Review on the Role of IRF6 in the Pathogenesis of Non-syndromic Orofacial Clefts

Si Di ZHANG¹, Yue YOU¹, Mei Lin YAO¹, Bing SHI¹, Zhong Lin JIA¹

Non-syndromic orofacial clefts (NSOCs) are the most common craniofacial malformation. In the complex aetiology and pathogenesis of NSOCs, genetic factors play a crucial role and IRF6, located at chromosome 1q32.2, is the best documented NSOC susceptibility gene. IRF6 is a key factor in oral maxillofacial development and known to contribute the most in NSOCs. It is essential to conduct a complete review of the existing results on IRF6 to further understand its role in the pathogenesis of NSOCs. Thus, the present authors summarised the research progress on the mechanism of IRF6 in NSOCs from both genetic and functional perspectives in this review.

Keywords: IRF6, genetics, non-syndromic orofacial cleft (NSOC)


Non-syndromic orofacial clefts (NSOCs) have a global prevalence of around 1/700 and are the most common craniofacial malformation. The aetiology and pathogenesis are very complex, including genetic factors, environmental factors and their interaction, with genetic factors playing a particularly crucial role¹. Syndromic orofacial clefts (SOCs) are usually monogenic or oligogenic disease with the characteristics of pedigree inheritance. In contrast, the genetic aspect of NSOCs presents polygenic characteristics.² Some of the genes may play a major role, whereas most play a minor role, resulting in the occurrence of NSOCs after accumulating to a certain extent.

The aetiology of NSOCs has been investigated using a variety of methods, including linkage analysis, candidate gene association studies, cytogenetics, animal models and expression studies. Since it was first found to be associated with NSOC in 2004², IRF6 has been recognised as the most contributing and well-documented of dozens of NSOC susceptibility genes/loci identified thus far.

As the encoding gene of interferon regulatory factor (IRF) 6, IRF6 is located at chromosome 1q32.2 (Fig 1). It first attracted attention in orofacial cleft studies because it was found to be the major pathogenic gene of Van der Woude syndrome (VWS; OMIM 119300), the most common form of SOC characterised by cleft lip and palate and lip pits.² This research also demonstrated the involvement of IRF6 in orofacial development.² At present, aetiological IRF6 variants have been found in at least 70% of patients with VWS.³ Hypotheses and studies on the association between IRF6 and NSOCs were thus initiated, and remarkable progress has been made.

Not only is IRF6 an important factor for craniofacial development⁴, but it has also been suggested that it might even be associated with the severity of orofacial clefts⁵, hence the critical importance of understanding of the existing research on IRF6 to gain a deeper under-

¹ State Key Laboratory of Oral Diseases & National Clinical Research Center for Oral Diseases & Department of cleft lip and palate, West China Hospital of Stomatology, Sichuan University, Chengdu, P.R. China.

Corresponding author: Dr Zhong Lin JIA, State Key Laboratory of Oral Diseases, West China Hospital of Stomatology, Sichuan University, No.14, 3rd Section, Renmin Nan Road, Chengdu 610041, P.R. China. Tel: 86-28-85503462. Email: zhonglinjia@sina.com

This project was supported by the major frontier issues of the National Science Funds of China (no. 81271118 and no. 81600849).
standing of its function in the mechanism of NSOC. Therefore, the present authors review the research progress of IRF6 in the pathogenesis of NSOC in genetic and functional studies.

**Genetic studies**

Genetic studies have identified many susceptibility variants in the IRF6 gene that might increase the risk of NSOCs under certain environmental conditions, but the potential influence of these variants on different populations and subtypes leads to a significant difference in their association with NSOCs. In this section, we will review related genetic studies according to the progress of research methods at different stages.

**Linkage analysis**

Linkage analysis maps susceptible genes to specific genomic regions by using the relationship between genetic markers (known locations) carried by family members and genetic disease, since susceptible genes associated with the disease are very likely to be linked to genetic markers.6

Zucchero et al2 proposed that there is a significant association between IRF6 and the risk of non-syndromic cleft lip with or without cleft palate (NSCL/P). They found that single nucleotide polymorphism (SNP) rs2235371 was significantly associated with the risk of non-syndromic cleft lip and palate (NSCLP) in Asian and South American populations.2 Located in exon 7 of the IRF6 gene, rs2235371 is involved in encoding valine or isoleucine at the position 274 amino acid (V274I) in the conserved protein binding domain of IRF6 protein.

In 2005, a linkage disequilibrium between rs2013162 and rs2235375 and NSCL/P was found in an Italian population, confirming the contribution of IRF6 to the aetiology of NSCL/P in the southern European population for the first time7, and later in the white population.8 In 2007, Jakobsen et al9 found a 6.5 MB linkage region in the vicinity of IRF6 within 1q32.1 and q32.2, which may contain genes or non-coding region elements that regulate IRF6 and increase the risk of NSCL/P.

Due to locus heterogeneity and limitations in the size and number of large multiple families (two or more family members with NSOC), the lack of repeated studies on linkage analysis makes it difficult to achieve genome-wide significant differences in the results of individual studies. Marazita et al10 conducted a meta-analysis of previous genome-wide linkage studies and found 16 significant linkage regions including 1q32. A subsequent family study identified six chromosomal regions associated with NSOC, of which IRF6 or its adjacent SNPs were the most significant.11 Studies conducted on different populations confirmed the contribution of IRF6 to the aetiology of NSCL/P, suggesting that further study is worthwhile.

**Candidate gene association studies**

Incorporating a large number of sporadic cases without relying on large multiple families, candidate gene association studies were conducted to determine whether the candidate gene was associated with NSOC by comparing the allele frequency differences between NSOC and control samples.12 Considering that the selection of candidate genes is the premise of association study, pathogenic genes of SOC and genes related to cranial and maxillofacial developmental signalling pathways were frequently used as sources for screening candidate genes of NSOC.

IRF6 initially attracted attention in 2004 as the main pathogenic gene of VWS.2 Subsequent studies explored the association between IRF6 and NSOC and obtained remarkable results. The first polymorphism found to be associated with NSOC in IRF6 was rs2235371 and the association was then confirmed in subsequent association studies carried out on different populations.8,13,14 It was also the first polymorphism that was significantly associated with NSCL/P in Asian and American Indian populations15,16 and complete left NSCL/P in the Brazilian Caucasian population.17 In 2008, Suazo et al18 found a linkage imbalance between SNPs rs2235371, rs764093, rs2236909, rs2235375 and NSCL/P in the Chilean population. The rs2235371 CT and CT/TTP genotypes were associated with a significant reduction
in risk. Larrabee et al proposed that rs2235371 may play a role in the pathogenesis of NSOC by affecting the function of IRF6 protein, and confirmed this association was caused by common rather than rare variants; however, the SNP characteristics of rs2235371 and its effect on protein function still need further exploration, and no association was found between rs642961 and NSOC.

Located in an enhancer of approximately 10,000 base pairs upstream of IRF6, rs642961 (G > A) may interfere with IRF6 expression by disrupting the binding site of AP-2alpha gene which is usually the enhancer of IRF6. Rahimov et al found the correlation between this SNP and NSOC in a European population first, and the association was more significant in non-syndromic cleft lip only (NSCLO). Brito et al confirmed that rs642961 was weakly associated with NSCLO in a Brazilian population, but found no association between rs642961 and IRF6 transcription level in oribicularis oris muscle mesenchymal stem cells (MSCs), suggesting that it may function during a specific period of embryogenesis.

Some studies proposed that A allele at rs642961 was associated with NSCL/P. Pan et al found that the AG and AG/AA genotypes of rs642961 were associated with increased risk of NSOC, especially NSCL/P and NSCLP. These findings require further confirmation in subsequent studies; however, studies of different populations or even the same ethnic population may achieve the opposite result. A population study demonstrated a lack of association between the A allele of rs642961 and Swedish NSCLO subset; however, another study suggested that they were associated with NSOC in a Brazilian population and that the G/A haplotype increased the risk for both children and mothers. The complexity of the ethnic mix of the Brazilian population appears to be an important confounding factor for this difference.

Located 27 bp downstream of exon 6 of IRF6, rs2235375 may alter the splicing process with functional annotations, suggesting that it is associated with the decrease in IRF6 expression. It was positively correlated with NSCL/P in Italy, Belgium, China, Norway, Chile, and Brazil. C allele at rs2235375 was associated with NSCL/P risk in the Chilean population, and its biological role at the level of craniofacial development needs further investigation; however, in a study of a south Indian population, rs2235375 was associated with an increased risk of non-syndromic cleft palate only (NSCPO) rather than NSCL/P.

The association between NSCL/P and rs2013162 (Ser153Ser) (which is a silent variant located in exon 5 of IRF6) was revealed in Belgian and Italian populations, respectively, in 2005. Xu et al showed that both rs2013162 and rs2235371 were closely associated with an increased risk of NSCL/P in northeast China; however, neither Birnbaum et al nor Huang et al found a significant association between rs2013162 and NSOC, which may be mainly related to population heterogeneity.

Studies on some susceptibility SNPs are limited and the results are controversial, indicating that more validation studies on different population samples are needed. Srichomthong et al found that IRF6 G820A was significantly associated with NSCL/P in a Thai population and was responsible for 16.7% of the genetic contribution to NSCL/P. Tang et al confirmed that A allele at G820A may increase the risk of NSCPO in China, and proposed the possibility of linkage disequilibrium between rs2235371 and other pathogenic variants. In 2009, an association study conducted in western China suggested that rs2235373, rs4844880 and rs2073485 were significantly associated with NSCL/P, whereas rs599021 was associated with NSCPO. These SNPs may be directly involved in transcription or indirectly involved in the replication and transcription of IRF6 gene through regulatory elements, affecting protein functions and thus regulating oral and maxillofacial development. It is also interesting that there was a gene–environment interaction between rs2235373 GG genotype and maternal history of abortion. Although rs590223 (A > G) regulates the transcription level of IRF6 in hepatocytes in vivo, its association with NSOC is still controversial. In the same year, Nikopensius et al conducted a case-control association study on NSCPO in a northeastern European population and found rs17389541, located about 8kb upstream of IRF6, was significantly associated with NSCPO susceptibility.

Several candidate gene association studies have demonstrated that IRF6 interacts with some genes in the mechanism of NSCLP. A gene–gene interaction analysis in 2013 indicated that the combination of rs2073485, rs2235371 or rs2236909 in IRF6 and rs17176643 in PAX9 may increase the risk of NSCL/P.

Some studies combined linkage analyses with candidate gene association studies. Park et al considered the genotype and double risk of specific SNPs in IRF6 first and identified its association with increased risk of NSCL/P in Asian trios. Diercks et al first demonstrated a strong association between rs2235371, rs1856161, rs2235377 and NSOC in a Honduran population which was enhanced in NSCL/P. Rahimov et al proposed that rs2235371 was not directly related to NSOC, and that there might be pathogenic variants of IRF6 regulatory element in the region of strong linkage disequilibrium.
with rs2235371. The contribution of functional variant rs642961 to the pathogenesis of NSCLO was about 18%.21 Blanton et al.40 conducted linkage analysis and association studies in Hispanic and non-Hispanic white NSOC families, respectively. They found that the association between rs2235371 and NSOC was more significant in non-Hispanic white single families, whereas the association between rs642961 and NSOC was not confirmed.40 Jugessur et al.26 confirmed that rs2235371 and rs2013162 were associated with NSOC, which has not been widely confirmed.24,41 Craniofacial measurements showed that rs2235371, rs2013162 or other pathogenic variants within the linkage disequilibrium region played an important role in the variation of nasolabial soft tissue morphology in a normal range in an East Asian population.42

Although linkage analyses and candidate gene association studies are increasingly effective in identifying common loci and specifying genes involved in complex traits, it is difficult to identify specific pathogenic variants. Convincing statistical evidence and functional data are needed to prove that a specific variant in a linkage disequilibrium region containing many strongly correlated SNPs is a pathogenic mutation.

Genome-wide association study (GWAS)

The results of genetic studies of NSOC in the last 10 years varied widely. Complex heterogeneity and confounding factors make linkage analysis and association studies difficult to replicate. The extensive development of the genome-wide association study (GWAS) has brought new breakthroughs to the genetic research of NSOC.

The GWAS can be used for large-scale groups. By selecting genetic markers at the whole genome level, the GWAS compares the differences of SNP allele frequency between NSOC samples and controls, using genetic analysis to identify NSOC-associated SNPs and susceptibility genes by location, linkage disequilibrium and bioinformatics functional annotation.43 Compared with linkage analysis, the GWAS has lower requirements for research samples and does not require large multiple family samples. Compared with candidate gene association studies, which rely more on the pre-setting of susceptible genes and include a few hundred SNPs at most, GWAS can screen out new susceptible genes and chromosomal regions in the whole genome without prior construction of a hypothesis.44 Over the past decade, GWAS and correlative meta-analysis have reported more than 40 NSOC susceptibility genes, far surpassing traditional family linkage analyses and candidate gene association studies, and furtherly identified IRF6 as a candidate gene for the pathogenesis of NSOC.45

In 2010, Beaty et al.46 conducted a GWAS in Asian and European NSOC trios and found four SNPs: rs2013162, rs2073485, rs861020 and rs10863790, which reached genome-wide significance. The first GWAS in a Chinese population was conducted in 2015 and confirmed that 1q32.2 was associated with NSCL/P.15 CCCTC binding factor (CTCF) chromatin interaction analysis revealed the interaction signal between the regulatory region containing rs2235371 and DNA sequence in multiple cell lines, suggesting that rs2235371 may be involved in chromatin activity. Thus, the role of rs2235371 may be complex and more studies are needed to elucidate its underlying mechanisms.

However, GWAS has limitations: it focuses mainly on common variants and does not involve rare variants. Compared to common variants that have a very limited effect on disease development, rare variants may have a greater effect on the severity and earlier onset of NSOC.47,48 On the other hand, since SNPs are mainly selected from HapMap data and usually influenced by strong linkage disequilibrium, the candidate genes/loci reported by the GWAS may be different from the actual situation. Therefore, the susceptible regions identified by the GWAS need to be located more accurately.

Research progress in the post-GWAS era

It is gradually realised that known genetic susceptibility genes and loci can only explain a small part of the heritability of NSOC since the complexity and high genetic heterogeneity of NSOC and the limited recognition ability of GWAS. There may be undiscovered genetic factors associated with NSOC in the IRF6 region, such as rare variants, gene–environment interactions and epigenetic inheritance, which may play a more important role in the development of NSOC.49 Advances in sequencing technology have resolved this issue.

First-generation sequencing, also known as Sanger sequencing, was developed from the dideoxy terminal termination method and is the gold standard for obtaining nucleic acid sequence information at present; however, it is difficult to carry out on a large scale due to its high operating costs. Next-generation sequencing enables a comprehensive analysis of the genome with the ability to perform massively parallel sequencing at a faster speed and lower cost. Compared with whole genome sequencing, whole exome sequencing (WES) and target region sequencing (TRS) have been more widely used in genetic studies of NSOC due to their economy and efficiency.
Exons and untranslated regions in the human genome only account for 1% to 2% of total sequences, containing up to 85% of disease-related variants. WES screens for pathogenic variants in coding regions that may affect the function of protein products, representing a new breakthrough in the genetic aetiology of NSOC. As it is more efficient, comprehensive and specific than GWAS, WES plays an increasingly important role in the study of the aetiology of genetic diseases with the development of cheap sequencing technology. The combination of WES and GWAS helps to explore the aetiology of NSOC further.

In 2008, Pegelow et al50 sequenced IRF6 exons in Swedish multiple NSCL/P families and no disease-related mutation was detected. In 2018, Zhao et al51 conducted a WES on a Han NSCLP patient and identified a new rare mutation of IRF6 (c.26G > A; p.Arg9Gln) that affected the structure of IRF6 to some extent by causing residue changes. A rare synonymous mutation (p.Ser307Ser; g.209963979, G>A; c.921C>T) was identified as a possible aetiology of NSCL/P in a WES in 2020.52 This mutation is located at exon 7 and may affect the binding of external splicing silencing element to the main splicing regulator. A new IRF6 pathogenic mutation (c.961C > T; p.Val321Met) was detected in a Chinese Han NSCLP family. As a conserved codon in many species, it also caused changes in residues and altered the structure of IRF6 to some extent.53

TRS, also known as targeted sequencing, provides more comprehensive coverage of targeted regions. It is a cost-effective way to obtain comprehensive information outside the coding area.TRS found a laterality difference in IRF6: 26 SNPs were associated with the difference between unilateral and bilateral NSCL/P, which is one of the genetic sources of phenotypic heterogeneity of NSOC.54 Sequencing of nine exons and untranslated regions at the 5’ and 3’ ends of IRF6 in African NSOC patients revealed that 92% of potentially pathogenic exon and splice site mutations occurred in exons 4 and 7.55 TRS of Chinese, Philippine, American and European trios revealed a G × G interaction between IRF6 and MAFB polymorphisms, which was most significant in European trios.55

The susceptibility regions identified by GWAS can be deeply sequenced to mine rare variants and identify functional variants in coding and non-coding sequences. Although TRS can further supplement genetic information in susceptible regions, genome-wide detection can obtain the most complete genomic information and have the potential to fully reveal the molecular changes of NSOC, which is an ideal state for NSOC genetic research; however, large-scale whole genome sequencing studies of NSOC that have been reported so far are still very limited. Identified variants are believed to be only a part of the overall heritability of NSOC and it is thought that there are other common or rare variants in the IRF6 region that have not been discovered yet in relation to maxillofacial development and NSOC aetiology.

One of the strategies of NSOC genetics research in the sequencing era is to effectively identify potential susceptibility sites, clarify the biological mechanisms through functional analysis and verify them in large and independent populations. As sequencing throughput and precision increase and costs decrease, the technology will be used more widely. An important trend in NSOC genetics research in the future may be to use sequencing technology to accurately locate susceptible regions based on further mining GWAS data, detect rare and functional variants more comprehensively and conduct in-depth studies on biological mechanisms.

It is important to note that a study published in 2019 found that infants with NSCL/P had a higher methylation level at IRF6 promoter regions than controls, suggesting that abnormal methylation of the promoter region may contribute to the development of NSCL/P.56

Functional studies

IRF6 contains 10 exons, among which exons 1, 2 and 10 are not involved in protein coding (Fig 1). As a protein-coding gene, the product protein of IRF6 is one of nine members of the family of interferon regulatory factor (IRF) proteins and has two conserved domains: a highly conserved DNA-binding domain and a less conserved protein-binding domain termed Smad-interferon regulatory factor–binding domain (SMIR), both of which are critical for its function.57

IRF6 is involved in the normal development of craniofacial structures by regulating epithelial differentiation and palatal fusion.

IRF6 is a key factor in the development of the lip and palate that plays an important role in craniofacial development and participates in the development of skin and external genitalia. Furthermore, as a key determinant in the proliferation-differentiation conversion of keratinocyte, IRF6 is required for the differentiation of mammalian skin, mammary gland and oral epithelium, and involved in wound healing and migration.

Kondo et al4 detected high expression of IRF6 mRNA in fused palate, tooth bud, hair follicle, genitalia and the
medial edge of skin in mice, and found that the haploid deficiency would impair the normal development of the maxillofacial region and lead to orofacial cleft, confirming that IRF6 is a key factor in the development of the lip and palate and is also involved in the development of the skin and external genitalia.

In 2006, Richardson et al.58 mutated Arg84 to cysteine (R84C) and constructed an Irf6R84C/R84C homozygous mutant mouse model that showed postnatal death. There were significant abnormalities in epidermal development in Irf6+/R84C heterozygous mutant mice: a dense basal layer, less differentiation in the expanded basal layer and no granular layer or keratinised layer, indicating excessive epidermal hyperplasia and an abnormal differentiation process that may lead to epithelial adhesions that block the mouth and cause cleft lip and palate. In the same year, Ingraham et al.59 used a mouse model deficient for Irf6 and observed abnormal skin, limbs and craniofacial morphology (Fig 2). They then conducted a histological and gene expression analysis on deficient mice and found that the main defect of mouse skin was abnormal proliferation and differentiation of keratinocytes, confirming that Irf6 is necessary for regulating keratinocyte proliferation and terminal differentiation.59 In vitro culture experiments supported Irf6 as a necessary condition for keratinocyte differentiation and the authors speculated that epidermal adhesion in the oral cavity of mice was due to the absence of normal keratinocytes, whereas skeletal abnormalities were secondary to defective epidermal differentiation.60

Thompson et al.61 conducted mouse model studies and proposed a different view, suggesting that Irf6 is involved in regulating bone differentiation and mineralisation during craniofacial bone development. In 2021, Girouisi et al.62 used keratinocytes derived from human skin and oral mucosa to construct Irf6 knockout cell lines. They found that keratinocytes lacking Irf6 have poor cohesion, and most of the differentially expressed proteins may be related to differentiation, intercellular adhesion and immune response.62 In 3D skin cultures, loss of Irf6 resulted in severe keratinocyte differentiation defects while cell growth rates remained constant. These results suggested that Irf6 deficiency disrupts epithelial homeostasis by altering the colony morphology, migration pattern and differentiation potential of human keratinocytes.62

Palatal development in vertebrates involves migration of cranial neural crest, fusion of facial processes and extension of cartilaginous framework, while cleft palate is a structural birth defect resulting from palatal dysplasia. Irf6 is expressed in ectodermal fusions of the upper lip and primary palate in mice and chicks, but only in the secondary palate in developing mice.63 Similar expression patterns also exist in human craniofacial structures. Irf6 is expressed in bilateral maxillary process chondrocytes and the maxillary processes need to fuse with the lateral nasal process in the centre. Dougherty et al.64 conducted a detailed analysis of palate development in zebrafish and revealed that Irf6 is involved in palate morphogenesis by regulating the fusion of facial processes.

**Exploration of Irf6’s participation mechanism in epithelial differentiation and palate fusion**

Given that in vitro and in vivo experiments have confirmed that Irf6 plays an important role in regulating epithelial differentiation and palatal fusion in craniofacial development, subsequent studies on the molecular mechanism by which Irf6 functions will help to describe and integrate the molecular pathway of lip and palate morphogenesis further.

Kousa et al.65 constructed a mouse model of Irf6 gene knockout and restored the expression of Irf6 in embryonic basal epithelium. They found that severity of the disease was significantly reduced and the death of mice after birth could be prevented, but the adhesion between the palate and tongue could not be saved completely, suggesting the importance of cell-independent Irf6 expression in peritum.65 In the same year, another study by Kousa et al.66 confirmed the interaction of Irf6 and Spry4 signals in mouse peridermal development. Through microarray analysis and in vivo transplantation, de la Garza et al.67 proposed that the encoding gene of Grhl3 is the direct effector of Irf6 in percutaneous differentiation. In addition, Irf6 protein interacts directly with Nme1 and Nme2 to regulate the availability or localisation of the Nme1/2 complex and the dynamic behaviour of the epithelia during lip and palate development. Missense mutations in Irf6 or Nme can lead to cleft lip and palate by damaging the ability of Irf6 to bind to Nme protein 66.68 Ferretti et al.69 suggested that Irf6 is involved in the conserved Pbx-Wnt-P63-Irf6 regulatory pathway in mammals that controls facial morphogenesis by promoting epithelial apoptosis. In addition, the integration of Irf6 and Jagged2 signalling is essential for controlling palatal adhesion and fusion competence.70

Palatal organ culture experiments showed that Irf6 was involved in the regulation of Tgfb3 on epithelial mesenchymal transformation and palatal fusion during embryonic palatal development.71 This result supported the role of Irf6 in the transforming growth
factor-β (TGF-β) signalling pathway, a fundamental developmental pathway. Considering that deletion of Irf6 can lead to cranial fracture and mandibular dysplasia in mice while mutations of Twist1 gene can lead to cranial fracture, mandibular dysplasia and cleft palate, Fakhouri et al. constructed monozygotic and double heterozygotic Irf6 and Twist1 genes (Irf6+/−; Twist1+/−) mouse embryo models, respectively. Homozygous deletion of Irf6 resulted in multiple bone defects in the mandible and limbs of mice; however, no expression of Irf6 was detected in bone tissue, suggesting the existence of intercellular communication. Although monozygous mice were almost normal, some dizygous mice (Irf6+/−; Twist1+/−) mice developed severe hypoplasia of the mandible, leading to holoacrania and cleft palate. It is suggested that genes encoding IRF6 and Twist1, two transcription factors essential for craniofacial development, may lead to craniofacial diseases through gene–gene interaction. A study in 2019 proposed that IRF6 and TAK1 synergistically promote HIPK2 activation and stimulate apoptosis during palatal fusion. IRF6 can also regulate the expression of ESRP1, which overlaps in mouse oral-facial skin and zebrafish periderm, nasofrontal ectoderm and oral epithelium, and controls vertebrate midfacial morphogenesis through Irf6-ESRP1/2 regulatory axis. Another study in 2020 indicated that SPECC1L cytoskeletal protein may function downstream of IRF6 in palatogenesis.

In addition, mice carrying both p63 heterozygous deletion and Irf6 knockout mutation R84C showed abnormal ectoderm development, which resulted in cleft palate. In vitro culture of primary keratinocytes from patients with cleft palate and functional studies found that P63 is a key regulatory molecule in palate development and binds to the upstream enhancer element to transactivate IRF6. Mutations in this enhancer element are associated with increased susceptibility to cleft palate. In mouse and human oral and maxillofacial development, p63 and IRF6 act within a regulatory ring that coordinates epithelial cell proliferation and differentiation during normal palatal development. IRF6 or p63 mutations cause this ring to break, resulting in fissure. On the other hand, Kousa et al. found that IRF6 and AP2A interact to regulate epidermal development, and rs642961 located in the enhancer element interferes with its binding with AP2A protein, affecting epithelial cell development and increasing the risk of NSOC. They proposed that P63 and AP-2α may play a synergistic role in IRF6 regulation.

Conclusions and future research

Numerous genetic and functional studies have provided sufficient evidence for the contribution of IRF6 in the pathogenesis of NSOC; however, studies have been conducted on different ethnic groups and population strati-
fication has become a major source of confounding factors in genetic studies, leading to possible differences in the results of studies of polymorphisms in different ethnic groups. In addition, although many polymorphisms in IRF6 were associated with NSOC in genetic studies, it is still uncertain which of them are pathogenic due to the existence of linkage disequilibrium. The inclusion of other risk factors and genes would also be useful for the improvement of disease prediction models and occurrence risk assessment of NSOC.

**Conflicts of interest**

The authors declare no conflicts of interest related to this study.

**Author contribution**

Dr Si Di ZHANG contributed to the literature collection, draft and revision; Dr Yue YOU contributed to the conception and draft; Dr Mei Lin YAO contributed to the draft; Dr Bing SHI contributed to supervision; Dr Zhong Lin JIA contributed to the conception, supervision and manuscript revision. All authors read and approved the final manuscript.

(Received Jun 05, 2023; accepted Oct 23, 2023)

**References**


