-66,521,660, UCSC Genome Browser, March 2006 assembly**.**

QPCR: Total RNA was isolated from cell lines with TRIzol Reagent (Invitrogen) and cDNA synthesized with SuperScript II (RNase H^-) reverse transcriptase (Invitrogen) according to manufacturer's instructions. QPCR analysis was carried out on a DNA Engine Opticon 2 (Biorad, Göteborg, SE) using LightCycler FastStart DNA Master^{PLUS} SYBR GreenI (Roche, Hvidovre, DK). Primer sequences used for real-time PCR analysis are shown above. Primers were purchased from TAG Copenhagen (Copenhagen, DK). All primers are intron-spanning and designed using Oligo6 software (Cascade, CO). The identity of all PCR products and primer sequence specificity was confirmed by BLAST analysis (NCBI, NIH, http://www.ncbi.nlm.nih.gov/BLAST/) and melting curve analysis.

Sample	Gender	Expression values *SOX9*
PRS 1	XY	5.4
PRS 2	XY	5.4
PRS 3	XX	1.7
PRS 4	XY	5.3
PRS 5	XX	7.1
PRS mean (range)		5.0 (1.7-7.1)
PRS 6 t(2;17)	XX	13.7
C 1	XY	10.6
C 2	XX	34.6
C 3	XY	10.4
C 4	XY	15.0
C 5	XX	22.4
C 6	XY	40.2
C 7	XX	22.5
C 8	XX	1.0
C 9	XY	21.4
C 10	XX	4.0
C 11	XY	31.9
C mean (range)		19.5 (1.0-40.2)

Table 2 *SOX9* expression in lymphoblasts (QPCR). The expression value of *SOX9* was normalised to the mean values of the housekeeping genes *UBC, YWHAZ, B2M, G6PD* and *GAPDH*. Mean values were calculated according to the method described in **Pfaffl MW**, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations. Biotechnol Lett. 2004;**26**:509-15.
The experiments were run in duplicates and repeated twice.
PRS 1-6: six patients with Pierre Robin Sequence, PRS 6 is the translocation patient. C 1-11: controls.

Sample	Gender	Expression *KCNJ2* A	Expression *KCNJ2 B*
PRS 1	XY	2.5	1.2
PRS 2	XY	3.0	1.0
PRS 3	XX	2.0	1.6
PRS 4	XY	2.7	3.0
PRS 5	XX	1.0	2.1
PRS mean (range)		2.2 (1.0-3.0)	1.8 (1.0-3.0)
PRS 6 t(2;17)	XX	3.7	2.5
C 1	XY	8.2	2.1
C 2	XX	70.3	23.7
C 3	XY	2.4	1.2
C 4	XY	10.8	4.3
C 5	XX	31.2	16.3
C 6	XY	17.1	5.6
C 7	XX	35.2	20.0
C 8	XX	6.7	1.1
C 9	XY	27.6	30.4
C 10	XX	7.6	2.3
C 11	XY	70.3	29.1
C mean (range)		26.1 (2.4-70.3)	12.3 (1.1-30.4)

Table 3 *KCNJ2* expression in lymphoblasts (QPCR). As the expression of *KCNJ2* was low, QPCR were performed with two different primer sets, A and B (Table 1). The expression value of *KCNJ2* was normalised to the mean values of the housekeeping genes *UBC*, *YWHAZ, B2M, G6PD and GAPDH*. Mean values were calculated according to the method described in **Pfaffl MW**, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations. Biotechnol Lett. 2004;**26**:509-15.
The experiments were run in duplicates and repeated twice.
PRS 1-6: six patients with Pierre Robin Sequence, PRS 6 is the translocation patient. C 1-11: controls.

Paper III

Original article to Journal of Medical Genetics

Title: **Expression analyses of cleft lip and palate tissue point to a role for osteopontin and immune related factors in cleft lip and palate**
Running title: **Gene expression in cleft lip and palate tissue**

Authors:
Linda P. Jakobsen, M.D.[1, 2]
Rehannah Borup, M.Sc., Ph.D. [3]
Janni Vestergaard, M.Sc.[1]
Lars A. Larsen, M.Sc., Ph.D. [1]
Kasper Lage, M.Sc.[4]
Lisa Leth Maroun, M.D.[5]
Carsten U. Niemann, M.D., Ph.D. [6]
Mikael Andersen, M.D.[2]
Mary A. Knudsen, M.D., Ph.D.[2]
Kjeld Møllgård, M.D., Ph.D.[7]
Niels Tommerup, M.D., Ph.D.[1]

1) Wilhelm Johannsen Centre for Functional Genome Research, Institute of Medical Biochemistry and Genetics, Panum Institute, Copenhagen, Denmark.
2) Department of Plastic- and Reconstructive Surgery and Burns Unit, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark.
3) Department of Clinical Biochemistry, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark.
4) Center for Biological Sequence Analysis (CBS), BioCentrum-DTU, Technical University of Denmark.
5) Department of Pathology, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark.
6) Department of Haematology, Granulocytlaboratoriet, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark.
7) Institute of Medical Anatomy, Panum Institute, Copenhagen, Denmark.

Correspondance to:
Linda P. Jakobsen, Wilhelm Johannsen Centre for Functional Genome Research, Institute of Medical Biochemistry and Genetics, Panum Institute 24.4., Blegdamsvej 3, 2200 Copenhagen N, Denmark.
Tel: +45 35327836
Fax: +45 35327845
mail: linda@imbg.ku.dk

Wordcount: 3289
Abbreviations: CP, isolated cleft palate; CL, isolated cleft lip; CLP, combined cleft lip and palate; CL/P, cleft lip and/or palate; COGENE , Craniofacial and Oral Gene Expression Network; QPCR, quantitative PCR; MGI, Mouse Genome Informatics

Abstract

Objective: Epidemiological studies point to different aetiologies in cleft lip and palate subgroups. To gain more insight into these differences, expression profiles in the three cleft lip and palate subgroups were studied.

Design: Affymetrix expression profiles in palate tissue from patients with isolated cleft palate (CP) was compared to expression in palate tissue from combined cleft lip and palate (CLP), and expression profiles in lip tissue from patients with isolated cleft lip (CL) was compared to CLP. Subsequently, the results were analysed by quantitative PCR (QPCR), and by immunohistochemical staining of craniofacial tissue from human embryos. Also, the expression profiles were compared to publicly available gene expression profiles from human embryonic craniofacial tissue (COGENE).

Results and conclusion: Osteopontin (*SPP1*) and other immune related genes were significantly higher expressed in palate tissue from patients with CLP compared to CP. Immunohistochemical staining against selected genes, *SPP1,* chemokine receptor 4 (*CXCR4*) and serglycin (*PRG1*) in sections of craniofacial tissue from human embryos, detected expression of all three genes in the palatal shelves during fusion. Our study suggests that osteopontin and genes involved in the immune response may play a role in the human craniofacial development, and that molecular processes involved in palatal clefting may differ according to the type of cleft.

Keywords: Gene expression, cleft lip and palate, osteopontin (*SPP1*), immune response

Introduction

Cleft lip and/or palate (CL/P) is a common congenital malformation affecting approximately 2 per 1000 newborns worldwide[1] and causing a major psychosocial and economic burden for the families and the society. Gaining insight into the aetiology of CL/P is important as it may lead to improved diagnosis, counselling and perhaps preventive measures in the future.

Both genes and environmental factors contribute to the complex aetiology of CL/P, although most studies point to genetic factors as the major determinants.[2]

Isolated cleft palate (CP) and cleft lip and palate (CLP) may have different aetiologies, as CP and CLP seldom occur in the same families[3], and isolated cleft lip (CL) and CLP display different sex ratios and prevalence of associated malformations.[4]

Rationalizing that these epidemiological differences reflects differences in the underlying molecular processes, we compared the expression profiles in lip and palate tissue removed during surgery from patients with isolated types of orofacial cleft: CL, CP and CLP.

Material

The study was approved by the local scientific ethics committee. Patients with non-syndromic CL, CLP or CP of Caucasian origin were included in the study. Tissue samples were collected during the primary cleft operation, if the parents had given their informed written consent. The patients were operated by one of three surgeons at the same hospital. The children with CL and CLP are operated when they are approximately four months old and the children with CP are operated when they are 16 months old.

To achieve optimal healing and hence cosmetic and functional results, a small piece of mucosa are routinely removed during the operation from the cleft region, either lip and/or palate tissue depending on the CL/P type. In each of the three CL/P subgroups, samples were obtained from three patients. As six tissue samples were collected from CLP patients (three lip and three palate tissue samples), the following two types of comparisons were possible in the total of 12 samples: Expression in lip tissue from patients with CL compared to lip tissue from CLP, and expression in palate tissue from CP compared to CLP. Samples from the lips are primarily composed of non-keratinised epithelium and subepithelial stroma. In embryological terms the lip samples were taken from mucosa developed from the maxillary (lateral) process on the cleft side, not from the frontonasal (median) process. Samples from the palate were taken from the soft palate mucosa at the borders of the cleft and are composed of same cell types as the lip tissue.

Methods

Affymetrix GeneChip analysis: Total RNA was extracted from the tissue using the *mir*Vana miRNA Isolation Kit (Ambion, http://www.ambion.com/) according to the manufacturer's recommendations. The quality and quantity of the extracted total RNAs were assessed by agarose gel electrophoresis, spectrophotometric ultraviolet (UV) absorbance at 260/280 nm, and Agilent Bioanalyzer analysis. Gene expression was performed on Affymetrix GeneChip Human Genome U133A Plus 2.0 Array (Affymetrix,

Inc., Santa Clara, CA). Labelling, hybridization, washing and image scanning were performed according to the Affymetrix standard protocol (supplementary information). Estimation, normalisation of gene expression values and hierachical clustering were performed using dCHIP software.[5] Transcripts were identified as being differentially expressed if the comparison of groups yielded a fold change higher than 1.5 (using the 90% lower confidence interval of the fold change), a difference of means higher than 50 and a $P<.05$ (two-tailed, two-sample unequal variance t-test, Welch test).

Validation by Quantitative PCR (QPCR): QPCR was performed in nine genes in two samples from each of the four groups of tissue; lip tissue from CL and CLP and palate tissue from CP and CLP (for information on cDNA synthesis and primer sequences, see supplementary information). The genes were selected on the basis of their function (according to UCSC Genome Browser, RefSeq summaries and Gene Ontology annotations (GO)). Specific primers for the *CD14* antigen and immunoglobulin kappa and lambda constants could not be designed, as the *CD14* antigen only has one small intron (88 basepairs) and the immunoglobulin kappa and lambda constants have several splice variants.

Validation by Immunohistochemistry: A total of seven human embryos (6^{th} to 8^{th} week) and five human fetuses (9^{th} to 11^{th} week) were examined. The human embryos and fetuses, which ranged from 7 – 60 mm crown-rump length (CRL) corresponding to 6^{th} – 11^{th} ovulation weeks, were obtained from spontaneous and legal abortions or ectopic pregnancies. Informed consent was obtained. Immediately following the operation, the embryos and fetuses were dissected into appropriate blocks and fixed for 12 to 24 hours at 4^{o}C in one of the following fixatives: 10% neutral buffered formalin, 4% Formol-Calcium, Lillie's or Bouin's fixatives. The specimens were dehydrated with graded alcohols, cleared in xylene and paraffin embedded. Serial sections, 3-5 μm thick, were cut in transverse, sagittal or horisontal planes and placed on silanised slides. Representative sections of each series were stained with hematoxylin and eosin or with toluidine blue.

The paraffin sections from tissue blocks containing the craniofacial region (including the palatal shelves) were dewaxed, rehydrated and washed in Tris-buffer saline (TBS: 0.05M Tris, 0.15M NaCl, pH 7.6) with 0.01% Nonidet P-40 (TBS/Nonidet)). All sections were incubated in 1.5% H_2O_2 in TBS/Nonidet for 15 min. to block endogenous peroxidase activity, and then in 10% normal goat serum in TBS/Nonidet for 30 min. to block non-specific binding. Sections were incubated overnight at 4ºC with the primary antibody, diluted in 10% goat serum. Sections were incubated with (1) human monoclonal anti-*CXCR4* Ig, diluted 1:1500 (RD systems MAB173), (2) rabbit polyclonal anti-*SPP1* Ig, diluted 1:600 (abcam ab33046), and (3) rabbit polyclonal anti-serglycin Ig, diluted 1:400 (kindly provided from the laboratory of Dr. Niels Borregaard).[6] The primary antibodies were detected using the peroxidase revelation method (DakoCytomation EnVision+DualLink System), according to the manufacturer´s recommendations. As negative controls, sections were incubated with only the secondary antibody.

In silico validation: The Craniofacial and Oral Gene Expression Network (COGENE) is a publicly available database, comprising gene expression profiles from human embryonic craniofacial structures. Comparing our Affymetrix data with expression profiles in COGENE, provided us with information on whether our differentially expressed genes were present in human embryonic craniofacial structures. COGENEs expression profiles consist of Affymetrix GeneChip analysis and Serial Analysis of Gene Expression (SAGE) profiles. Gene expression in palate tissue was compared to COGENE Affymetrix expression profiles from the following structures, since these structures develop into palatal tissue: 4^{th} week pharyngeal arch 1, 5^{th} week pharyngeal arch 1, 6^{th} week maxilla, and 8.5^{th} week palatine shelves. Expression in lip tissue was compared to COGENE Affymetrix expression profiles from: 4^{th} week pharyngeal arch 1, 5^{th} week pharyngeal arch 1, 6^{th} week maxilla, 4^{th} week frontonasal prominence, 5^{th} week frontonasal prominence, 6^{th} week median nasal prominence, and 8.5^{th} week upper lip, as these structures develop into lip tissue. A few COGENE SAGE profiles were relevant for the lip development: 4^{th} week frontonasal prominence, 5^{th} week frontonasal prominence, and 8.5^{th} week upper lip. When the clone (or tag in SAGE) was present in COGENE, absent or not available, it was labelled P, A or N.A. respectively. A SAGE tag had to be identified twice or more to be considered present (Tables 1 and 2).
Signaling pathways were searched for using a human protein interaction network, which we have previously constructed using several network databases (Lage et al., in preparation, see supplementary information for details on the protein interaction network). Also, for the identified genes, phenotype data on mouse mutants was searched for in Jackson Laboratories Mouse Genome Informatics (MGI), as a craniofacial phenotype in mouse could support our findings.

Results

Affymetrix GeneChip analysis: We identified 11 genes differentially expressed in the palate tissue (*CD14* antigen, serglycin (*PRG1*), immunoglobulin lambda constant 2, osteopontin (*SPP1*), chemokine receptor 4 (*CXCR4*), adenosine monophosphate deaminase 2 (isoform L) (*AMPD2*), regeneration associated muscle protease, hypothetical protein MGC27165 and gb:AJ239383, immunoglobulin kappa constant, likely ortholog of rat vacuole membrane protein 1 and hypothetical protein LOC124773) (Table 1).
19 genes were differentially expressed in the lip tissue (bone marrow stromal cell antigen 2 (*BST2*), SMART/ HDAC1 associated repressor protein (*SPEN*), deiodinase, iodothyronine type II (*DIO2*), cytochrome c oxidase subunit VIIa polypeptide 1 (*COX7A1*), butyrophilin, subfamily 3, member A3 (*BTN3A3*), crystallin, alpha B (*CRYAB*), gb:M27487, gb:AA770596, major histocompatibility complex, class II, DQ alpha 2 (*HLA-DQA1*), heat shock protein, alpha-crystallin-related, B6 (*HSPB6*), protein kinase C and casein kinase substrate in neurons 3 (*PACN3*), EGF-like-domain, multiple 6 (*EGFL6*), calsequestrin 1 (*CASQ1*), gb:AF116709, nuclear factor I/X (CCAAT-binding transcription factor) (*NFIX*), kinase interacting stathmin (*UHMK1*), hypothetical protein (*CMYA1*), hypothetical protein MGC34032, EH domain binding protein 1-like 1 (*EHBP1L1*)) (Table 2).

Quantitative PCR: While QPCR validated the Affymetrix GeneChip results in the palate tissue, only five of the six genes tested in lip tissue showed a non-significant trend. *PACN3* was found expressed at similar levels in lip tissue from isolated CL and from CLP by QPCR, and was therefore not verified (Table 1 and 2).

Immunohistochemistry: The early human embryo has a common oronasal cavity, but between the 6th and the 10th week, the developing palate separates the oral from the nasal cavity. Immunostaining of frontal sections in that time interval for *SPP1* and *CXCR4* showed a characteristic, and in many respects complementary reactivity during initial growth and elevation of the palatal shelves (Fig 1 A and 1 B), and during fusion of the palatal processes and removal of the medial edge epithelium at the site of fusion (Fig 1 C and 1 D). *PRG1* protein expression became apparent during the palatal fusion and medial edge epithelium breakdown (Fig 1 E).

In silico validation: CD14, SPP1, CXCR4, AMPD2 and the immunoglobulin kappa and lambda constants found in the palate tissue were present in COGENEs Affymetrix and/or SAGE profiles, while the following genes identified in the lip tissue were present in COGENES profiles: *BST2, SPEN, DIO2, HLA-DQA1, PACN3, EGFL6, UHMK1* and the hypothetical protein *MGC34032*. Furthermore, two transcripts (AA770596 and AF116709) aligned to the genes myristoylated alanine-rich protein kinase C *(MARCKS)* and H19, imprinted maternally expressed untranslated mRNA *(H19)*, which were both present in COGENE.

Studying our human protein interaction network, we identified *CRYAB* and *HSPB6* as neighbours in a protein complex.[7] Furthermore, interaction between *SPP1* and *PRG1* was suggested.[8]
Several of the identified differentially expressed genes had a genetically altered mouse model (MGI), but no mouse model displayed craniofacial abnormalities.

Discussion

Our data points to the existence of molecular differences between CP, CLP and CL, supporting the epidemiological data in the field.

As tissue from normal controls was not available, the present approach merely enabled us to study possible *differences* in the molecular processes in CL/P subgroups, recognizing that CL/P subgroups may also *share* some processes, evidenced by the seldom occurrence of mixed clefting (CL+/-P *and* CP) in the same families.[9] Consequently, our study design does not provide a direct identification of CL/P candidate genes, but may point to gene families or signalling networks potentially important in CL/P.

For reasons of clarity, and as only genes identified in the palatal tissue were validated by QPCR (Table 1), we chose to focus on these genes and omit a discussion of the genes identified in the lip tissue (Table 2).

Secreted phosphoprotein 1, also referred to as osteopontin (*SSP1*), appears to be a promising candidate potentially involved in the cleft palate process. *SSP1* was expressed at significantly higher levels in the palate tissue from the CLP patients compared to the CP patients with Affymetrix GeneChip analysis and QPCR, and it was also identified in human embryonic palate.

SSP1 is involved in ossification and the immunologic response (as a cytokine involved in the T-helper cell type 1 response) (GO), and is present in COGENEs expression data from human embryonic craniofacial tissues. This may indicate that *SSP1* play a role in palate development, and perhaps a different role in CP compared to CLP.

SSP1 has been emphasized as important for palate formation in two recent microarray expression studies on fetal murine palates, since it showed significant changes in gene expression during craniofacial development.[10 11] *SSP1* transgene mice do not display CL/P, but altered osteoclast formation and wound healing (MGI).

We identified a protein interaction between *SPP1* and *PRG1*, and using immunohistochemistry, *PRG1* was also identified in the human embryonic palate during fusion. *PRG1* has recently been shown to inhibit bone mineralization in vitro.[12]

Interestingly, as *SPP1*, several other factors related to the immune system were differentially expressed. *CD14* antigen, *CXCR4*, and immunoglobulin kappa and lambda constants were all significantly higher expressed in palate tissue from CLP patients compared to CP, but only *CXCR4* was verified by QPCR, as specific primers for *CD14* antigen, and immunoglobulin kappa and lambda constants could not be designed. Moreover, immunohistochemistry showed that *CXCR4* was present in the medial edge epithelium in sections of human embryonic palate.

According to hospital records, one out of the three patients in each cleft palate subgroup (CP and CLP) providing the tissue samples for the GeneChip analysis, had received antibiotics postoperatively, because of clinical signs of upper respiratory track infection, so this could not explain the different expression profiles of the immune related factors in the two groups of CP and CLP patients. Consequently, we studied six

additional samples of palate tissue (samples from three patients with CP and three patients with CLP) by routine light microscopy to look for signs of inflammation in the palate. All but one sample showed sign of a patchy mild or moderate chronic inflammatory response in the subepithelial stroma (fig 2 A and 2 B). The reason for this is not known. According to hospital records, only one of the children (with CP) who provided these tissue samples showed clinical signs of infection postoperatively and received antibiotics. A previous study of palatal tissue from CP patients found dense fibrous reactions, but no inflammation.[13]

The role of immunology in relation to CL/P is presently unknown, but the idea that the same molecule may have different functions depending on the time and place is not new. Biochemical pathways may play dual roles in development and immunology, exemplified by interleukins which play a role in both immune response and stem cell development.[14] Animal studies support a role for immune related factors playing a role in craniofacial development: Immune stimulation following teratogenic exposure in pregnant mice decreased the frequency of fetal abnormalities, including cleft palate.[15] This observed "rescue" of fetal abnormalities was considered mediated by *TGFβ3*, *IGFs* and *ILs*.[16] *TGFβ3* is a known immuno-modulatory cytokine[17] and is considered a CL/P candidate gene based on association- and linkage studies and mouse models.[1]

Although autoimmune diseases are generally not associated with CL/P and only a few syndromic CL/P forms, Kabuki- and Velocardiofacial syndrome (MIM [147920] and [192430]) show altered immune response, an epidemiological study has pointed to an association between maternal infection during pregnancy and increased incidence of CL/P, suggesting that altered immune system may influence fetal development.[18]

A few already recognised CL/P genes have functions related to the immune response; interferon regulating factor 6 (*IRF6*) and poliovirus receptor-like 1 (*PVRL1*). The exact function of *IRF6* is presently unknown, but other types of *IRFs* regulate interferons, which are produced by cells upon viral infection, and they mediate an antiviral response.[19] Mutations in *IRF6* cause both Van der Woude syndrome, VWS (MIM [119300]) and Popliteal Pterygium Syndrome, PPS (MIM [119500])[20] and common variants in *IRF6* contribute to the aetiology in non-syndromic CL/P as well.[21] *PVRL1,* a cell surface receptor for alpha-herpesviruses is involved in CL/P; Cleft lip/palate-ectodermal dysplasia syndrome, CLPED1 (MIM [225060]) is caused by homozygous mutations in *PVRL1* and heterozygous mutations may contribute to the risk of non-syndromic CL/P.[22-24]

Finally, in the palate tissue samples, adenosine monophosphate deaminase 2 (*AMPD2*) displayed a significant higher expression in CLP compared to CP by Affymetrix and QPCR analysis. *AMPD2* is present in skeletal muscle and involved in the purine metabolism. Although the evidence is scarce, adenosine analogues have been shown to have embryo-lethal effects when administered in mice.[25]

The current lack of knowledge on the molecular processes leading to craniofacial development implies certain limitations to the conclusions that can be drawn from an expression study like this, and we need to

consider the potential biases thoroughly. Although the immunohistochemical stainings in human embryonic craniofacial tissue sections support our findings, we do not know whether the tissue samples collected are in fact from the craniofacial region where the clefting process is taking place, and since the tissue is collected and analysed post-natally, the cleft-causing process may have ended.
Moreover, as the samples from the CP and CLP patients are from different age groups, this may result in confounding biases in the studied expression profiles. Tables 1 and 2 show that the genes are expressed at higher levels in CLP compared to CP in the palatal tissue except for the regeneration associated muscle protease, and at lower levels in CLP compared to CL in the lip tissue, except for the hypothetical protein *MGC34032*. The reason for this skewed expression pattern is unknown, but may be explained by the genes acting together in signaling networks not known yet. The existence of signaling networks is supported by our protein interaction data that identified interactions between *CRYAB* and *HSPB6* and *SPP1* and *PRG1*, respectively.

Evidence from CL/P linkage studies has provided us with a range of chromosomal loci. A recent meta-analysis of 13 genome wide linkage studies yielded significant linkage to six loci on five chromosomes (1p12-13, 6p23, 6q23-25, 9q21, 14q21-24, and 15q15) and moreover the meta-analysis calculations of all 13 genome wide scans added 10 additional loci (1q32, 2q32-35, 3p25, 7p12, 8p21, 8q23, 12p11, 17q21, 18q21 and 20q13).[26] Only the gene *AMPD2* (1p13.1) identified in our GeneChip analysis is located in one of these CL/P loci. As significant linkage is difficult to obtain in complex diseases (due to locus heterogeneity), this does not exclude the other of our differentially expressed genes from being potentially important in CL/P aetiology.

In conclusion, by Affymetrix Genechip analysis and staining with selected antibodies on sections from human embryonic palates, we found supportive evidence that osteopontin (*SPP1*), chemokine receptor 4 (*CXCR4*) and serglycin (*PRG1*) may play a role in the development of the palate, and this may be a different role in the two types of cleft palate; CP and CLP. How factors normally related to the immune response, may also be involved in the development of the palate is an issue which needs to be explored further. Our data may serve as a starting point from where the differences in the molecular processes of the CL/P subgroups can be studied.

Gene	Accession number	Chromosomal locus	Biological functions	Affy fold change (CP to CLP)	QPCR fold change	COGENE Affy/ SAGE
CD14 antigen	NM_000591	5q31.3	Phagocytosis, apoptosis, inflammatory response, immune response,cell surface receptor linked signal transduction	2.14		A/P
Serglycin (*PRG1*)	J03223, NM_002727	10q21.3	May neutralize hydrolytic enzymes	2.12-2.48		N.A./A
Immunoglobulin lambda constant 2	M87790, AA680302, AV698647	22q11.22	Immune response, antigen binding	18.90-19.62		P/P
Osteopontin/secreted phosphoprotein 1 (*SPP1*)	M83248	4q22.1	Ossification, anti-apoptosis, inflammatory response, cell adhesion and signaling, immune cell chemotaxis, T-helper 1 type immune response	4.46	5.3 ($p<.05$)	P/P
Chemokine receptor 4 (*CXCR4*)	AF348491, AJ224869	2q21.3	Signal transduction, G-protein coupled receptor protein signaling pathway	1.92-2.93	2.0 ($p<.05$)	A/P
Adenosine monophosphate deaminase 2 (isoform L) (*AMPD2*)	AI916249	1p13.3	Nucleotide metabolism purine, ribonucleoside monophosphate biosynthesis	1.67	1.9 ($p<.05$)	A/P
Regeneration associated muscle protease	AI671186	11p13	Proteolysis	-2.07		N.A./A
Hypothetical protein MGC27165 and gb:AJ239383	S55735/ AJ239383	14q32.33	Immune response (function from *IGHG1* and *IGHA1*)	4.67-11.19		N.A./P
Immunoglobulin kappa constant	AW575927, BC005332, BG485135, M63438	2p11.2	Immune response	9.46-11.05		N.A./P
Likely ortholog of rat vacuole membrane protein 1	BF674052	17q23.1	-	2.65		N.A./P
Hypothetical protein LOC124773	AI190160	17q23.2	-	6.13		N.A./A

Table 1 Significant differently ($p<.05$) expressed genes on Affymetrix GeneChip in palate tissue from patients with cleft palate only (CP) versus palate tissue from patients with cleft lip and palate (CLP), selected genes validated by QPCR.
Accessions numbers are from the Affymetrix GeneChip analysis.
Functions are selected from the processes listed in Gene Ontology (GO) or Refseq summary from UCSC Genome Bioinformatics site.
The COGENE database was searched for expression in the relevant tissues (see Methods) in Affymetrix chip analyses and SAGE libraries, A, absent; P, present; N.A.; not available.

Gene	Accession number	Chromosomal locus	Biological functions	Affy fold change (CL to CLP)	QPCR fold change	COGENE Affy/ SAGE
Bone marrow stromal cell antigen 2 (*BST2*)	NM_004335	19p13.11	Humoral immune response, cell signalling, development	-2.11	-2.1, ns	A/P
SMART/ HDAC1 ass. repressor protein (*SPEN*)	AL524033	1p36.13	Transcription, regulation of transcription, DNA-dependent, Notch signaling pathway	-1.90	-1.2, ns	N.A./P
Deiodinase, iodothyronine type II (*DIO2*)	U53506	14q31.1	Activates thyroid hormone	-2.49	-1.6, ns	A/P
Cytochrome c oxidase subunit VIIa polypeptide 1 (*COX7A1*)	NM_001864	19q13.12	Generation of precursor metabolites and energy, electron transport	-2.02		N.A./A
Butyrophilin, subfamily 3, member A3 (*BTN3A3*)	NM_006994	6p22.1	Belongs to the immunoglobulin superfamily, *BTN/MOG*	-1.70		A/A
Crystallin, alpha B (*CRYAB*)	AF007162	11q23.1	Protein folding, muscle contraction, tyrosine kinase signaling pathway, visual perception	-2.10	-2.0, ns	N.A./A
gb:M27487	M27487 (Aligns to *HLA-DPA1*)	6p21.32 and 6p25.1	Immune response, antigen presentation and processing (function from *HLA-DPA1*)	-2.13		N.A./A
gb:AA770596	AA770596 (aligns to part of *MARCKS*)	6q22.1	Cell motility (function from *MARCKS*)	-2.06		N.A./P (*MARCKS* present in SAGE)
Major histocompatibility complex, class II, DQ alpha 2 (*HLA-DQA1*)	X00452	6p21.32	Immune response, antigen presentation and processing	-3.71		N.A./P
Heat shock protein, alpha-crystallin-related, B6 (*HSPB6*)	AL551046	19q13.12	Protein folding, response to unfolded protein	-2.81		N.A./A
Protein kinase C and casein kinase substrate in neurons 3 (*PACN3*)	NM_016223	11p11.2	Endocytosis, negative regulation of endocytosis	-1.75	1.0, ns	N.A./P (SAGE: 4th week posterior rhombomeres)
EGF-like-domain, multiple 6 (*EGFL6*)	NM_015507	Xp22.2	Regulation of cell cycle, proliferation and developmental processes	-2.65	-1.6, ns	N.A./P
Calsequestrin 1 (*CASQ1*)	NM_001231	1q23.2	Calcium ion binding	-3.02		N.A./A
gb:AF116709	AF116709 (Aligns to H19 and JMJD2C)	11p15.5 and 9p24.1	*H19* (imprinted maternally expressed untranslated mRNA) and *JMJD2C* (jumonji domain containing 2C)	-1.90		N.A./P (*H19* present in SAGE)
Nuclear factor I/X (CCAAT-binding	AI817698	19p13.13	DNA replication, regulation of transcription	-2.14		A/A

transcription factor) (*NFIX*)						
Kinase interacting stathmin (*UHMK1*)	AI249980	1q23.3	Amino acid phosphorylation, cell cycle arrest, regulation of protein export from nucleus	-1.91		N.A./P
Hypothetical protein (*CMYA1*)	AW755250	3p22.2	-	-2.57		N.A./A
Hypothetical protein MGC34032	AA001450, AW020413	1p31.1	Expressed in retina and cochlea	2.00-5.08		N.A./P
EH domain binding protein 1-like 1 (*EHBP1L1*)	AK092750	11q13.1	Expressed in small intestine	-2.08		N.A./A

Table 2 Significant differently ($p<.05$) expressed genes on Affymetrix GeneChip in lip tissue, from patients with cleft lip only (CL) versus lip tissue from patients with cleft lip and palate (CLP), selected genes validated by QPCR (ns is non significant).
Accessions numbers are from the Affymetrix GeneChip analysis.
Functions are selected from the processes listed in Gene Ontology (GO) or Refseq summary from UCSC Genome Bioinformatics site.
The COGENE database was searched for expression in the relevant tissues (see Methods section) in Affymetrix chip analyses and SAGE libraries, A, absent; P, present; N.A.; not available.

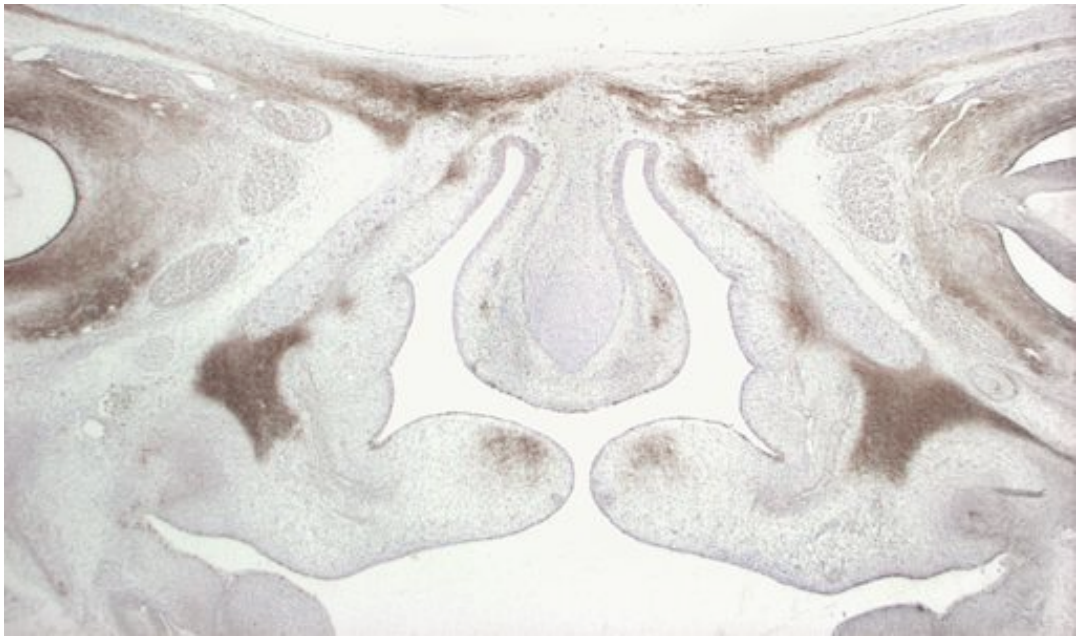

Figure 1A

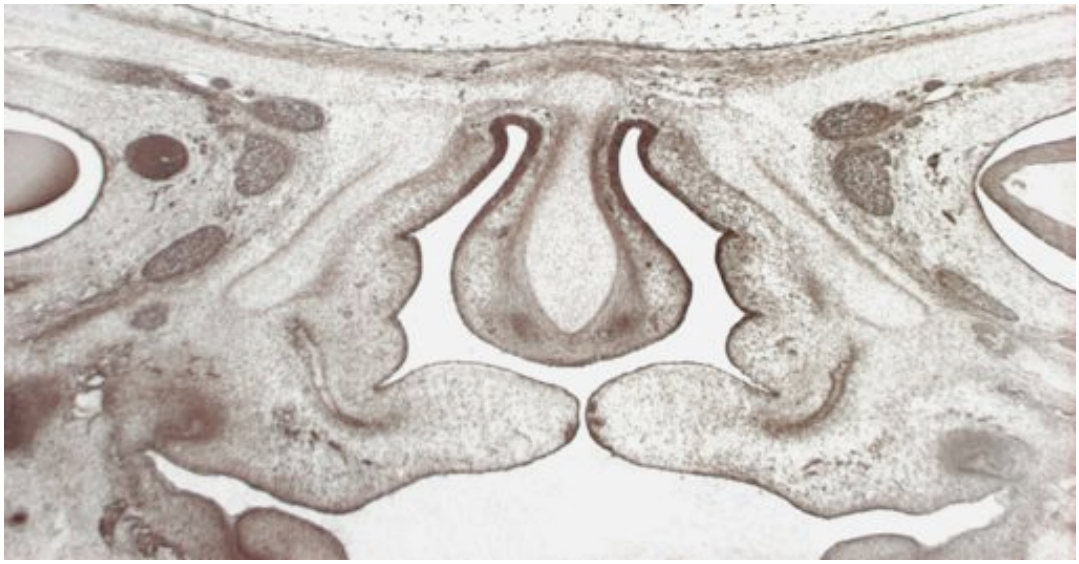

Figure 1B

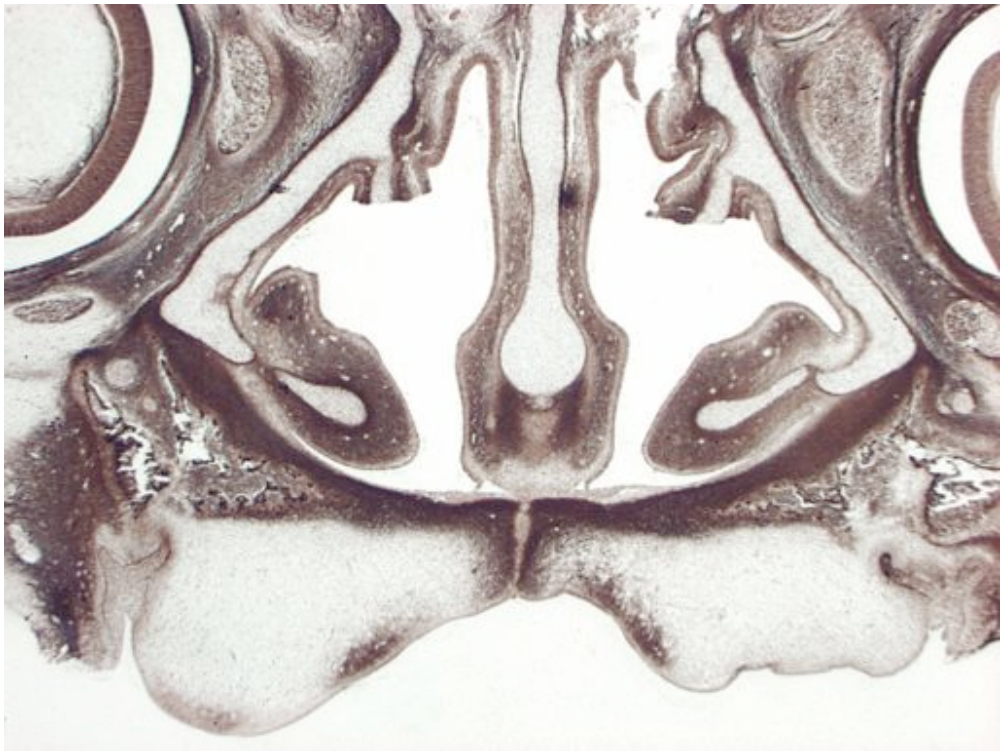

Figure 1C

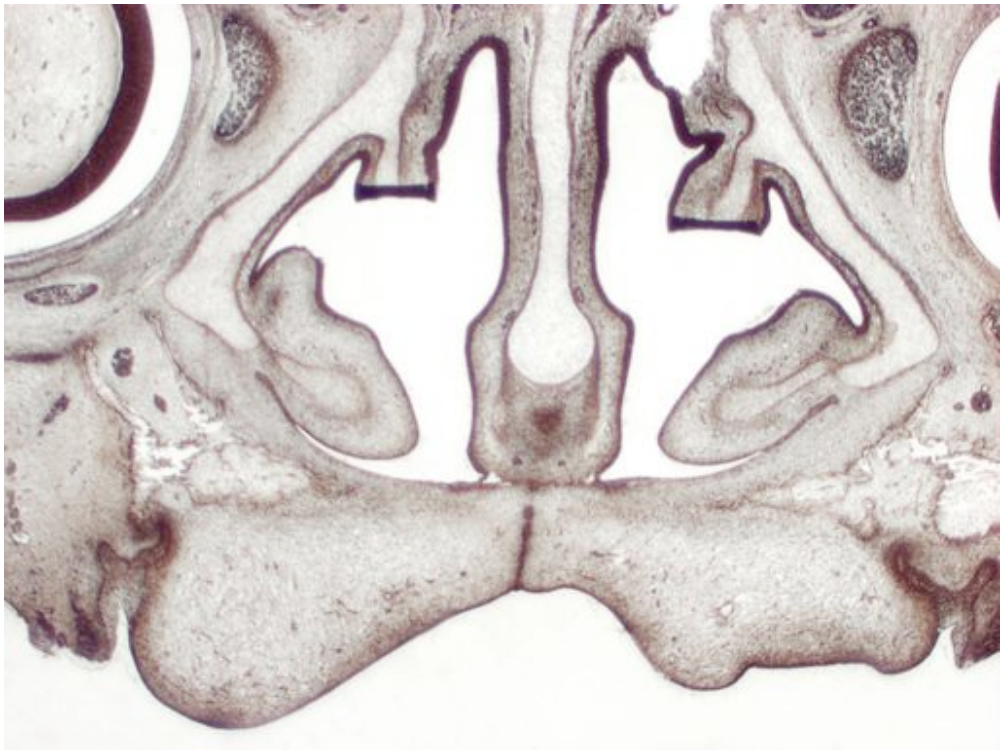

Figure 1D

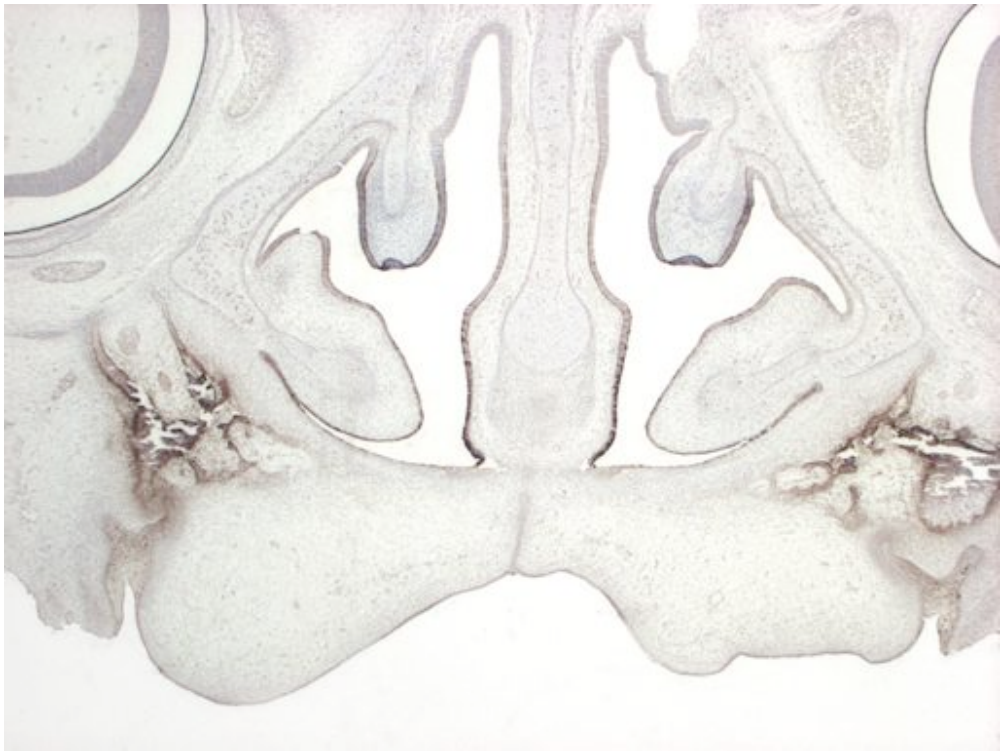

Figure 1E

Figure 1A-E Immunohistochemical staining of human embryonic craniofacial sections. A. During elevation of the palatal shelves osteopontin (*SPP1*) became apparent in the palatal mesenchyme and **B.** Chemokine receptor 4 (*CXCR4*) was present in the epithelial lining of the oronasal cavity and the palatal shelves.
C. During fusion of the palatal shelves *SPP1* is present in the palatal mesenchyme and **D.** *CXCR4* is present in the epithelium of the oronasal cavity, including the medial edge epithelium, where the palatal shelves meet and fuse. **E.** Serglycin (*PRG1*) protein expression became apparent during palatal fusion in the epithelial lining of the oronasal cavity and in the medial edge epithelium. (Hematoxylin-eosin or toluidine blue staining, primary magnification ×1.25).

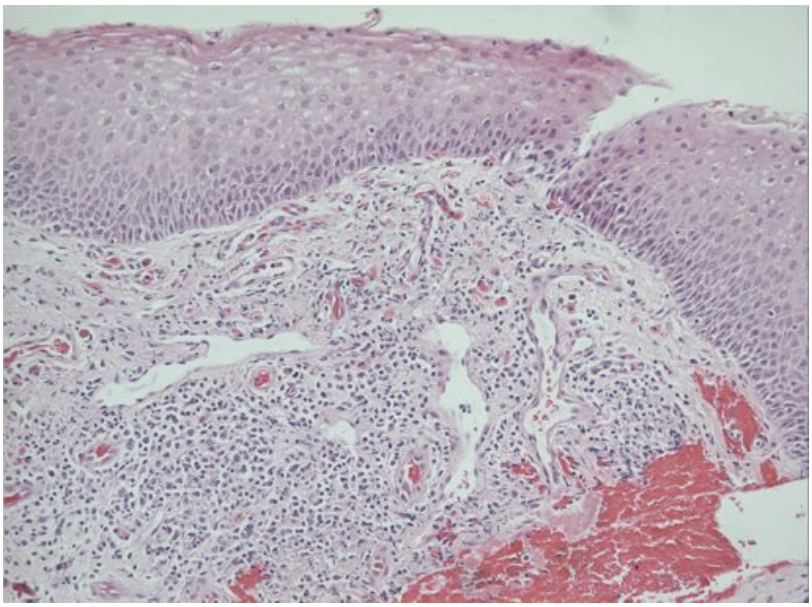

Figure 2A

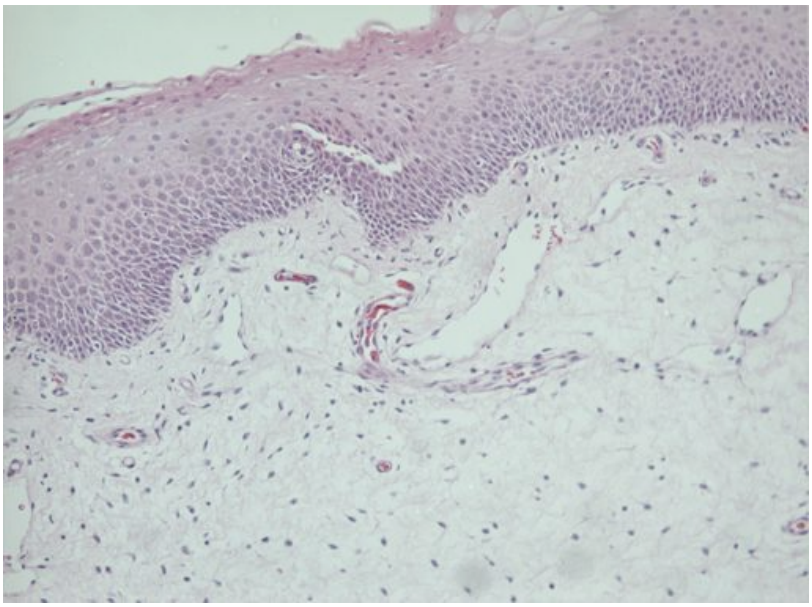

Figure 2B

Figure 2 Routine microscopy of palatal tissue A. Moderate chronic inflammation in the subepithelial stroma of the palatal tissue from one of the patients with cleft lip and palate **B**. Adjacent tissue from the same specimen showing no inflammation, indicating the inflammation was patchy.

Acknowledgments

We thank the patients and their families for participating in the study. The work was partly supported by grants from the Augustinus Foundation, the Vanfoere Foundation, Doctor Johan Boserup and Lise Boserups Foundation, Aase and Ejnar Danielsens Foundation and The Research Foundation of the Queen Louise Childrens Hospital. Wilhelm Johannsen Centre for Functional Genome Research was established by the Danish National Research Foundation.

Authors´affiliations

Linda P. Jakobsen, Wilhelm Johannsen Centre for Functional Genome Research, Institute of Medical Biochemistry and Genetics, Panum Institute and Department of Plastic- and Reconstructive Surgery and Burns Unit, Rigshospitalet, University Hospital of Copenhagen, Denmark.

Rehannah Borup, Department of Clinical Biochemistry, Rigshospitalet, University Hospital of Copenhagen, Denmark.

Kasper Lage, Center for Biological Sequence Analysis (CBS), BioCentrum-DTU, Technical University of Denmark.

Lisa Leth Maroun, Department of Pathology, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark.

Carsten U. Niemann, Department of Haematology, Granulocytlaboratoriet, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark.

Mikael Andersen, Department of Plastic- and Reconstructive Surgery and Burns Unit, Rigshospitalet, University Hospital of Copenhagen, Denmark.

Mary A. Knudsen, Department of Plastic- and Reconstructive Surgery and Burns Unit, Rigshospitalet, University Hospital of Copenhagen, Denmark.

Kjeld Møllgård, Institute of Medical Anatomy, Panum Institute, Copenhagen, Denmark.

Niels Tommerup, **Lars A. Larsen** and **Janni Vestergaard,** Wilhelm Johannsen Centre for Functional Genome Research, Institute of Medical Biochemistry and Genetics, Panum Institute, Copenhagen, Denmark.

Conflict of interest: none declared.

Electronic-Database Information

The URLs for data presented herein are as follows:

Craniofacial and Oral Gene Expression Network (COGENE), http://hg.wustl.edu/cogene/

Gene Ontology (GO), http://www.geneontology.org/

Mouse Genome Informatics (MGI), Jackson Laboratories, http://www.informatics.jax.org/

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/omim/

UCSC Genome Bioinformatics site, March 2006 assembly, http://www.genome.ucsc.edu/

References

1. **Murray JC**. Gene/environment causes of cleft lip and/or palate. *Clin Genet* 2002;**61**:248-56.
2. **Christensen K**, Mitchell LE. Familial recurrence-pattern analysis of nonsyndromic isolated cleft palate-a Danish Registry study. *Am J Hum Genet* 1996;**58**:182-90.
3. **Fogh-Andersen P.** Inheritance of Harelip and Cleft Palate. Copenhagen: Arnold Busck; 1942.
4. **Harville EW**, Wilcox AJ, Lie RT, Vindenes H, Abyholm F. Cleft lip and palate versus cleft lip only: are they distinct defects? *Am J Epidemiol* 2005;**162**:448-53.
5. **Li C,** Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A* 2001;**98**:31-6.
6. **Niemann CU**, Cowland JB, Klausen P, Askaa J, Calafat J, Borregaard N. Localization of serglycin in human neutrophil granulocytes and their precursors. *J Leukoc Biol* 2004;**76**:406-15.
7. **Sugiyama Y**, Suzuki A, Kishikawa M, Akutsu R, Hirose T, Waye MM, Tsui SK, Yoshida S, Ohno S. Muscle develops a specific form of small heat shock protein complex composed of MKBP/HSPB2 and HSPB3 during myogenic differentiation. *J Biol Chem* 2000;**275**:1095-104.
8. **Rual JF**, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, Klitgord N, Simon C, Boxem M, Milstein S, Rosenberg J, Goldberg DS, Zhang LV, Wong SL, Franklin G, Li S, Albala JS, Lim J, Fraughton C, Llamosas E, Cevik S, Bex C, Lamesch P, Sikorski RS, Vandenhaute J, Zoghbi HY, Smolyar A, Bosak S, Sequerra R, Doucette-Stamm L, Cusick ME, Hill DE, Roth FP, Vidal M. Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 2005;**437**:1173-8.
9. **Neilson DE**, Brunger JW, Heeger S, Bamshad M, Robin NH. Mixed clefting type in Rapp-Hodgkin syndrome. *Am J Med Genet* 2002;**108**:281-4.
10. **Brown NL**, Knott L, Halligan E, Yarram SJ, Mansell JP, Sandy JR. Microarray analysis of murine palatogenesis: temporal expression of genes during normal palate development. *Dev Growth Differ* 2003;**45**:153-65.
11. **Mukhopadhyay P**, Greene RM, Zacharias W, Weinrich MC, Singh S, Young WW Jr, Pisano MM. Developmental gene expression profiling of mammalian, fetal orofacial tissue. *Birth Defects Res A Clin Mol Teratol* 2004;**70**:912-26.
12. **Theocharis AD**, Seidel C, Borset M, Dobra K, Baykov V, Labropoulou V, Kanakis I, Dalas E, Karamanos NK, Sundan A, Hjerpe A. Serglycin Constitutively Secreted by Myeloma Plasma Cells Is a Potent Inhibitor of Bone Mineralization in Vitro. *J Biol Chem* 2006;**281**:35116-35128.
13. **Stal S**, Hicks MJ. Classic and occult submucous cleft palates: a histopathologic analysis. *Cleft Palate Craniofac J* 1998;**35**:351-8.
14. **Barrow PC**. Reproductive toxicology studies and immunotherapeutics. *Toxicology* 2003;**185**:205-12.
15. **Holladay SD**, Sharova LV, Punareewattana K, Hrubec TC, Gogal RM Jr, Prater MR, Sharov AA. Maternal immune stimulation in mice decreases fetal malformations caused by teratogens. *Int Immunopharmacol* 2002;**2**:325-32.
16. **Sharova LV**, Gogal RM Jr, Sharov AA, Chrisman MV, Holladay SD. Immune stimulation in urethane-exposed pregnant mice increases expression level of spleen leukocyte genes for TGFbeta3 GM-CSF and other cytokines that may play a role in reduced chemical-induced birth defects. *Int Immunopharmacol* 2002;**2**:1477-89.
17. **Wahl SM**, Vazquez N, Chen W. Regulatory T cells and transcription factors: gatekeepers in allergic inflammation. *Curr Opin Immunol* 2004;**16**:768-74.
18. **Metneki J**, Puho E, Czeizel AE. Maternal diseases and isolated orofacial clefts in Hungary. *Birth Defects Res A Clin Mol Teratol* 2005;**73**:617-23.
19. **Taniguchi T**, Takaoka A. The interferon-alpha/beta system in antiviral responses: a multimodal machinery of gene regulation by the IRF family of transcription factors. *Curr Opin Immunol* 2002;**14**:111-6.
20. **Kondo S**, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, de Lima RL, Daack-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, Lammer EJ, Aylsworth AS, Ardinger HH, Lidral AC, Pober BR, Moreno L, Arcos-Burgos M, Valencia C, Houdayer C, Bahuau M, Moretti-Ferreira D, Richieri-Costa A, Dixon MJ, Murray JC. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nat Genet* 2002;**32**:285-9.

21. **Zucchero TM**, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S, Schutte BC, Marazita ML, Murray JC. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *N Engl J Med* 2004;**351**:769-80.
22. **Suzuki K**, Hu D, Bustos T, Zlotogora J, Richieri-Costa A, Helms JA, Spritz RA. Mutations of PVRL1, encoding a cell-cell adhesion molecule/herpesvirus receptor, in cleft lip/palate-ectodermal dysplasia. *Nat Genet* 2000;**25**:427-30.
23. **Sozen MA**, Suzuki K, Tolarova MM, Bustos T, Fernandez Iglesias JE, Spritz RA. Mutation of PVRL1 is associated with sporadic, non-syndromic cleft lip/palate in northern Venezuela. *Nat Genet* 2001;**29**:141-2.
24. **Avila JR**, Jezewski PA, Vieira AR, Orioli IM, Castilla EE, Christensen K, Daack-Hirsch S, Romitti PA, Murray JC. PVRL1 variants contribute to non-syndromic cleft lip and palate in multiple populations. *Am J Med Genet* 2006;**140**:2562-70.
25. **Clark RL**, Eschbach K, Cusick WA, Heyse JF. Interactions between caffeine and adenosine agonists in producing embryo resorptions and malformations in mice. *Toxicol Appl Pharmacol* 1987;**91**:371-85.
26. **Marazita ML**, Murray JC, Lidral AC, Arcos-Burgos M, Cooper ME, Goldstein T, Maher BS, Daack-Hirsch S, Schultz R, Mansilla MA, Field LL, Liu YE, Prescott N, Malcolm S, Winter R, Ray A, Moreno L, Valencia C, Neiswanger K, Wyszynski DF, Bailey-Wilson JE, Albacha-Hejazi H, Beaty TH, McIntosh I, Hetmanski JB, Tuncbilek G, Edwards M, Harkin L, Scott R, Roddick LG. Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35. *Am J Hum Genet* 2004;**75**:161-73.

Supplementary information

Affymetrix GeneChip analysis

Two µg of total RNA was used to synthesise double-stranded cDNA with the Superscript Choice system (Invitrogen) using an oligo(dT) primer containing a T7 RNA polymerase promoter (GenSet). The cDNA was used as the template for an in vitro transcription reaction to synthesise biotin-labeled antisense cRNA (IVT labelling Kit, Affymetrix Inc.). After fragmentation at 94°C for 35 min. in fragmentation buffer (40 mM Tris, 30 mM magnesium acetate, 10 mM potassium acetate), the labeled cRNA was hybridised for 16 h to Affymetrix HG-U133 Plus 2.0 arrays (Affymetrix Inc.), that contain 54,613 probe sets. The arrays were washed and stained with phycoerythrin-streptavidin (SAPE) using the Affymetrix Fluidics Station 450, and the arrays were scanned in the Affymetrix GeneArray 3000 scanner to produce raw data image files.

QPCR analysis

cDNA was synthesised with superscript II (RNase H^-) reverse transcriptase (Invitrogen, Carlsbad, CA) and QPCR analysis was performed in a DNA Engine Opticon 2 (Biorad, Hercules, CA) using LightCycler FastStart DNA master^{PLUS} SYBR GreenI kit (Roche, Indianapolis, IN). Primers were purchased from TAG Copenhagen (Copenhagen, DK). All primers are intron-spanning and designed using Oligo6 software (Cascade, CO). The identity of PCR products and primer sequence specificity was confirmed by BLAST analysis (NCBI, NIH, http://www.ncbi.nlm.nih.gov/BLAST/) and melting curve analyses. Gene expression levels were normalised to the expression of beta-2-microglobulin precursor (*B2M*).

Primer sequences (from 5' to 3') were:

B2M, forward (F)-TGTGCTCGCGCTACTCTCTC, reverse (R)- CTGAATGCTCCACTTTTTCAATTCT;

BST2, F-GATGGAGTGTCGCAATGTCA, R-GCCCAGCAGCACAATCA;

SPEN, F-GTCCTGGCCCATCGGT, R-TTGGCCTGCAGGTAAGTGAT;

DIO2, F-AGAGGGACTGCGCTGCGTCT, R-CTGGCAGCTGGCTCGTGAAA;

CRYAB, F-GTTGGGAGATGTGATTGAGG, R-GCTTCTCTTCACGGGTGAT;

PACN3, F-GGGTGAGGGCACTCTATGAC, R-GCCCTCAGCCTAGACTCGT;

EGFL6, F-AGACATTGGCCGATTGAA, R-ATGCCAGAGGTCCTATGATG;

SPP1, F-CTAAGAAGTTTCGCAGACCT, R-CTATCAATCACATCGGAATG;

CXCR4, F-TCCAGTAGCCACCGCATCT, R-TGCCCACAATGCCAGTTAAG;

AMPD2, F-CTCAGCTATCACCGGAATC, R-TCTTTACCTTGTGCGAGAAG.

Protein interactions

We have previously constructed a human protein interaction network (Lage et al. , in preparation) by downloading data from MINT[1], BIND[2], IntAct[3], KEGG[4] and Reactome[5], and increasing coverage using interolog data (the transfer of protein interactions between orthologous protein pairs in different organisms)[6] from 17 eukaryotic organisms. Orthology was assigned using the Inparanoid database[7] with strict thresholds.

1. **Zanzoni A,** Montecchi-Palazzi L, Quondam M, Ausiello G, Helmer-Citterich M, Cesareni G. MINT: a Molecular INTeraction database. *FEBS Lett* 2002;**513**:135-40.

2. **Bader GD**, Betel D, Hogue CW. BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res* 2003;**31**:248-50.

3. **Hermjakob H**, Montecchi-Palazzi L, Lewington C, Mudali S, Kerrien S, Orchard S, Vingron M, Roechert B, Roepstorff P, Valencia A, Margalit H, Armstrong J, Bairoch A, Cesareni G, Sherman D, Apweiler R. IntAct: an open source molecular interaction database. *Nucleic Acids Res* 2004;**32(**Database issue**)**:D452-5.

4. **Kanehisa M**, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* 2006;**34**(Database issue):D354-7.

5. **Joshi-Tope G**, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath GR, Wu GR, Matthews L, Lewis S, Birney E, Stein L. Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res* 2005;**33**(Database issue):D428-32.

6. **Walhout AJ**, Sordella R, Lu X, Hartley JL, Temple GF, Brasch MA, Thierry-Mieg N, Vidal M. Protein interaction mapping in C. elegans using proteins involved in vulval development. *Science* 2000;**287**:116-22.

7. **O'Brien KP**, Remm M, Sonnhammer EL. Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res* 2005;**33 (**Database Issue**)**:D476-80.

Paper IV

Clinical report to American Journal of Medical Genetics

Title: **Suggestive linkage to a neighboring region of *IRF6* in a cleft lip and palate multiplex family**

Authors:
Linda P. Jakobsen, M.D.[1, 2]
Reinhard Ullmann, M.Sc., Ph.D. [3]
Klaus W. Kjaer, M.D., Ph.D.[1]
Mary A. Knudsen, M.D., Ph.D.[2]
Niels Tommerup, M.D., Ph.D.[1]
Hans Eiberg, M.Sc. [4]

1) Wilhelm Johannsen Centre for Functional Genome Research (WJC), Institute of Medical Biochemistry and Genetics (IMBG), Panum Institute, Copenhagen, Denmark.
2) Department of Plastic- and Reconstructive Surgery and Burns Unit, Rigshospitalet, University Hospital of Copenhagen, Denmark.
3) Department of Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Berlin, Germany.
4) Institute of Medical Biochemistry and Genetics (IMBG), Panum Institute, Copenhagen, Denmark.

Correspondance to:
Linda P. Jakobsen, Wilhelm Johannsen Centre for Functional Genome Research (WJC), Institute of Medical Biochemistry and Genetics (IMBG), Panum Institute 24.4., Blegdamsvej 3, 2200 Copenhagen N, Denmark.
Telephone: +45 35327836
Fax: +45 35327845
mail: linda@imbg.ku.dk

Acknowledgement:
We thank the family for their participation in the study. We thank Anne Mette Mikkelsen for skilful technical assistance. The work was partly supported by grants from the Augustinus Foundation, the Vanfoere Foundation, Doctor Johan Boserup and Lise Boserups Foundation, Aase and Ejnar Danielsens Foundation, Jacob Madsen and wife Olga Madsens Foundation and Queen Louise Childhospital Research Foundation. Wilhelm Johannsen Centre for Functional Genome Research is established by the Danish National Research Foundation.

Wordcount: 2100
Abbreviations: CL/P, Cleft lip and/or palate; CGH, Comparative Genome Hybridization; kb, kilobase; LOD, logarithm of odds; Mb, megabase; NSCLP, non-syndromic CL/P; SNP, single nucleotide polymorphism; STS, Sequence Tagged Site

Abstract

Background: Cleft lip and/or palate (CL/P) is a common congenital malformation with a complex etiology, as many genes and environmental factors play a role. One way to overcome the issue of complexity is by studying CL/P families separately, as relatively few loci are expected to contribute to the CL/P etiology in each family.

Results: A 10K Affymetrix SNP genome scan followed by fine mapping with microsatellite markers in a CL/P multiplex family, suggested linkage (maximum LOD of 2.73, θ=0.0) to a 6.5 Mb interval at 1q32.1-1q32.2 next to, but excluding *IRF6* from the interval. Sequencing of selected genes and comparative genome hybridization in the most severely affected family member did not reveal any mutations or genomic aberrations.

Discussion: Mutations in the gene *IRF6* (1q32.2) can cause syndromic CL/P, Van der Woude syndrome, and several association studies have shown that *IRF6* polymorphisms are associated with non-syndromic CL/P. However, in this CL/P multiplex family, the coding part of *IRF6* was excluded from the linkage interval, whereas part of the non-coding landscape potentially controlling *IRF6* expression was included. Our data suggest that an unidentified CL/P gene or genes neighboring of *IRF6*, or non-coding regulatory elements exert an effect on *IRF6* expression, and cause CL/P in this multiplex family.

Keywords: cleft lip and palate, linkage, non-coding regulatory elements, *IRF6*

Introduction

Cleft lip and/or palate (CL/P) is a common congenital malformation affecting approximately 2 per 1000 newborns worldwide. The etiology of CL/P is complex, as many genetic and environmental factors play a role (Murray, 2002). Although some success in identifying CL/P genes has been achieved by the delineation of syndromes with orofacial clefting (Stanier and Moore, 2004), the overall etiology in non-syndromic CL/P (NSCLP) is mostly unknown. A variety of approaches are used to study the genetic etiology in CL/P including animal models, association- and linkage studies, cytogenetic and expression studies.

CL/P linkage studies have been performed for decades. The first locus (*F13A*, 6p25.1) for NSCLP was identified in 1987 (Eiberg et al., 1987), and subsequently many loci have been suggested. A major obstacle in CL/P linkage studies are the lack of replicative studies, caused by locus heterogeneity, as it has been pointed out that 2-14 different loci may be involved in the etiology of CL/P (Schliekelman and Slatkin, 2002). Moreover, CL/P linkage studies also suffer from lack of significant results (LOD scores >3), partly as a consequence of locus heterogeneity and partly because of the lack of large multiplex families (two or more affected family members). Most often LOD scores of 2-3, which is only suggestive of linkage, must be accepted in complex diseases as CL/P (Lander and Kruglyak, 1995; Altmüller et al., 2001). Only two recent studies have yielded genome wide significant linkage in NSCLP; a meta-analysis comprising 13 genomewide scans identified 16 loci with significant results (Marazita et al., 2004) and a study in two Indian families reported significant linkage to locus 13q33.1-34 (Radhakrishna et al., 2006).

Here we present a multiplex NSCLP family with suggestive linkage (maximum LOD score of 2.73) to a 6.5 Mb locus on 1q32.1-1q32.2, *omitting* the well known CL/P gene, *IRF6*, (Kondo et al., 2002; Zucchero et al., 2004) from the linkage interval. This could point to the existence of other CL/P genes or non-coding regulatory elements in this neighboring region of *IRF6*, exerting effects on *IRF6*.

Patients and Methods

The study was approved by the Local Scientific Ethics Committee and written informed consents have been obtained from the family members. The patient photographed has agreed to have his picture published in a scientific paper. The NSCLP multiplex family (fig.1) has had five affected members (fig. 2 is a picture of affected family member III-1), but person II-9 died in 1996 and DNA from this person was not available. This family displays different types of non-syndromic cleft lip with or without palate. The affected family members III-1 and IV-1 had unilateral cleft lip and palate, III-7 had bilateral cleft lip and palate, and IV-2 had unilateral cleft lip. There is a history of psoriasis in some family members (III-2, III-3, the mother of I-2 and a sister of III-9), moderately elevated blood pressure (II-1, II-3, II-5 and II-7) and celiac disease in IV-2. None of these diseases co-segregate with the CL/P phenotype. No other associated malformations were identified in the family, especially no lower lip pits or oligo- or hypodontia, which is seen in patients with *IRF6* mutations (van der Woude syndrome (VWS, MIM #119300) and popliteal pterygium syndrome (PPS, MIM # 119500)) were found.

Genomic DNA was isolated from whole blood according to standard protocols. The genome wide scan was performed with a 10 K Affymetrix single nucleotide polymorphisms (SNPs) chip (a service performed at Institute of Medical Genetics, Humboldt University Berlin, Berlin). Fine mapping of the region was carried out with microsatellite DNA markers (for PCR primer sequences and chromosome positions of the microsatellite markers, see table 1). The dinucleotide repeat markers "#5" (F:AGGCTCTTCCCTGATACACG; R: GTTTCTTTCCTGCCTGGTTG) and "#2" (F:AATTGCAGGAATGTGGAACC; R: ATCCAAGTTGCTGTGAATGC) (~232 bp and ~165 bp, respectively), were designed as publicly available microsatellite markers were not informative in this chromosomal region.

LOD score was calculated (for microsatellite markers) using the linkage program MLINK (FASTLINK) (Elston and Stewart, 1971; Schäffer et al., 1994).

PCR products of the translated regions of the genes *IRF6, SOX13, FMOD, OPTC, IKBKE* and the microRNAs mir 29b-2, 29c and 135b were sequenced in both directions in the most severely affected family member (III-7) with bilateral CLP (PCR primer sequences and conditions are available upon request) using ABI Big Dye version 1.1 and an ABI3100 sequencing machine (Applied Biosystems, Foster City, CA). The DNA sequences were analyzed using Chromas (Technelysium Pty. Ltd., Tewantin).

Comparative Genome Hybridization (CGH) was carried out in family member (III-7) with bilateral CLP, using a submegabase resolution whole genome tiling path BAC array consisting of the human genome high resolution 32 k re-arrayed clone set (BACPAC Resources), the 1 Mb Sanger set (Wellcome Trust Sanger Institute) and a set of 390 subtelomeric clones (assembled by members of the COST B19 initiative: Molecular Cytogenetics of solid tumors) (Erdogan et al., 2006). Deviant signal intensity ratios involving three or more neighboring BAC clones were considered as genomic aberrations unless they were fully covered by a known DNA copy number variant, as listed in the Database of Genomic Variants.

Results

The genome wide scan suggested linkage to a region between SNPs rs2185781 and rs1567190 at 1q32.1-1q32.2 (NCBI dbSNP Build 126 and UCSC Genome Bioinformatics site, March 2006 assembly). The linkage interval was confirmed and fine mapped with microsatellite markers (#5 and #2) to a 6.5 Mb interval at chromosome 1 position 201338592-207914577 (fig. 1 and table 1). A maximum two-point LOD score of 2.73 (θ=0.0) was calculated for marker D1S2796 (assuming dominant inheritance, a disease-allele frequency of 0.001 and affected only). Attempting to narrow the linkage interval further, we also typed the markers D1S2727, D1S504 and D1S245, and two dinucleotide repeat markers (primers not shown), but these markers were not informative in the family.

The 6.5 Mb region encompasses more than 50 known genes (fig. 3). Based on a literature search we chose to sequence the translated regions of the genes *IRF6, SOX13, FMOD, OPTC, IKBKE* and the three microRNA genes 29b-2, 29c and 135b. However, we detected no mutations, nor did we detect any DNA copy number change by array-CGH.

Discussion

Albeit only suggestive linkage was found, it is interesting that the linkage region is very close to, but does not include *IRF6* (position 208027885-208046102), as *IRF6* is separated from our linkage interval (201338592-207914577) by more than 100 kb. *IRF6* is the prime candidate gene of interest in this region, as the syndromic forms of CL/P, VWS and PPS, are caused by mutations in *IRF6* (Kondo et al., 2002). *IRF6* is a member of a group of eight transcription factors (*IRF1-8*) involved in the immune response (Taniguchi et al., 2001) and *IRF6* mutant mice display defective epidermal development (Ingraham et al., 2006).

VWS may mimic NSCLP in humans, as no lip pits are found in ~15% of patients with VWS (Burdick et al., 1985). Moreover, an association between NSCLP and different variants in and around *IRF6* has been identified, however, specific disease-causing mutations have not been identified (Zucchero et al., 2004). Likewise, sequencing of *IRF6* in individual III-7 (the person with bilateral CL/P) revealed no coding mutations, but a few common variants.

The present linkage study point to an involvement of locus 1q32 in NSCLP, although not the direct involvement of *IRF6* in NSCLP. Supportive of this, is the meta-analysis of CL/P genome scans, where significant linkage to region 1q32 was identified (Marazita et al., 2004), while only weak linkage (LOD<1.0) to the *IRF6* locus itself was found (Zucchero et al., 2004).

It is obvious to consider whether the surrounding region of *IRF6* harbors genes or non-coding regions with a regulatory effects on *IRF6*, as regulatory elements exerting long-range effects have been found more than 1 Mb from the gene, and even within the intronic region of a neighboring gene (Kleinjan and van Heyningen, 2005). At the 3´end of *IRF6*, a ~1.5 Mb gene empty region is located (~position 206.4-207.9 Mb), harboring several highly conserved and potentially regulatory regions.

Association studies of *IRF6* suggest that the region surrounding *IRF6* may be of importance in NSCLP, but the results are inconsistent. Zucchero and colleagues (2004) found association between several SNPs and NSCLP in a long linkage block extending from 40 kb 5` to at least 100 kb 3´ of *IRF6*. A SNP 100 kb 3´ of *IRF6 (*rs2235543, position 207927291) was significantly associated with NSCLP in Filipino triads, but not in Danish or Iowa triads (the border of our linkage interval closest to *IRF6* is only 13 kb away from this SNP). In addition in this comprehensive study, three conserved regions 3´of *IRF6* (81 kb, 103 kb and 117 kb) were sequenced in a group of patients, but no mutations were detected.

The association of *IRF6* and NSCLP has been confirmed in a European population, but the association of the SNP 100 kb 3´of *IRF6* (rs2235543) was not confirmed (Ghassibe et al., 2005; Scapoli et al., 2005), nor did Houdayer and colleagues (2001) detect association to a marker (D1S491) located 200 kb 3´of *IRF6*. The lack of significant association of SNPs 100 kb 3´of *IRF6* in the European population could indicate that this region does not provide a major contribution to the CL/P etiology in this population.

Lessons from other studies have taught us that selecting candidate disease genes based on known functions is a difficult task. No obvious candidate genes are located in our linkage interval (fig. 3), as none of the genes have been related to clefting in mice (Mouse Genome Informatics (MGI)) or in humans (Online Mendelian Inheritance in Man (OMIM)). Although not much is known about the function of *SOX13*, we chose to sequence the translated exons of this gene, as mutations in a related gene (*SOX9*) are causing campomelic dysplasia (MIM #114290) which is a severe skeletal dysplasia involving cleft palate. *FMOD* and *OPTC* were sequenced as both genes interact with collagen fibrils, and *FMOD* regulate TGF*B* activities (functions from UCSC Genome Bioinformatics site), and we also chose to sequence *IKBKE*, an IkappaB kinase interacting with *IRF3*, a member of the *IRFs* (Fitzgerald et al., 2003; Sharma et al., 2003). *IKBKE* shows some similarity with *IKK1* (or *CHUK*), another IkappaB kinase, and *Ikk1*-deficient mice may display cleft palate (Li et al., 1999). No mutations in the coding regions of *SOX13, IKBKE, FMOD* or *OPTC* were encountered. Furthermore, at the resolution of BAC arrays (~100 kb), we could exclude the presence of microdeletions or microduplications within the 6.5 Mb linkage interval.
This interval also includes three microRNA genes, mir 29b-2, 29c and 135b. MicroRNAs are short non-coding RNAs involved in down regulation of genes at the translational level by binding to specific target sites in the 3´-UTR of the mRNA (Berezikov and Plasterk, 2005). Since mutations in microRNAs could potentially alter the expression of microRNA target genes, we sequenced mir 29b-2, 29c and 135b, but detected no sequence alterations. As it was only the translated regions that were sequenced, we cannot exclude the mutations in non-coding regions of the genes.

Our data suggest that the gene *IRF6* should not be the only CL/P gene to focus on in the 1q32 region. Either *IRF6* regulatory elements, a gene acting in the same signaling pathway as *IRF6,* or a gene with a distinct function in CL/P could be located in the presented linkage interval neighboring of *IRF6*.

Electronic-Database Information

The URLs for data presented herein are as follows:

Database of Genomic Variants, http://projects.tcag.ca/variation/

FASTLINK, http://softlib.rice.edu/fastlink.html

Mouse Genome Informatics (MGI), Jackson Laboratories, http://www.informatics.jax.org/

NCBI, dbSNP Build 126, http://www.ncbi.nlm.nih.gov/SNP/

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/

UCSC Genome Bioinformatics site, March 2006 assembly, http://www.genome.ucsc.edu/

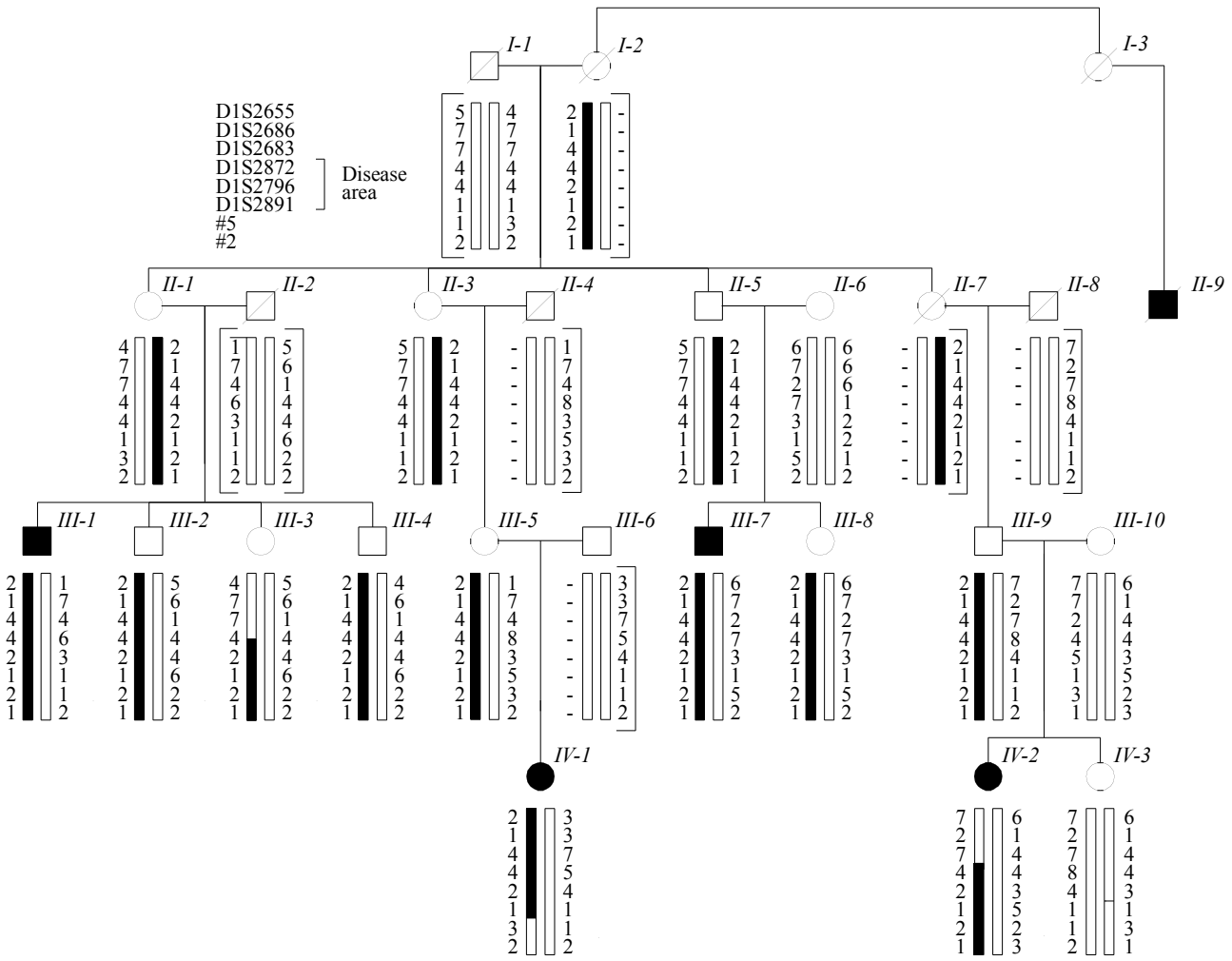

Figure 1. Pedigree of the CL/P multiplex family. Affected individuals are shown with blackened symbols. The black bars indicate the haplotype associated with CL/P. See table 1 for information on microsatellite markers. As D1S2655, D1S2686 and D1S2683 show recombination (in IV-2) and marker #5 and #2 show recombination (in IV-1), the linkage interval is between D1S2683 and #5.

Figure 2. Family member III-1 with operated unilateral cleft lip and palate.

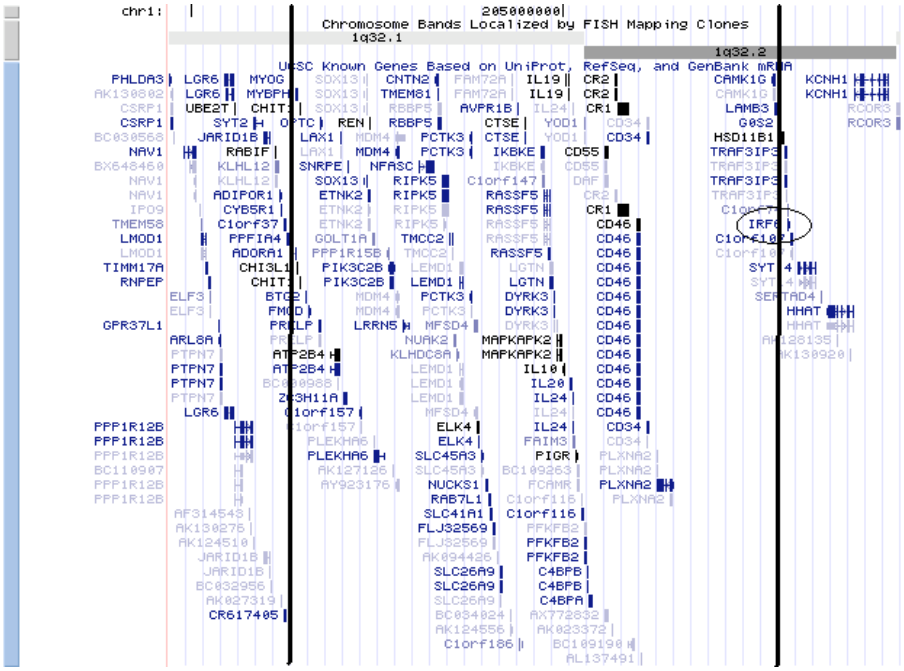

Figure 3. The ~ 6.5 Mb linkage interval, chromosome position 201338592-207914577, showing more than 50 known genes. The lines indicate the borders of the linkage interval identified by microsatellite markers. Note that *IRF6* is located outside the border of the interval (circled) and 3´of *IRF6* is a ~1.5 Mb gene empty region located (UCSC Genome Bioinformatics site, March 2006 assembly).

Markers	Primers (5´-3´)	Chromosome position
D1S2655	F: AGGGTCCCCAAAGAGCCTTC R: ATGGCAGCACATCCTGCTTC	200831887- 200832229
D1S2686	F: GGGACTTTTGTATTGACTGAC R: GTAGTTTGCTGAGCAGAGG	200903162-200903596
D1S2683	F: TGCCTTGTCTTCAAGAGC R: GCAGTGACAGGAATCTGG	201338227- 201338592
D1S2872	F: GGTACATGGGGGATCG R: CGGGACATCAGTGAGG	203092038- 203092348
D1S2796	F: TGCACCACTCTACTCCACCT R: AGGCTGATGGCTTAGTCTGT	205826035- 205826266
D1S2891	F: ACTGCTTATTCGGAGTTGGA R:CCAAGAGTTTTCTTAGCAAATCAC	206519051- 206519413
#5	F: AGGCTCTTCCCTGATACACG R: GTTTCTTTCCTGCCTGGTTG	207914577-207914808
#2	F: AATTGCAGGAATGTGGAACC R: ATCCAAGTTGCTGTGAATGC	207980256-207980420

Table 1. Information on microsatellite markers. #5 and #2 are dinucleotide repeats. F is forward primer, R is reverse primer. Chromosome position according to UCSC Genome Bioinformatics site, March 2006 assembly. The position of *IRF6* is 208027885-208046102.

References

Altmüller J, Palmer LJ, Fischer G, Scherb H, Wjst M. 2001. Genomewide scans of complex human diseases: true linkage is hard to find. Am J Hum Genet 69:936-50.

Berezikov E, Plasterk RH. 2005. Camels and zebrafish, viruses and cancer: a microRNA update. Hum Mol Genet 14:R183-R190.

Burdick AB, Bixler D, Puckett CL. 1985. Genetic analysis in families with van der Woude syndrome. J Craniofac Genet Dev Biol 5:181-208.

Eiberg H, Bixler D, Nielsen LS, Conneally PM, Mohr J. 1987. Suggestion of linkage of a major locus for nonsyndromic orofacial cleft with F13A and tentative assignment to chromosome 6. Clin Genet 32:129-132.

Elston RC, Stewart J. 1971. A general model for the genetic analysis of pedigree data. Hum Hered 21:523-542.

Erdogan F, Chen W, Kirchhoff M, Kalscheuer VM, Hultschig C, Müller I, Schulz R, Menzel C, Bryndorf T, Ropers H-H, Ullmann R. 2006. Impact of low copy repeats on the generation of balanced and unbalanced chromosomal aberrations in mental retardation. CGR 115:247-253.

Fitzgerald KA, McWhirter SM, Faia KL, Rowe DC, Latz E, Golenbock DT, Coyle AJ, Liao SM, Maniatis T. 2003. IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway. Nat Immunol 4:491-496.

Ghassibe M, Bayet B, Revencu N, Verellen-Dumoulin C, Gillerot Y, Vanwijck R, Vikkula M. 2005. Interferon regulatory factor-6: a gene predisposing to isolated cleft lip with or without cleft palate in the Belgian population. Eur J Hum Genet 13:1239-1242.

Houdayer C, Bonaiti-Pellie C, Erguy C, Soupre V, Dondon MG, Burglen L, Cougoureux E, Couderc R, Vazquez MP, Bahuau M. 2001. Possible relationship between the van der Woude syndrome (vWS) locus and nonsyndromic cleft lip with or without cleft palate (NSCL/P). Am J Med Genet 104:86-92.

Ingraham CR, Kinoshita A, Kondo S, Yang B, Sajan S, Trout KJ, Malik MI, Dunnwald M, Goudy SL, Lovett M, Murray JC, Schutte BC. 2006. Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6). Nat Genet 38:1335-1340.

Kleinjan DA, van Heyningen V. 2005. Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet 76:8-32.

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, de Lima RL, Daack-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, Lammer EJ, Aylsworth AS, Ardinger HH, Lidral AC, Pober BR, Moreno L, Arcos-Burgos M, Valencia C, Houdayer C, Bahuau M, Moretti-Ferreira D, Richieri-Costa A, Dixon MJ, Murray JC. 2002. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. Nat Genet 32:285-289.

Lander E, Kruglyak L. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241-247.

Li Q, Lu Q, Hwang JY, Buscher D, Lee KF, Izpisua-Belmonte JC, Verma IM. 1999. IKK1-deficient mice exhibit abnormal development of skin and skeleton. Genes Dev 13:1322-1328.

Marazita ML, Murray JC, Lidral AC, Arcos-Burgos M, Cooper ME, Goldstein T, Maher BS, Daack-Hirsch S, Schultz R, Mansilla MA, Field LL, Liu YE, Prescott N, Malcolm S, Winter R, Ray A, Moreno L, Valencia C, Neiswanger K, Wyszynski DF, Bailey-Wilson JE, Albacha-Hejazi H, Beaty TH, McIntosh I, Hetmanski JB, Tuncbilek G, Edwards M, Harkin L, Scott R, Roddick LG. 2004. Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35. Am J Hum Genet 75:161-173.

Murray JC. 2002. Gene/environment causes of cleft lip and/or palate. Clin Genet 61:248-256.

Radhakrishna U, Ratnamala U, Gaines M, Beiraghi S, Hutchings D, Golla J, Husain SA, Gambhir PS, Sheth JJ, Sheth FJ, Chetan GK, Naveed M, Solanki JV, Patel UC, Master DC, Memon R, Antonarakis GS, Antonarakis SE, Nath SK. 2006. Genomewide scan for nonsyndromic cleft lip and palate in multigenerational Indian families reveals significant evidence of linkage at 13q33.1-34. Am J Hum Genet 79:580-585.

Scapoli L, Palmieri A, Martinelli M, Pezzetti F, Carinci P, Tognon M, Carinci F. 2005. Strong evidence of linkage disequilibrium between polymorphisms at the IRF6 locus and nonsyndromic cleft lip with or without cleft palate, in an Italian population. Am J Hum Genet 76:180-183.

Schäffer AA, Gupta SK, Shriram K, Cottingham RW Jr. 1994. Avoiding Recomputation in Linkage Analysis. Human Heredity 44;225-237.

Schliekelman P, Slatkin M. 2002. Multiplex relative risk and estimation of the number of loci underlying an inherited disease. Am J Hum Genet 71:1369-1385.

Sharma S, tenOver BR, Grandvaux N, Zhou GP, Lin R, Hiscott J. 2003. Triggering the interferon antiviral response through an IKK-related pathway. Science 300:1148-1151.

Stanier P, Moore GE. 2004. Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. Hum Mol Genet 13:R73-81.

Taniguchi T, Ogasawara K, Takaoka A, Tanaka N. 2001. IRF family of transcription factors as regulators of host defense. Annu Rev Immunol 19:623-655.

Zucchero TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S, Schutte BC, Marazita ML, Murray JC. 2004. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. N Engl J Med 351:769-780.