Gene 635 (2017) 69–76

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

A novel G to A transition at initiation codon and exon-intron boundary of PAX9 identified in association with familial isolated oligodontia

Tanmoy Sarkar[a], Rajesh Bansal[b], Parimal Das[a,*]

[a] Centre for Genetic Disorders, Institute of Science, Banaras Hindu University, India
[b] Faculty of Dental Sciences, Institute of Medical Sciences, Banaras Hindu University, Varanasi 221 005, India

ARTICLE INFO

Keywords:
Initiation codon
Haplotype
PAX9 minigene
Haploinsufficiency
Oligodontia

ABSTRACT

Several studies on experimental animals indicate that the process of organogenesis crucially depends upon the spatiotemporal dose of certain critical bio-molecules. Tooth development is also not an exception. While most of the knowledge regarding the molecular mechanism of tooth development comes from the studies on mouse model, pathogenic variations identified in human tooth agenesis also provide valuable information on mammalian tooth development. Until now five major candidate genes have been identified for tooth agenesis in human. Among them, *PAX9* plays the crucial role in tooth development and in non-syndromic congenital tooth agenesis. In this study, microsatellite and SNP based genotyping identifies a disease specific haplotype block, which includes *PAX9* gene, segregates with autosomal dominant tooth agenesis phenotype. Direct sequencing of *PAX9* identifies a novel heterozygous G to A transition at the third base (c.3G > A) of initiation codon leading to ATG to ATA shift in all affected individuals which is absent in all unaffected relatives and 200 control chromosomes. Further, *in vitro* functional analysis creating *PAX9* minigene construct did apparently show no effect on the splice-site migration. It is therefore proposed that haploinsufficiency of PAX9 is the causal factor for tooth agenesis in this family.

1. Introduction

Congenital tooth agenesis (CTA) is one of the most common craniofacial developmental disorders with known (OMIM ID # 106600, # 604625, # 150400, # 616724, # 617073, # 617275 and # 313500) and unknown (OMIM ID % 602639 and % 610926) genetic aetiology. CTA may be associated with a number of malformation syndromes or may be nonsyndromic affecting only tooth organ. Depending upon the population prevalence of tooth agenesis varies from 2.6% to 11.3% and females show significantly higher frequency than males in almost all studied population (Polder et al., 2004; De Coster et al., 2009). Among different tooth type, third molar agenesis is most common (> 20%) followed by mandibular second premolar (2.91–3.22%), maxillary lateral incisor (1.55–1.78%), maxillary second premolar (1.39–1.61%). On the other hand, maxillary central incisor is least frequently missing (0.00–0.01%) followed by mandibular first molar (0.00–0.02%), and mandibular canine (0.01–0.03%). Depending upon the number of missing teeth, tooth agenesis can be classified as hypodontia (< 6 missing teeth), oligodontia (≥ 6 missing teeth) and anodontia (missing of all teeth), those entire categories exclude 3rd molars, the Wisdom teeth. Recent study in Japanese population indicates that sibling recurrence risk for hypodontia is 24.5% (95% CI: 13.8–38.3%) whereas for oligodontia it is 43.8% (95% CI: 26.4–62.3%) (Machida et al., 2015). The first gene identified to be associated with non-syndromic CTA was *MSX1* (Vastardis et al., 1996), however, after identification of *PAX9* associated with tooth agenesis, it becomes one of the major candidate genes for familial non-syndromic CTA (Stockton et al., 2000; Abid et al., 2017; Liang et al., 2016). Subsequently, more genes namely *AXIN2* and *WNT10A* were identified as candidate genes for tooth agenesis (Lammi et al., 2004; Kantaputra and Sripathomsawat, 2011). Recently *WNT10A* has been considered as the major candidate gene for sporadic tooth agenesis (van den Boogaard et al., 2012; Mostowska et al., 2013). *EDA*, which was mostly associated with Ectodermal Dysplasia syndrome, recently reported being associated with non-syndromic CTA (Tao et al., 2006; Tarpey et al., 2007; Sarkar et al., 2014).

The highest number of mutations that have been found to be associated with non-syndromic tooth agenesis is mapped on *PAX9* gene. *PAX9* is a member of protein family codes for a group of transcription factors with a characteristic paired domain for DNA binding. PAX family of genes are critical for the development of several vertebrate organs (Lang et al., 2007). Homozygous *PAX9* null mouse dies soon after birth and showed arrested tooth development at bud stage along

* Corresponding author at: Centre for Genetic Disorders, Institute of Science, Banaras Hindu University, Varanasi 221 005, India.
E-mail addresses: parimal@bhu.ac.in, hellow_parimal@yahoo.com (P. Das).

http://dx.doi.org/10.1016/j.gene.2017.08.020
Received 29 June 2017; Accepted 23 August 2017
Available online 25 August 2017

with abnormalities of pharyngeal pouch derivatives and limb. This indicates that *PAX9* is essential for survival and also for the development of craniofacial organs and teeth in the mouse. Several mutations right from haploinsufficiency to loss of function of *PAX9* have been reported to be associated with non-syndromic tooth agenesis in human, which established the role of *PAX9* during tooth development (Stockton et al., 2000; Nieminen et al., 2001; Das et al., 2002; Frazier-Bowers et al., 2002; Das et al., 2003).

During the process of tooth development *PAX9* along with *MSX1* plays an essential role in the establishment of odontogenic potential from dental epithelium to dental mesenchyme. Both the genes expressed in dental mesenchyme and interact with each other to activate one of their known downstream targets Bmp4. This subsequently establishes enamel knot, a major signalling centre that controls further tooth development.

In this present study, we have identified and characterised a novel pathogenic variation in *PAX9* responsible for the autosomal dominant type of non-syndromic oligodontia in an Indian family.

2. Materials and methods

2.1. Patient family identification and control sampling

All the affected subjects and their family members participated in this study belong to an Indian family (DEN 11). Clinical examination of teeth, hair, skin and nails were carried out to identify any abnormalities in those ectodermal organs. Information regarding sweating activity and thermotolerance were collected through interview. All those clinical examinations were performed by a dental practitioner from Faculty of Dental Sciences, Sir Sundar Lal Hospital, Banaras Hindu University. Written informed consent was obtained from all the members who were enrolled in this study prior to collect Orthopantomograph (OPG) and 8 ml of peripheral venous blood in a heparinized syringe. Affection status was confirmed and categorised using OPGs. The Blood sample from unaffected, unrelated, healthy control individuals was collected from blood bank of Sir Sundar Lal Hospital, Banaras Hindu University after examining their teeth. All the study protocols and subject consent were approved by Institutional Review Board of Banaras Hindu University (No. Dean/2008-09/484).

2.2. DNA isolation and targeted linkage analysis

Genomic DNA from affected and unaffected individuals were isolated from peripheral venous blood samples using the standard protocol described by Grimberg et al. (1989) with few modifications (Grimberg et al., 1989). To determine the association between locus underlying non-syndromic CTA and known candidate genes, targeted linkage analysis was performed using genotype of microsatellite markers flanking each of the three major candidate genes *viz. PAX9* (D14S275, D14S70, D14S288 and D14S276), *WNT10A* (D2S325, D2S2382, D2S126 and D2S396) and *AXIN2* (D17S787, D17S944, D17S949 and D17S785) for 8 affected and 4 unaffected individuals of the DEN11 family. All microsatellite primers were selected from ABI Prism Linkage Mapping Set v2.5 MD10 panel, PCR-amplified and resolved using ABI 3130 Genetic Analyzer according to manufacturer's protocol. Markers were genotyped using ABI Gene Mapper® software V4. Two-point parametric linkage analysis was performed using FastLink v4.1 program of easyLINKAGE Plus v5.00 software package. Marshfield map (sex-averaged) were used to calculate LOD score using dominant mode of inheritance, disease allele frequency 0.01% and penetrance 95%.

2.3. Variation screening and haplotype analysis

To identify pathogenic variation by direct DNA sequencing, all the exons and exon-intron boundaries of *PAX9* and *MSX1* were amplified using polymerase chain reaction (PCR) from genomic DNA of 8 affected 4 unaffected individuals. PCR conditions and primers were described previously (Sarkar et al., 2014). Unincorporated primers and dNTPs were removed with exonuclease I and recombinant Shrimp alkaline phosphatase (rSAP) (USB Affymetrix, USA) treatment. Purified PCR products were labelled with ABI Big Dye Terminator V3.1 cycle sequencing kit followed by capillary electrophoresis and automated base calling using ABI 3130 Genetic Analyzer and Sequencing analysis software V5.2 (Applied Biosystems, USA) respectively according to manufacturer's protocol. Any sequence mismatch was detected by aligning sequencing reads with available National Center for Biotechnology Information (NCBI) Gen Bank DNA sequence database using NCBI-Basic Local Alignment Search tool (BLAST). Other than variation screening genotype of 5 SNP markers (rs2073244, rs2073245, rs2073246, rs12881240 and rs4904210) located within *PAX9* gene and genotype of 4 microsatellite markers (D14S275, D14S70, D14S288 and D14S276), flanking *PAX9* were used for haplotype analysis.

2.4. Genotyping of c.3G > A in affected and control individuals

c.3G > A variation identified in *PAX9* was verified by direct DNA sequencing in all affected and unaffected family members of this family. To cross verify this change DNA fragment including variant locus was amplified using Q5 high-fidelity DNA polymerase (NEB, USA) from heterozygous affected individual followed by adenylation using Taq DNA polymerase. Adenylated fragment was cloned into the pGEM-T-easy vector (Promega) using manufacturer's protocol. Plasmid DNA was isolated from 10 colonies selected at random and sequenced using Sanger method as mentioned above. Further, this variation destroys restriction site of *Bsr*DI (5′GCAATGNN3′). Primers (5′*PAX9*-16F 5′CCACGTTGCTGCTTAGATTG3′ and 5′*PAX9*-16R 5′CTGACCCTTAGCG TGTTCCT3′) were designed to amplify a 392 bp DNA fragment. The amplicon was digested with 3 U of *Bsr*I for 1 h at 65 °C and resolved in 3% agarose gel containing 0.5 μg/ml of EtBr. Amplicon containing wild type allele (5′…GCAATGGG…3′) was digested into 225 bp and 167 bp fragment while variant allele (5′…GCAATAGG…3′) remains undigested. 100 individuals within 25 to 30 years of age (18 females and 82 males) of same ethnic background were examined for this potentially novel variation using the above PCR-RFLP method or by direct sequencing.

2.5. PAX9 minigene construction

c.3G > A variation is located on the exon-intron boundary of exon 2 and intron 2 and may have affect splicing activity. To study this, a 1561 bp long DNA fragment containing 270 bp from exon 2, whole intron 2 (803 bp) and 488 bp from exon 3 was amplified from heterozygous affected individual using Q5 high fidelity DNA polymerase (NEB, USA) and following primers: P1 (forward), 5′-TACAAGGACGACGATGACAAGCCAGTGAGTGATAGACGGAGC-3′ with partial FLAG tag flanking at 5′ side and P3 reverse 5′-TAATATC**TCTAGA**GCTGGGGTACGAGTAGATGT-3′ with *Xba*I site (Bold) at 3'end. First round PCR product was re-amplified with P2 (forward) primer 5′-TAATAT**AAGCTT**GCCATGGACTACAAGGACGA CGATGACAAG-3′ having complete FLAG tag and *Hin*dIII site (Bold) and P3 (reverse primer). Both the PCRs were performed in ABI verity 96 well thermal cycler (Applied Biosystems, USA) programmed with initial denaturation at 98 °C for 30 s, followed by 35 cycles consists of denaturation at 98 °C for 10 s and annealing-extension at 72 °C for 2 min, and one final extension at 72 °C for 10 min. Second round PCR product was adenylated by incubating at 72 °C for 30 min in presence of 200 μM of dATP and Taq DNA polymerase. Adenylated PCR product was gel extracted and ligated with pGEM-T-easy vector (Promega) using manufacturer's protocol and *E. coli* DH5α cells were transformed with recombinant plasmid DNA. From positive colonies, Plasmid DNA was extracted and confirmed by restriction endonuclease digestion with *Bsr*DI followed by Sanger sequencing. To study the splicing activity of