Journal of Oral Biology and Craniofacial Research 13 (2023) 453–460

Contents lists available at ScienceDirect

Journal of Oral Biology and Craniofacial Research

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

# Missense polymorphisms potentially involved in mandibular prognathism

Amin Kalmari [a], Abasalt Hosseinzadeh Colagar [a,*], Mohammadkazem Heydari [a], Valiollah Arash [b]

[a] Department of Molecular and Cell Biology, Faculty of Science, University of Mazandaran, Babolsar, PC:47416-95447, Mazandaran, Iran
[b] Department of Orthodontics, School of dentistry, Babol University of Medical Sciences, Babol, PC: 47176-47745, Mazandaran, Iran

ARTICLE INFO

*Keywords:*
*DUSP6*
*FBN3*
Mandibular prognathism
Missense polymorphisms
*PLXNA2*

ABSTRACT

*Objective:* The current study aimed to identify and analyze missense single nucleotide polymorphisms (SNPs) that can potentially cause mandibular prognathism.
*Methods:* After reviewing the articles, 56 genes associated with mandibular prognathism were identified and their missense SNPs were retrieved from the NCBI website. Several web-based tools including CADD, PolyPhen-2, PROVEAN, SNAP2, PANTHER, FATHMM, and PON-P2 were used to filter out harmful SNPs. Additionally, ConSurf determined the level of evolutionary conservation at positions where SNPs occur. I-Mutant2 and MUpro predicted the effect of SNPs on protein stability. Furthermore, to investigate the structural and functional changes of proteins, HOPE and LOMETS tools were utilized.
*Results:* Based on predictions in at least four web-based tools, the results indicated that *PLXNA2-rs4844658*, *DUSP6-rs2279574*, and *FBN3-rs33967815* are harmful. These SNPs are located at positions with variable or average conservation and have the potential to reduce the stability of their respective proteins. Moreover, they may impair protein activity by causing structural and functional changes.
*Conclusions:* In this study, we identified *PLXNA2-rs4844658*, *DUSP6-rs2279574*, and *FBN3-rs33967815* as potential risk factors for mandibular prognathism using several web-based tools. According to the possible roles of PLXNA2, DUSP6, and FBN3 proteins in ossification pathways, we recommend that these SNPs be investigated further in experimental research. Through such studies, we hope to gain a better understanding of the molecular mechanisms involved in mandible formation.

## 1. Introduction

Mandibular prognathism is a type of skeletal abnormality that can be diagnosed by analyzing facial profile and soft tissue relationships. It is characterized by excessive growth of the lower jaw, resulting in concave facial profile, frontal teeth negative overjet, and improper lip contact.[1,2] Mandibular prognathism seems to be the main cause of skeletal Class III malocclusion and can be recognized at an early age. As the condition becomes more apparent with growth, patients often require orthodontic and surgical treatments.[3] Some negative effects of this abnormality in a patient's life include low masticatory efficiency, impaired speech expression, decreased self-confidence, unpleasant profile, and severe psychological handicap.[4] The prevalence of this disorder varies across different ethnic populations, with less than 1% of Caucasians affected, while about 15% of Asians, particularly Chinese and Japanese, are affected.[5] The occurrence of mandibular prognathism may be influenced by environmental, epigenetic, and genetic factors. Enlarged tonsils, nasal breathing difficulties, posture, habitual head position, endocrine disturbances, trauma, and instrumental deliveries are among the environmental factors that can play an important role in the development of this disorder.[6] To date, various genetic linkage analyses and genome-wide association studies have reported a correlation between numerous genes and loci with mandibular prognathism. Some of these genes are *MATN1*, *COL2A1*, *FGFR2*, *PLXNA2*, *ARHGAP21*, *ADAMTS1*, *GHR, MYO1H, FBN3,* and *DUSP6*.[7] Many of these genes are involved in mandible formation by affecting osteogenesis. During intramembranous ossification, the mesenchymal tissue in the intermediate portion of the mandible is directly transformed into bone, whereas the distal region of the mandible is formed through endochondral ossification, another bone-forming process during embryonic development.[8] Endochondral

* Corresponding author.
*E-mail addresses:* a.kalmari@yahoo.com (A. Kalmari), ahcolagar@umz.ac.ir, acolagar@yahoo.com (A. Hosseinzadeh Colagar), mohamadkazem_hs@yahoo.com (M. Heydari), v.arash@mubabol.ac.ir (V. Arash).

https://doi.org/10.1016/j.jobcr.2023.05.007
Received 13 August 2022; Received in revised form 18 March 2023; Accepted 10 May 2023

ossification involves the initial differentiation of mesenchymal cells into chondrocytes, followed by the eventual replacement of the cartilage by bone tissue.[9] The distinct patterns of ossification in different regions of the mandible emphasize the complex nature of mandibular development and the importance of a precise genetic and molecular control of this process.[10] Therefore, identifying genetic factors and comprehending their role in mandible formation can aid in the accurate diagnosis of mandibular prognathism, while also enabling the advancement of novel treatment approaches for this disorder.[11]

Single nucleotide polymorphisms (SNPs) are single base changes in a DNA sequence that have a significant allelic frequency (≥1%) in the human population. They are one of the most popular genetic markers to identify the association between a particular gene and a certain disease.[12] Among the various types of SNPs, missense ones are very important. They change the codon region of a specific gene, resulting in a new amino acid being inserted into the protein sequence. This new residue may have unique properties that potentially impact the overall structure and function of the protein.[13]

To date, no computational studies have been conducted to identify harmful missense SNPs in genes associated with mandibular prognathism. Therefore, the aim of this study was to use various web-based tools to identify and report SNPs with disease-causing potential.

## 2. Methodology

### 2.1. Identification of genes & SNPs retrieval

All of the genes associated with mandibular prognathism were extracted from two review articles.[7,14] Information on genes and SNPs (SNP ID, alleles, and global allelic frequency) was obtained from the National Center for Biotechnology Information (NCBI) website (https://www.ncbi.nlm.nih.gov/). For this study, only missense SNPs with a global allelic frequency above 0.1 in the 1000 genomes project were screened. The screened SNPs were analyzed using various web-based tools, as shown in Fig. 1.

### 2.2. Finding harmful SNPs

Seven different web-based tools were used to filter out harmful missense SNPs. Combined Annotation Dependent Depletion (CADD) is a tool that provides a detailed analysis of SNPs along with a "Phred score" value (https://cadd.gs.washington.edu/). A score of 20 or more predicts that the analyzed SNP may be among the top 1% of deleterious polymorphism in the human genome.[15] Polymorphism Phenotyping v2 (PolyPhen-2) uses structural and comparative evolutionary considerations to predict the impact of missense SNPs on the function of human proteins (http://genetics.bwh.harvard.edu/pph2/). PolyPhen-2 scores can be interpreted as 0.0 to 0.15 for benign SNPs and 0.15 to 1 for damaging SNPs.[16] Protein Variation Effect Analyzer (PROVEAN) predicts changes in a protein's biological function upon missense SNP (http://provean.jcvi.org/index.php). A polymorphism score of equal to or below −2.5 in PROVEAN is generally considered "deleterious".[17] SNAP2 (https://rostlab.org/services/snap2web/) is a tool based on a machine learning device that predicts the effect of SNPs on protein function by taking into account a variety of sequences and variant features.[18] Protein Analysis Through Evolutionary Relationships (PANTHER) website has a tool that employs a related but distinct metric based on "evolutionary preservation" for predicting missense SNPs (http://www.pantherdb.org/). Polymorphisms that may play a causal role in human disease are considered "damaging" in this tool.[19] Functional Analysis through Hidden Markov Models (FATHMM) uses a species-independent method with optional species-specific weightings for prediction the functional effects of missense SNPs (http://fathmm.biocompute.org.uk/inherited.html). This tool categorizes SNPs into two groups, namely "neutral" and "damaging", based on their disease-causing potential.[20] PON-P2 (http://structure.bmc.lu.se/PON-P2/) utilizes amino acid features, Gene Ontology annotations, and evolutionary conservation to classify missense SNPs into pathogenic, neutral, or unknown groups.[21]

### 2.3. Stability prediction of SNPs on proteins

To predict the effect of SNPs on protein stability, two web-based

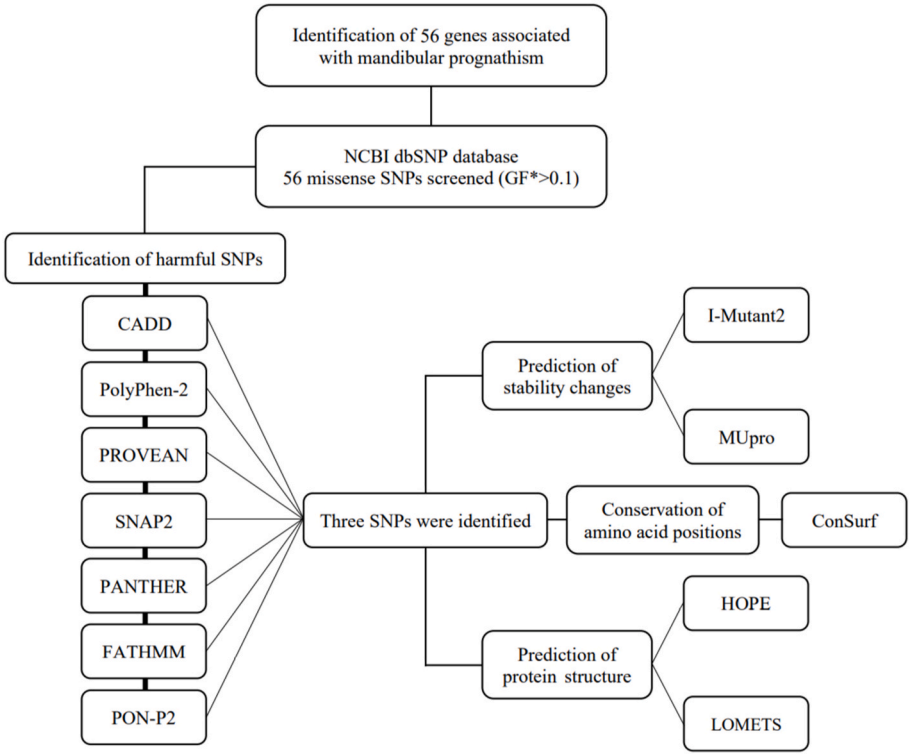


**Fig. 1.** Diagrammatic representation of methodology (*global frequency).

tools, I-Mutant2 (https://folding.biofold.org/i-mutant/i-mutant2.0.html) and MUpro (http://mupro.proteomics.ics.uci.edu/) were used. I-Mutant2 is based on the Support Vector Machine that predicts the changes in the stability of a protein, upon missense SNPs. This tool provides a Reliability Index (RI) as an output, which ranges from 0 to 10, with 10 representing the highest level of reliability.[22] MUpro can estimate the changes in protein stability resulting from missense SNPs. If this tool predicts delta G (ΔG) as negative, it suggests that the analyzed SNP may destabilize the protein.[23]

### 2.4. Conservation analysis

ConSurf (https://consurf.tau.ac.il/) calculates the level of evolutionary conservation of amino acids at a specific position and gives the output as a score ranging from one to nine. A score in a range of 1–3 indicates variable conservation, 4–6 indicates average conservation, and a score in a range of 7–9 signifies the highly conserved positions of amino acids.[24]

### 2.5. Structural analysis of proteins

To investigate the impact of SNPs on the structure and 3D modeling of proteins, two web-based tools were used. The first tool, HOPE, collects structural information from several sources and combines them to build a report (https://www3.cmbi.umcn.nl/hope/). It predicts and characterizes the structural effects of amino acid substitutions on protein domains.[25] The second tool, Local meta-threading-server (LOMETS), is a meta-server method for protein structure prediction and function annotation (https://zhanggroup.org/LOMETS/). The 3D models are built using high-scoring target-to-template alignments from nine threading programs.[26] The wild and mutant models were generated by LOMETS, and the structures were visualized using UCSF CHIMERA software.

## 3. Results

### 3.1. Genes and SNPs selection

According to the detailed examination of two review articles (mentioned in Section 2.1), 56 genes associated with mandibular prognathism were identified and selected. However, we excluded collagen-family genes (*COL1A1* and *COL2A1*) in the present study based on our previous report.[27] By analyzing all these genes in the NCBI dbSNP database, 56 missense SNPs were screened in 27 different genes (Table 1). Missense SNPs with a global allelic frequency above 0.1 were not found in the following genes: *MATN1*, *KAT6B*, *EPB41*, *FOXO3*, *HDAC4*, *HOXC*, *LTBP2*, *MMP13*, *TBX5*, *TGFB3*, *RUNX2*, *PTGS2*, *ADAMTSL1*, *CALN1*, *FGF12*, *FGF20*, *FGF3*, *FGFR1*, *JAG1*, *NBPF8*, *NUMB*, *RORA*, *SMAD6*, *TCF21*, and *WNT3A*. Although the NCBI dbSNP database reports some SNPs in *PSEN2*, *MYH1*, MYH2, and *MYH7* genes as missense, their global frequencies suggest otherwise.

### 3.2. Identification of harmful SNPs

All screened SNPs were analyzed using seven web-based tools as shown in Table 2. The terms *deleterious*, *damaging*, *effect*, and *pathogenic* were considered to be indicative of harmful SNPs that can affect the structure and function of the protein. Conversely, the terms *neutral*, *benign*, and *tolerated* indicated SNPs that were not suitable for this study. Out of 56 SNPs, five cases were identified as harmful in more than three web-based tools. After identifying three suitable SNPs in *FBN3* gene (rs33967815, rs12150963, and rs7245429), we decided to further analyze the most harmful ones, but this does not mean that the other two SNPs are unimportant. In total, according to Table 2, *PLXNA2-rs4844658, DUSP6-rs2279574*, and *FBN3-rs33967815* were selected for further analysis using other tools.

### 3.3. Effect of SNPs on proteins stability

The analysis of three SNPs, *PLXNA2-rs4844658*, *DUSP6-rs2279574*,

**Table 1**
Global frequency and allelic change of 56 screened SNPs.

| Gene | SNP(s) | Allele(s) | Global frequency[a] | Gene | SNP(s) | Allele(s) | Global frequency[a] |
|---|---|---|---|---|---|---|---|
| *KRT7* | rs6580870 | A > G | A = 0.1280; G = 0.8720 | *NBPF9* | rs200319336 | C > A, G, T | C = 0.8934; T = 0.1066 |
| | rs2608009 | G > A, C, T | G = 0.1078; C = 0.8922 | | rs71231749 | A > C | A = 0.8630; C = 0.1370 |
| *MYO1H* | rs11611277 | C > A, T | C = 0.8450; A = 0.1550 | | rs1227241524 | A > C, T | A = 0.8618; T = 0.1382 |
| | rs3825393 | T > C, G | T = 0.2953; C = 0.7047 | *NCOR2* | rs2229840 | C > G, T | C = 0.8391; T = 0.1609 |
| *PLXNA2* | rs4844658 | T > C | T = 0.8814; C = 0.1186 | *NOTCH3* | rs35769976 | C > A, G, T | C = 0.8898; G = 0.1102 |
| | rs3748735 | C > T | C = 0.8604; T = 0.1396 | | rs1044009 | G > A, C | G = 0.3706; A = 0.6294 |
| | rs2782948 | C > A, T | C = 0.6354; T = 0.3646 | *NOTCH4* | rs915894 | T > G | T = 0.6012; G = 0.3988 |
| *SSX2IP* | rs1057746 | A > C, G, T | A = 0.4874; G = 0.5126 | | rs520692 | T > C | T = 0.7322; C = 0.2678 |
| *ADAMTS1* | rs428785 | C > A, G, T | C = 0.3109; G = 0.6891 | | rs422951 | T > C | T = 0.6569; C = 0.3431 |
| *ALPL* | rs3200254 | T > A, C, G | T = 0.7330; C = 0.2670 | *FGFR2* | rs755793 | A > G | A = 0.8744; G = 0.1256 |
| *ARHGAP21* | rs3748222 | T > C | T = 0.5617; C = 0.4383 | *FGF23* | rs7955866 | G > A | G = 0.8526; A = 0.1474 |
| | rs1127893 | C > G, T | C = 0.5064; G = 0.4926 | *MYH3* | rs2285477 | C > T | C = 0.4135; T = 0.5865 |
| *DUSP6* | rs2279574 | C > A, G, T | C = 0.5337; A = 0.4663 | *MYH8* | rs8069834 | A > G, T | A = 0.5146; G = 0.4854 |
| | rs770087 | A > C | A = 0.7825; C = 0.2175 | *RASA2* | rs34558081 | T > A, G | T = 0.8938; G = 0.1062 |
| *EVC* | rs6414624 | T > A, C | T = 0.2552; C = 0.7448 | *HSPG2* | rs3736360 | C > T | C = 0.8239; T = 0.1761 |
| | rs2302075 | C > A, G, T | C = 0.1554; A = 0.8446 | | rs2291827 | G > A | G = 0.8478; A = 0.1522 |
| | rs2291157 | A > C | A = 0.8958; C = 0.1042 | | rs897471 | G > A, C | G = 0.3085; A = 0.6915 |
| | rs1383180 | G > A, T | G = 0.6985; A = 0.3015 | *IGF1* | rs35767 | A > C, G, T | A = 0.3037; G = 0.6963 |
| *EVC2* | rs6820907 | G > A | G = 0.7813; A = 0.2187 | *FBN3* | rs35025963 | C > A, G, T | C = 0.7005; T = 0.2995 |
| | rs4689278 | T > A, C | T = 0.7348; C = 0.2652 | | rs33967815 | C > T | C = 0.8622; T = 0.1378 |
| | rs730469 | T > C | T = 0.4824; C = 0.5176 | | rs12975322 | C > A, G, T | C = 0.7173; T = 0.2827 |
| *GHR* | rs6180 | A > C, G | A = 0.5555; C = 0.4445 | | rs12608849 | G > A | G = 0.6905; A = 0.3095 |
| *BEST3* | rs61747221 | G > A | G = 0.8620; A = 0.1380 | | rs12150963 | G > A, C | G = 0.8678; C = 0.1322 |
| *EP300* | rs20551 | A > G | A = 0.7766; G = 0.2234 | | rs7257948 | C > A | C = 0.4205; A = 0.5795 |
| *ERLEC1* | rs2287345 | G > C | G = 0.8914; C = 0.1086 | | rs7246376 | G > A | G = 0.6727; A = 0.3273 |
| *GLI2* | rs72971975 | A > G | A = 0.8608; G = 0.1392 | | rs7245429 | G > C, T | G = 0.5142; T = 0.4858 |
| | rs12711538 | G > A | G = 0.4329; A = 0.5671 | | rs4804063 | T > C, G | T = 0.6923; C = 0.3077 |
| | rs3738880 | G > A, T | G = 0.4910; T = 0.5090 | | rs3829817 | C > G, T | C = 0.8163; T = 0.1837 |

[a] Global allelic frequency above 0.1 in 1000 genomes project.

**Table 2**
Analysis of 56 screened SNPs by seven web-based tools.

| Gene names | SNP ID[a] | CADD | PolyPhen-2 | PROVEAN | SNAP2 | PANTHER | FATHMM | PON-P2 |
|---|---|---|---|---|---|---|---|---|
| *KRT7* | rs6580870 | Neutral | Benign | Neutral | Effect | Benign | Tolerated | Neutral |
| | rs2608009 | Neutral | Benign | Neutral | Neutral | Benign | Tolerated | Neutral |
| *MYO1H* | rs11611277 | Neutral | Benign | Neutral | Neutral | Benign | Damaging | Neutral |
| | rs3825393 | Neutral | Benign | Neutral | Effect | Benign | Tolerated | Unknown |
| *PLXNA2* | **rs4844658** | **Deleterious** | **Damaging** | **Deleterious** | **Neutral** | **Damaging** | **Tolerated** | **Unknown** |
| | rs3748735 | Neutral | Benign | Neutral | Neutral | Damaging | Tolerated | Unknown |
| | rs2782948 | Neutral | Benign | Neutral | Effect | Damaging | Tolerated | Neutral |
| *SSX2IP* | rs1057746 | Neutral | Benign | Neutral | Effect | Benign | Tolerated | Neutral |
| *ADAMTS1* | rs428785 | Neutral | Benign | Neutral | Neutral | Damaging | Tolerated | Neutral |
| *ALPL* | rs3200254 | Neutral | Benign | Neutral | Neutral | Unknown | Damaging | Neutral |
| *ARHGAP21* | rs3748222 | Neutral | Benign | Neutral | Neutral | Unknown | Tolerated | Unknown |
| | rs1127893 | Neutral | Benign | Neutral | Neutral | Unknown | Tolerated | Unknown |
| *DUSP6* | **rs2279574** | **Deleterious** | **Damaging** | **Neutral** | **Effect** | **Damaging** | **Tolerated** | **Pathogenic** |
| | rs770087 | Deleterious | Benign | Neutral | Effect | Benign | Tolerated | Neutral |
| *EVC* | rs6414624 | Neutral | Benign | Neutral | Neutral | Damaging | Tolerated | Unknown |
| | rs2302075 | Neutral | Benign | Neutral | Neutral | Benign | Tolerated | Neutral |
| | rs2291157 | Neutral | Benign | Deleterious | Neutral | Benign | Tolerated | Unknown |
| | rs1383180 | Deleterious | Damaging | Neutral | Neutral | Damaging | Tolerated | Unknown |
| *EVC2* | rs6820907 | Neutral | Damaging | Neutral | Neutral | Benign | Tolerated | Neutral |
| | rs4689278 | Neutral | Benign | Neutral | Neutral | Benign | Tolerated | Unknown |
| | rs730469 | Neutral | Benign | Neutral | Neutral | Benign | Tolerated | Neutral |
| *GHR* | rs6180 | Deleterious | Damaging | Neutral | Neutral | Damaging | Tolerated | Neutral |
| *BEST3* | rs61747221 | Neutral | Benign | Neutral | Effect | Unknown | Damaging | Unknown |
| *EP300* | rs20551 | Neutral | Benign | Neutral | Neutral | Benign | Damaging | Unknown |
| *ERLEC1* | rs2287345 | Neutral | Benign | Neutral | Neutral | Benign | Tolerated | Neutral |
| *GLI2* | rs72971975 | Neutral | Unknown | Unknown | Unknown | Unknown | Unknown | Unknown |
| | rs12711538 | Neutral | Benign | Neutral | Effect | Benign | Tolerated | Neutral |
| | rs3738880 | Neutral | Benign | Neutral | Neutral | Benign | Tolerated | Unknown |
| *NBPF9* | rs200319336 | Neutral | Unknown | Neutral | Neutral | Benign | Unknown | Unknown |
| | rs71231749 | Neutral | Unknown | Neutral | Effect | Benign | Unknown | Unknown |
| | rs1227241524 | Neutral | Unknown | Neutral | Neutral | Unknown | Unknown | Unknown |
| *NCOR2* | rs2229840 | Deleterious | Damaging | Neutral | Neutral | Damaging | Tolerated | Unknown |
| *NOTCH3* | rs35769976 | Neutral | Benign | Neutral | Neutral | Benign | Damaging | Unknown |
| | rs1044009 | Neutral | Benign | Neutral | Neutral | Benign | Tolerated | Neutral |
| *NOTCH4* | rs915894 | Neutral | Benign | Neutral | Neutral | Benign | Damaging | Unknown |
| | rs520692 | Neutral | Benign | Neutral | Effect | Benign | Damaging | Unknown |
| | rs422951 | Neutral | Benign | Neutral | Neutral | Benign | Damaging | Unknown |
| *FGFR2* | rs755793 | Neutral | Benign | Neutral | Effect | Benign | Tolerated | Unknown |
| *FGF23* | rs7955866 | Neutral | Benign | Neutral | Effect | Unknown | Tolerated | Neutral |
| *MYH3* | rs2285477 | Neutral | Benign | Neutral | Neutral | Benign | Tolerated | Neutral |
| *MYH8* | rs8069834 | Neutral | Benign | Neutral | Effect | Unknown | Tolerated | Neutral |
| *RASA2* | rs34558081 | Neutral | Unknown | Unknown | Effect | Unknown | Tolerated | Pathogenic |
| *HSPG2* | rs3736360 | Neutral | Benign | Neutral | Neutral | Unknown | Tolerated | Neutral |
| | rs2291827 | Deleterious | Benign | Deleterious | Neutral | Damaging | Tolerated | Unknown |
| | rs897471 | Deleterious | Damaging | Deleterious | Neutral | Benign | Tolerated | Unknown |
| *IGF1* | rs35767 | Neutral | Unknown | Unknown | Effect | Unknown | Unknown | Unknown |
| *FBN3* | rs35025963 | Neutral | Benign | Neutral | Effect | Unknown | Damaging | Neutral |
| | **rs33967815** | **Deleterious** | **Damaging** | **Deleterious** | **Effect** | **Unknown** | **Damaging** | **Neutral** |
| | rs12975322 | Neutral | Damaging | Neutral | Neutral | Unknown | Damaging | Neutral |
| | rs12608849 | Neutral | Benign | Neutral | Neutral | Unknown | Damaging | Neutral |
| | rs12150963 | Deleterious | Benign | Deleterious | Effect | Unknown | Damaging | Neutral |
| | rs7257948 | Neutral | Benign | Neutral | Neutral | Unknown | Tolerated | Neutral |
| | rs7246376 | Neutral | Benign | Neutral | Neutral | Unknown | Damaging | Neutral |
| | rs7245429 | Deleterious | Damaging | Deleterious | Neutral | Unknown | Damaging | Neutral |
| | rs4804063 | Neutral | Benign | Neutral | Neutral | Unknown | Damaging | Neutral |
| | rs3829817 | Neutral | Benign | Neutral | Effect | Unknown | Damaging | Neutral |

[a] Selected SNPs are **bolded**.

and *FBN3-rs33967815*, in I-Mutant2 and MUpro tools revealed that all of these SNPs can decrease the stability of their respective proteins (Table 3). Reliability Index and delta G values indicated that this effect on protein stability is more significant for *FBN3-rs33967815* in I-Mutant2 and for *PLXNA2-rs4844658* in MUpro.

### 3.4. Residue conservation analysis

The PLXNA2 (Accession: NP_079455.3), DUSP6 (Accession: NP_001937.2), and FBN3 (Accession: NP_115823.3) proteins have 1894, 381, and 2809 amino acids, respectively. The *PLXNA2-rs4844658* occurs at position 369, *DUSP6-rs2279574* at position 114, and *FBN3-rs33967815* at position 1614 of their respective proteins. Analysis of these positions in ConSurf revealed that two SNPs alter exposed residues with variable conservation (Table 3). While only amino acid 114 is buried in DUSP6 protein with an average level of conservation.

### 3.5. Analysis of structural effects

The HOPE tool revealed that selected harmful SNPs have the potential to impact the function of their related proteins. Specifically, *PLXNA2-rs4844658* (E369G) and *DUSP6-rs2279574* (V114L) can cause structural changes in the Sema and Rhodanese domains, respectively. These domains are crucial for protein activity and interact with other domains (Table 3). Additionally, *FBN3-rs33967815* may disturb the EGF-like 25 domain and abolish its function (Table 3) due to a Serine to Glycine replacement (G1614S). The names of mentioned domains were adapted from UniProt (https://www.uniprot.org/) website.

**Table 3**
Structural, stability, and conservation analysis of three harmful SNPs.

| SNP ID | I-Mutant2 | MUpro | ConSurf | HOPE |
|---|---|---|---|---|
| rs4844658 | Decrease (Reliability index: 2) | Decrease (ΔG: -1.26) | Conservation scale: 3 An exposed residue with variable conservation | (Glutamic-Acid into a Glycine) The mutated residue is located in a domain that is important for the activity of the protein and in contact with another domain that is also important for the activity. The interaction between these domains could be disturbed by the mutation, which might affect the function of the protein. |
| rs2279574 | Decrease (Reliability index: 7) | Decrease (ΔG: -0.17) | Conservation scale: 6 A buried residue with average conservation | (Valine into a Leucine) The mutated residue is located in a domain that is important for the activity of the protein and in contact with another domain that is also important for the activity. The interaction between these domains could be disturbed by the mutation, which might affect the function of the protein. |
| rs33967815 | Decrease (Reliability index: 8) | Decrease (ΔG: -0.87) | Conservation scale: 1 An exposed residue with variable conservation | (Glycine into a Serine) The mutation introduces an amino acid with different properties, which can disturb a specific domain and abolish its function. |

Furthermore, based on LOMETS models, all three harmful SNPs can change the 3D structure of their proteins (Fig. 2).

## 4. Discussion

Mandibular prognathism is a common facial disorder that is characterized by excessive growth of the lower jaw with or without hypoplasia of the upper jaw. The incidence of this disorder varies among ethnic groups, with the highest rate in the Asian population.[5] Although the etiology of mandibular prognathism is still unclear, several studies have shown that it is likely to be polygenic and multifactorial. Association studies, genome-wide association studies, and genome-wide linkage analysis have led to the discovery of numerous genes correlated with mandibular prognathism. In particular, single nucleotide polymorphisms have played an essential role in advancing the mentioned research, and their significance cannot be overstated.[7,14] The current study utilized several web-based tools to screen and analyze missense SNPs (global frequency above 0.1) in 56 genes associated with mandibular prognathism (Fig. 1). Finally, according to the results of Table 2, *PLXNA2-rs4844658*, *DUSP6-rs2279574*, and *FBN3-rs33967815* were selected for further investigation. Through in-depth analysis of these three SNPs, it was determined that they may have significant effects on the structure and proper function of their respective proteins.

*PLXNA2* is a member of the semaphorin receptor family located on chromosome band 1q32.2. This gene is expressed in various tissues during craniofacial development, such as the neural crest and developing cranial bones. The cells of the neural crest are derived from the developing nervous system and give rise the bones and cartilage of the face and skull.[28] PLXNA2 is a transmembrane protein that is thought to transduce signals from semaphorin-3A.[29] Both PLXNA2 and semaphorin-3A are expressed in osteoblast cells, and their binding to each other can stimulate osteoblast differentiation.[30] The interaction of two proteins may suppress the expression of parathyroid hormone-related peptide receptor 1 (PTH-R1) in human proliferative chondrocytes.[31] This is significant because the parathyroid hormone-related protein and its receptor PTH-R1 play a crucial role in regulating chondrocyte proliferation, differentiation, and apoptosis.[32] As most chondrocytes in the mandibular condylar cartilage undergo programmed death, the cartilage matrix is replaced by bone through endochondral ossification, ultimately forming the mandibular condyle.[33] Researchers conducted a genome-wide association study in the Japanese population and suggested 1q32.2 loci as a susceptibility region for the development of mandibular prognathism.[34] However, no experimental studies have investigated the association between *PLXNA2* SNPs and mandibular prognathism to date.

In this study, we identified *PLXNA2-rs4844658* as a harmful polymorphism (Table 2). This SNP occurs at position E369G with variable conservation, which decreases PLXNA2 stability and affects its functions (Table 3). Due to the functional and structural changes caused by *PLXNA2-rs4844658* (Fig. 2 A1-A2), the interaction between this protein and semaphorin-3A may be disrupted, leading to a potential reduction in the expression of PTH-R1. In mice, the functions of PTH-R1 cause a delay in the differentiation process of chondrocytes.[35] Therefore, decreasing PTH-R1 expression may promote the differentiation of chondrocytes and accelerate endochondral ossification, which can ultimately result in mandibular prognathism. To advance our understanding of the molecular mechanisms underlying mandibular prognathism, we suggest investigating the potential impact of *PLXNA2* SNPs, particularly *PLXNA2-rs4844658*, in experimental studies. Future research may reveal additional insights into the role of this gene in the development of mandibular prognathism.

*DUSP6* is a member of the dual specificity phosphatase subfamily located on chromosome band 12q21.33. This gene encodes a cytoplasmic enzyme that regulates the activity of mitogen-activated protein kinases (MAPKs) superfamily (including MAPK/ERK, SAPK/JNK, p38).[36] The absence or reduction of DUSP6 in mice can lead to the development of at least two distinct phenotypes, skeletal dwarfism and coronal craniosynostosis (a medical condition that may affect the development of the skull). These phenotypes are believed to be caused by improper regulation of MAPK/ERK signaling pathway.[37] The MAPK/ERK pathway is crucial for the formation of skeletal structures, as it controls the differentiation, proliferation, and activity of osteoblasts, chondrocytes, and osteoclasts.[38] During early mouse embryo development, the activation of the MAPK/ERK pathway relies heavily on fibroblast growth factors (FGFs). These FGFs activate FGF receptors (FGFRs), which are single-pass transmembrane proteins with intracellular tyrosine kinase activity.[39] MAPK/ERK pathway is stimulated by FGF/FGFR, which regulates the differentiation of chondrocytes. Since chondrogenesis is essential for endochondral ossification, FGF/FGFR signaling plays a crucial role in the development of facial bones.[40]

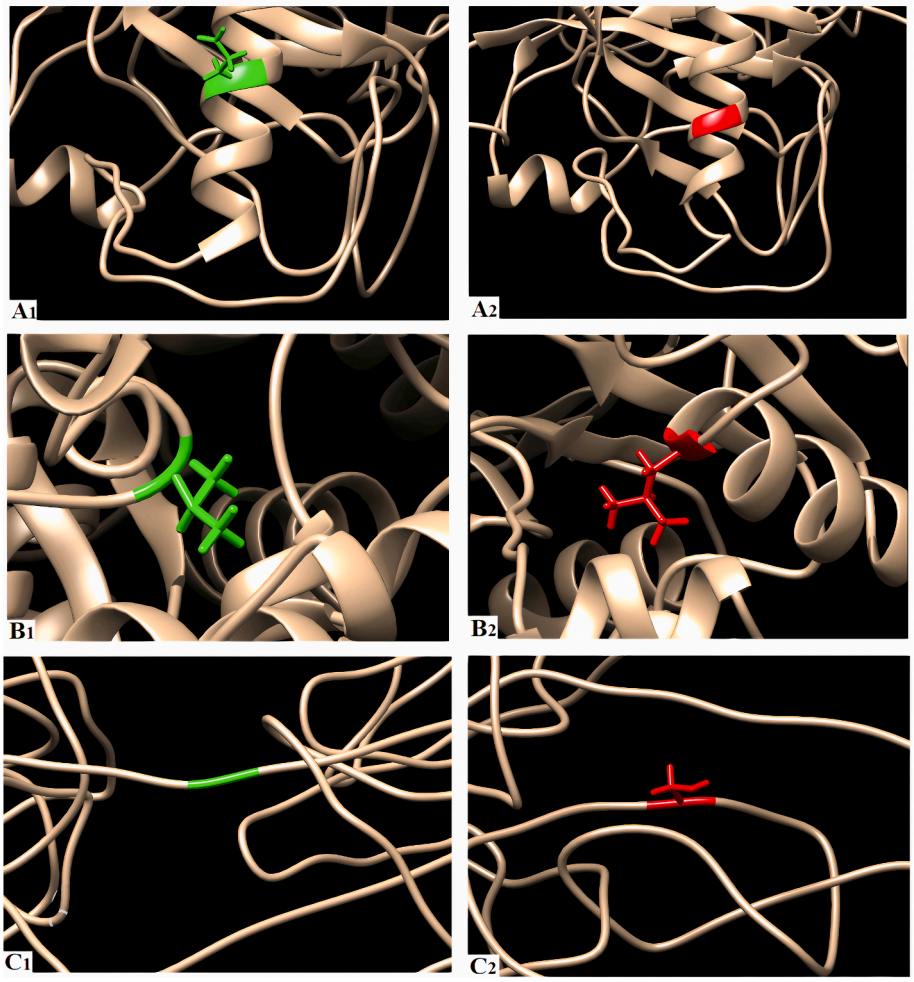


**Fig. 2.** Effect of harmful SNPs on the 3D structure of their corresponding proteins (the wild-type and the mutant residues are shown and colored green and red respectively). **A1-A2.** *PLXNA2-rs4844658* causes Glycine substitution for Glutamic-acid at position 369 (E369G). **B1-B2.** *DUSP6-rs2279574* leads to conversion of Valine into a Leucine at position 114 (V114L). **C1-C2.** *FBN3-rs33967815* results in the substitution of a Serine for a Glycine at position 1614 (G1614S).

Through whole-exome sequencing on five siblings from an Estonian family, researchers reported *DUSP6-rs139318648* (Serine182 Phenylalanine) as a significant variant in skeletal Class III malocclusion. [3]They suggested that *DUSP6* may be responsible for both mandibular prognathism and maxillary deficiency. The results of another study revealed significant dissimilarities in craniofacial morphology between individuals with the *DUSP6* mutation and non-mutation individuals, indicating a possible role of this gene in the development of Class III malocclusion.[41] In the present study, we screened *DUSP6-rs2279574* as a harmful polymorphism (Table 2). This SNP decreases protein stability and occurs at a critical position (V114L) that results in significant structural changes (Table 3). Specifically, the structure of this position, which normally features a buried residue with average conservation, shifts from a strand to a helix when the SNP is present (Fig. 2 B1-B2). Based on the significant structural changes mentioned, the phosphatase activity of DUSP6 may increase in the presence of ERKs. Moreover, there is a report indicating that Erk1/2 is able to inhibit chondrocyte differentiation.[42] With increased DUSP6 activity, Erk1/2 may experience a greater loss of function, which could accelerate chondrocyte differentiation and subsequent endochondral ossification, potentially leading to mandibular prognathism. Finally, we believe that investigating DUSP6-rs2279574 through experimental studies has the potential to contribute significantly to our understanding of mandibular prognathism and improve clinical outcomes for affected individuals.

*FBN3* encodes a member of the fibrillin protein family and is located on chromosome band 19p13.2. This fibrillin is expressed in the interstitial matrix of hypertrophic cartilage and is present throughout endochondral ossification of the ribs.[43] Recent research conducted on a cohort of Chinese families suggested that FBN3 variants may be implicated in the pathogenesis of Klippel-Trenaunay-Weber syndrome, a condition that can lead to skeletal overgrowth and asymmetry in the limbs.[44] In addition, genome-wide linkage analysis on Korean and Japanese sibling pairs has identified 19p13.2 as a susceptibility region for mandibular prognathism.[45] Despite these studies, the precise function of *FBN3* and the molecular pathways that may cause the aforementioned abnormalities, particularly mandibular prognathism, is still unknown. Here, we screened *FBN3-rs33967815* as a harmful polymorphism (Table 2). Our further analysis revealed that this SNP occurs in a position with variable conservation and reduces the stability of FBN3 (Table 3). Moreover, the SNP causes significant changes in the protein structure that could have an impact on its overall function (Fig. 2C1-C2). Based on these findings, our study suggests that it is worthwhile to investigate the potential association between *FBN3-rs33967815* and mandibular prognathism. Future studies could serve as an initial step in discovering the exact role of FBN3 in the pathways involved in mandibular formation.

Although we used various web-based tools to identify potential SNPs associated with mandibular prognathism in our study, it is important to note the limitations of this research method. Despite the convenience and accessibility of these tools, their algorithms may be based on restricted data, resulting in an incomplete understanding of the effects of SNPs on biological systems. Additionally, while changes in protein

stability and structure are important factors in disease development, epigenetic modifications or interactions with other genetic or environmental factors may also contribute to the occurrence of abnormalities. Furthermore, the effects of a SNP on a protein might be more complex than what computational tools can predict. Therefore, combining web-based tools with experimental validation can provide a more complete understanding of the genetic basis of a disease such as mandibular prognathism. By doing so, researchers may be able to identify new targets for treatment or prevention of this condition.

## 5. Conclusion

After analyzing missense SNPs with global frequency above 0.1 in 56 genes linked to mandibular prognathism, our study screened *PLXNA2-rs4844658*, *DUSP6-rs2279574*, and *FBN3-rs33967815* as harmful polymorphisms. Through further analysis, we found that all three SNPs can impact the stability, structure, and function of their respective proteins, and they occur at positions with variable or average conservation. Given the possible involvement of PLXNA2, DUSP6, and FBN3 proteins in ossification pathways, our study suggests that identified harmful SNPs may be potential causes of mandibular prognathism. These SNPs could be reported as a genetic molecular marker for mandibular-related disorders in future studies.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Declaration of competing interest

We wish to confirm no known conflicts of interest associated with this publication. There has been no significant financial support for this work that could have influenced its outcome.

## Please note

The following reference contains identifying information. So, we move it to the title page according to journal guidelines ("The main body of the paper should not include any identifying information, such as the author's names or affiliations").

27. Kalmari A, Heydari M, Hosseinzadeh Colagar A, Arash V. In silico analysis of collagens missense SNPs and human abnormalities. *Biochem Genet*. 2022:1-27. https://doi.org/10.1007/s10528-021-10172-6.

## References

1 Jacobson A, Evans WG, Preston CB, Sadowsky PL. Mandibular prognathism. *Am J Orthod*. 1974;66(2):140–171.
2 Jaruga A, Ksiazkiewicz J, Kuzniarz K, Tylzanowski P. Orofacial cleft and mandibular prognathism-human genetics and animal models. *Int J Mol Sci*. 2022;23(2):953.
3 Nikopensius T, Saag M, Jagomägi T, et al. A missense mutation in DUSP6 is associated with Class III malocclusion. *J Dent Res*. 2013;92(10):893–898.
4 Guan X, Song Y, Ott J, et al. The ADAMTS1 gene is associated with familial mandibular prognathism. *J Dent Res*. 2015;94(9):1196–1201.
5 Tassopoulou-Fishell M, Deeley K, Harvey EM, Sciote J, Vieira AR. Genetic variation in myosin 1H contributes to mandibular prognathism. *Am J Orthod Dentofacial Orthop*. 2012;141(1):51–59.
6 Jena AK, Duggal R, Mathur VP, Parkash H. Class-III malocclusion: genetics or environment? A twins study. *J Indian Soc Pedod Prev Dent*. 2005;23(1):27–30.
7 Doraczynska-Kowalik A, Nelke KH, Pawlak W, Sasiadek MM, Gerber H. Genetic factors involved in mandibular prognathism. *J Craniofac Surg*. 2017;28(5):e422–e431.
8 Parada C, Chai Y. Mandible and tongue development. *Curr Top Dev Biol*. 2015;115:31–58.
9 Mackie E, Ahmed YA, Tatarczuch L, Chen KS, Mirams M. Endochondral ossification: how cartilage is converted into bone in the developing skeleton. *Int J Biochem Cell Biol*. 2008;40(1):46–62.
10 Mizoguchi I, Toriya N, Nakao Y. Growth of the mandible and biological characteristics of the mandibular condylar cartilage. *Jpn Dent Sci Rev*. 2013;49(4):139–150.
11 Neela PK, Atteeri A, Mamillapalli PK, et al. Genetics of dentofacial and orthodontic abnormalities. *Glob Med Genet*. 2020;7(4):95–100.
12 Wang DG, Fan JB, Siao CJ, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science*. 1998;280(5366):1077–1082.
13 Sukhumsirichart W. Polymorphisms. In: *Genetic Diversity and Disease Susceptibility*. IntechOpen; 2018. https://doi.org/10.5772/intechopen.76728.
14 Gershater E, Li C, Ha P, et al. Genes and pathways associated with skeletal sagittal malocclusions: a systematic review. *Int J Mol Sci*. 2021;22(23), 13037.
15 Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47(D1):886–894.
16 Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013;76(1):7–20.
17 Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*. 2015;31(16):2745–2747.
18 Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants. *BMC Genom*. 2015;16(8):1–12.
19 Tang H, Thomas PD. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics*. 2016;32(14):2230–2232.
20 Shihab HA, Gough J, Cooper DN, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Hum Mutat*. 2013;34(1):57–65.
21 Niroula A, Urolagin S, Vihinen M. PON-P2: prediction method for fast and reliable identification of harmful variants. *PLoS One*. 2015;10(2), 0117380.
22 Capriotti E, Fariselli P, Casadio R, I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res*. 2005;33(suppl_2):306–310.
23 Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins: Struct, Funct, Bioinf*. 2006;62(4):1125–1132.
24 Ashkenazy H, Erez E, Martz E, Pupko T, ConSurf Ben-Tal N. 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res*. 2010;38(suppl_2):529–533.
25 Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinf*. 2010;11(1):1–10.
26 Wu S, Zhang Y. LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic Acids Res*. 2007;35(10):3375–3382.
27 This Reference Has Been Removed According to Journal Guidelines (The Manuscript Should Not Contain Any Identifying Information).
28 Bronner ME, LeDouarin NM. Development and evolution of the neural crest: an overview. *Dev Biol*. 2012;366(1):2–9.
29 Sabag AD, Smolkin T, Mumblat Y, et al. The role of the plexin-A2 receptor in Sema3A and Sema3B signal transduction. *J Cell Sci*. 2014;127(24):5240–5252.
30 Hayashi M, Nakashima T, Taniguchi M, Kodama T, Kumanogoh A, Takayanagi H. Osteoprotection by semaphorin 3A. *Nature*. 2012;485(7396):69–74.
31 Kajii TS, Oka A, Hatta M, Yamazaki J, Yamashita J, Iida J. PLXNA2 identified as a candidate gene by genome-wide association analysis for mandibular prognathism in human chondrocytes. *Biomed Rep*. 2018;9(3):253–258.
32 Clemens TL, Cormier S, Eichinger A, et al. Parathyroid hormone-related protein and its receptors: nuclear functions and roles in the renal and cardiovascular systems, the placental trophoblasts and the pancreatic islets. *Br J Pharmacol*. 2001;134(6):1113–1136.
33 Hinton RJ, Jing Y, Jing J, Feng JQ. Roles of chondrocytes in endochondral bone formation and fracture repair. *J Dent Res*. 2017;96(1):23–30.
34 Saito F, Kajii TS, Oka A, Ikuno K, Iida J. Genome-wide association study for mandibular prognathism using microsatellite and pooled DNA method. *Am J Orthod Dentofacial Orthop*. 2017;152(3):382–388.
35 Guo J, Chung UI, Kondo H, Bringhurst FR, Kronenberg HM. The PTH/PTHrP receptor can delay chondrocyte hypertrophy in vivo without activating phospholipase C. *Dev Cell*. 2002;3(2):183–194.
36 Ahmad MK, Abdollah NA, Shafie NH, Yusof NM, Razak SRA. Dual-specificity phosphatase 6 (DUSP6): a review of its molecular characteristics and clinical relevance in cancer. *Cancer Biol Med*. 2018;15(1):14.
37 Li C, Scott DA, Hatch E, Tian X, Mansour SL. Dusp6 (Mkp3) is a negative feedback regulator of FGF-stimulated ERK signaling during mouse development. *Development*. 2007;134(1):167–176.
38 Lu N, Malemud CJ. Extracellular signal-regulated kinase: a regulator of cell growth, inflammation, chondrocyte and bone cell receptor-mediated gene expression. *Int J Mol Sci*. 2019;20(15):3792.
39 Sarabipour S, Hristova K. Mechanism of FGF receptor dimerization and activation. *Nat Commun*. 2016;7(1):1–12.
40 Su N, Jin M, Chen L. Role of FGF/FGFR signaling in skeletal development and homeostasis: learning from mouse models. *Bone Res*. 2014;2(1):1–24.
41 Nowrin SA, Basri R, Alam MK, et al. Craniofacial morphology of Class III malocclusion with DUSP6 gene: mutation and non-mutation groups. *J Hard Tissue Biol*. 2016;25(3):247–256.
42 Ba P, Duan X, Fu G, Lv S, Yang P, Sun Q. Differential effects of p38 and Erk1/2 on the chondrogenic and osteogenic differentiation of dental pulp stem cells. *Mol Med Rep*. 2017;16(1):63–68.

43 Sabatier L, Miosge N, Hubmacher D, Lin G, Davis EC, Reinhardt DP. Fibrillin-3 expression in human development. *Matrix Biol*. 2011;30(1):43–52.
44 Liu HY, Zhou L, Zheng MY, et al. Diagnostic and clinical utility of whole genome sequencing in a cohort of undiagnosed Chinese families with rare diseases. *Sci Rep*. 2019;9(1):1–11.
45 Yamaguchi T, Park SB, Narita A, Maki K, Inoue I. Genome-wide linkage analysis of mandibular prognathism in Korean and Japanese patients. *J Dent Res*. 2005;84(3):255–259.