

Genetic Evaluation of the Class III Dentofacial in Rural and Urban Spanish Population by AI Techniques

Marta Muñoz, Manuel Rodríguez, M^a Encarnación Rodríguez,
and Sara Rodríguez

Abstract. The etiology of skeletal class III malocclusion is multifactorial, complex and likely results from mutations in numerous genes. In this study, we sought to understand genotype correlation of the class III dentofacial deformity in rural and urban spanish population of more than one generation. The genetic analyze was made using a Genome-wide scan. It will hold a novel classification using Artificial Intelligence techniques highlighting the difference between the two groups at the level of polymorphism. Our phenotypic and genetic analysis highlights that each group is unique.

Keywords: Skeletal class III malocclusion, genetic etiology, genome-wide scan, extraction of knowledge, artificial intelligence.

1 Introduction

Skeletal Class III malocclusion is a general morphological description of a diverse group of dentofacial conditions in which the mandibular teeth are forward in relationship to the maxillary teeth, resulting in an anterior crossbite or underbite. The term skeletal implies that the positions of the teeth are the result of underlying jaw relationships. This type of skeletal occlusal pattern is also referred to as true Class III or true mesiocclusion. These conditions are developmental to the extent that they are not recognizable at birth and by definition, until the individual is dentate, it is not possible to make a diagnosis of skeletal Class III malocclusion.

Marta Muñoz · Manuel Rodríguez · Sara Rodríguez
University of Salamanca, Salamanca, Spain
e-mail: {mmbruguier, mrrvm, srg}@usal.es

M^a Encarnación Rodríguez
Dentistry and Stomatology, Hnos. García Noblejas, Madrid
e-mail: maritarodrod@hotmail.com

As one would expect there is a higher incidence of this condition in the transitional and adult dentitions than there is in the primary dentition.

A Class III deformity can be an attribute of a syndrome, as in achondroplasia with associated midface deficiency resulting from a failure in the development of the cartilaginous nasal capsule or can merely be a manifestation of normal morphologic variation. Where a growth effect is responsible for the skeletal Class III problem, the affect can be primary and active, such as in acromegaly where an increased production of pituitary growth hormone acts on the condylar cartilage creating exuberant mandibular growth. Recent gene mapping and linkage analysis of individuals with achondroplasia and acromegaly have identified some of the responsible genes.

The Class III dentofacial deformity is clinically and genetically heterogeneous presenting with a distinct subphenotype and genotype in 2 cohorts. The main goal of this study is to conduct genome-wide scans followed by linkage analysis to identify the genetic loci associated with the Class III trait in rural and urban Spanish populations. The populations of individuals chosen to carry out the study are two groups of individuals of second generation rural and urban environments in Spain. It will hold a novel classification using Artificial Intelligence techniques highlighting the difference between Class III and normal individuals at the level of polymorphism.

The article is structured as follows: Section 2 makes a review of the problem description and the reasons that led to the realization of this research. Sections 3 introduces the most important requirements and functionalities of the AI (Artificial Techniques) used. Finally, some experimental results and conclusions are given in Sections 4 and 5.

2 Background

The genetic etiology of Class III malocclusion has been demonstrated in several studies. In [1], Bui et al. demonstrated that the Class III trait was inherited in an autosomal dominant fashion in the 12 families that they studied. This has been previously suggested by other studies [7][8]. Certain syndromic conditions with a genetic etiology, such as Crouzon syndrome, acromegaly and achondroplasia, have been described as presenting with skeletal Class III malocclusion [10][11][18].

In recent studies is reported results from the mutation on specific chromosomes and mutations in genes [16] [13], results in achondroplasia and amelogenesis. These recent advances have fallen on the heels of the Human Genome Project (HGP) that began in 1990. As a result of the HGP comprehensive genetic maps have been created that locate and identify genes underlying susceptibility to disease. Increasingly detailed knowledge of the human genome at the DNA level forms the basis of our understanding of genetic transmission and gene action. The HGP has mapped 30,000 genes thus far, and therefore provides the basis for genetic diagnosis and therapy.

The field of molecular genetics continues to improve and advance, it should be possible to identify relevant genetic markers for such traits. Skeletal Class III malocclusion or mandibular prognathism has been analyzed genetically [4][15][9].

Human studies have played a major role in the developing hypothesis that Class III malocclusion is at least in part due to genetic factors. Orofacial structures are significant in the development of the craniofacial complex and have been shown to be under genetic control, hence they should be considered in the etiology of the development of skeletal Class III malocclusion [7][8][3][14][12].

The existence of familial aggregation of mandibular prognathism (MP) suggests that genetic components play an important role in its etiology. A genetic etiology of class III malocclusion is suggested by many lines of evidence [5][2][6]. In Jena's report, a pair of monozygotic female twins were presented. The girls exhibited a marked similarity in facial appearance. They both had a similar dentition, but their occlusions were dissimilar to some extent. Twin 1, reverse overjet, overbite and class III molar relations were more severe than twin 2. Both twins had bilateral posterior crossbite. The cephalometric parameters did not reveal a very significant difference in skeletal morphology. Height of the anterior face was similar in both the twins, but posterior facial height was more in twin 2. Position of mandible in relation to anterior cranial base and Frankfort-horizontal plane was significantly different among the twins [5]. In this study, the concavity of the face (Angle of convexity) in twin 1 was more compared to twin 2.

Relatively more backward position of the maxilla (Angle SNA, N Perpendicular to point-A) and forward position of the chin (Angle SNB, N Perpendicular to Pog) contributed to such difference in the severity of the facial concavity. The antero-posterior position of the mandible in the twin study was influenced significantly by environmental factors. However, in a previous study undertaken, a report was made that the anterior-facial posterior position of the mandible is genetically determined. Anterior facial height of both twins was apparently equal. It showed that the height of the anterior face is genetically determined and did not play any role in the discordance of class III malocclusion. This is in agreement with the result from a study done by Townsend and Richards [17]. Another study investigated the role of genetic influences in the etiology of class III malocclusion [2]. In this study, a segregation analysis of 37 families of patients that were treated for mandibular prognathism, was performed. A study conducted by Yamaguchi, et al, in 2005 [19], utilized a genome-wide linkage analysis to identify loci susceptible to MP with 90 affected sibling-pairs in 42 families, comprised of 40 Korean sibling-pairs and 50 Japanese sibling-pairs. Two non-parametric linkage analyses, GENEHUNTER-PLUS and SIBPAL, were applied and detected nominal statistical significance of linkage to MP at chromosomes 1p36, 6q25, and 19p13.2. The best evidence of linkage was detected near D1S234 (maximum Zlr = 2.51, P = 0.0012). In addition, evidence of linkage was observed near D6S305 (maximum Zlr = 2.23, P = 0.025) and D19S884 (maximum Zlr = 1.93, P = 0.0089). This study while helpful relied on sibling pairs, which is less powerful than the family studies that we report in this publication.

The identification of the susceptible genes in the linkage regions will pave the way for insights into the molecular pathways that cause MP, especially overgrowth of the mandible, and may lead to the development of novel therapeutic tools.

3 AI Techniques

Conduct genome-wide scans followed by linkage analysis to identify the genetic loci associated with the Class III trait in the rural and urban populations. The GeneChip Mapping 10K 2.0 AssayR version by Affymetrix (Santa Clara, CA 2004), is a mapping tool designed to identify regions of the genome that are linked to or associated with a particular trait or phenotype. It is also useful for determination of allele frequencies in various populations and for mapping regions with chromosomal copy. This array system provides genotypes for 10,000 human single nucleotide polymorphisms (SNPs) on a single array.

The extraction of knowledge that is presented to the human expert is carried out using the J48 algorithm [21]. The J48 algorithm is the Java implementation of the C4.5 algorithm, an evolution of the original ID3 [20], whose main advantage is that it allows incorporate numerical attributes into the logical operations carried out in the test nodes. There are other alternatives for the generation of decision rules which operate similar to the decision trees, including RIPPER [22] and PART [23]. The J48 [20] algorithm attempts to minimize the width of the decision tree by using heavy search strategies. In summary, the algorithm defines two terms: gain and rate of gain with respect to the information $I(S)$ contained in a node S . Using only the gain criteria, attributes with multiple values are more highly favored given that they can more easily divide the elements into numerous subsets. To avoid the effect of favoring attributes with multiple values, the concept of gain rate is added.

$$I(S) = -\sum_{j=1}^n f_j^S \cdot \log(f_j^S) \quad (1)$$

Where f_j^S represents the relative frequency of class j in the node S .

The gain function is defined as follows:

$$G(S, B) = I(S) - \sum_{i=1}^l \frac{|S_i|}{|S|} I(S_i) \quad (2)$$

B represents the test that separates the modes and l the number of elements of the set.

To avoid favoring the partitions with branches containing few elements, the rate of gain term is introduced in the following manner:

$$P(S, B) = -\sum_{i=1}^l \frac{|S_i|}{|S|} \log\left(\frac{|S_i|}{|S|}\right) \quad (3)$$

Finally test B is selected from the previous tests since it maximizes the following criteria:

$$G(S, B) / P(S, B) \quad (4)$$

The system uses the equation 4 for calculating the nodes of the decision tree, the nodes represent the relevant SNPs that allow classifying the patients.

4 Experimental Results

The sample consists in 48 individuals in 4 families in rural environments and 25 individuals in 3 families urban environment were genotyped using 500 microsatellite markers prior to the refinement of the Single Nucleotide Polymorphism (SNP) genotyping methods. Some individuals belong to skeletal Class III malocclusion and other normal individuals. In order to amplify the DNA at these loci, polymerase chain reaction (PCR) was carried out using primers surrounding a previously identified locus. The initialization step took place at 95°C for 15 minutes and the DNA denaturation step took place at 95°C for 30 seconds. The annealing step took place at 62°C for 30 seconds. The elongation step took place at 72°C for 1 minute. Chromatography was then used to analyze the gene fragment. A chromatogram was analyzed to determine the size of the fragments (bp) as well as whether subjects were heterozygous or homozygous for a given allele. A chromatogram of a heterozygous individual typically exhibits two defined peaks with smaller leading and lagging peaks, while that of a homozygous individual typically exhibits one defined peak. The GeneChip Mapping 10K 2.0 AssayR version by Affymetrix (Santa Clara, CA 2004), is a mapping tool designed to identify regions of the genome that are linked to or associated with a particular trait or phenotype. It is also useful for determination of allele frequencies in various populations and for mapping regions with chromosomal copy.

An analysis by inspection revealed an autosomal dominant mode of inheritance among all the families. Rural family #1, is composed of 3 generations of both affected and unaffected individuals. Each generation exhibits an approximately equal number of affected and unaffected family members. In this family, more females than males exhibit the phenotype. This family appears to exhibit autosomal dominant mode of inheritance of the Class III trait. Rural family #2, comprises 4 generations of individuals affected and unaffected with the skeletal Class III trait. This family has twice as many females exhibit the phenotype. This family also appears to exhibit an autosomal dominant mode of inheritance. Rural family #3, is composed of 4 generations with approximately males and females affected the trait equally, exhibiting an autosomal dominant mode of inheritance. Rural family #4, comprises 2 generations with twice as many female affected as males and all are affected with the trait. The mode of inheritance is autosomal dominant. AI family #19 and #33 are both composed of 6 and 7 generations respectively, with an equal number of males and females affected with the skeletal Class III trait, thereby exhibiting an autosomal dominant mode of inheritance.

Urban family #1, is made up of three generations of both affected and unaffected individuals. Each generation exhibits an approximately equal number of affected and unaffected family members. This family has twice as many males as females exhibiting the phenotype and appears to exhibit an autosomal dominant mode of skeletal Class III malocclusion. Urban Family #2 in the urban cohort, is composed of 2 generations with an equal amount of males affected with the skeletal Class III trait as females. This family appears to exhibit autosomal dominant inheritance of the skeletal Class III malocclusion. Urban family # 3, has 5 generations with 4 generations affected. Males seem twice as likely to inherit the condition as females. This family appears to exhibit autosomal dominant inheritance of this trait. Urban Family #4, has 2 generations with both generations affected with skeletal Class III malocclusion. Twice as many males are likely to be affected than females. This family appears to exhibit an autosomal dominant mode of inheritance of the skeletal Class III phenotype. Urban Family # 5, has 3 generations with 1 generation of affected individuals. More males were affected than females. There was some uncertainty regarding the skeletal Class III affection status of some of the individuals, hence a statement regarding the mode of inheritance was not made for this family.

Urban family #6, had 3 generations with 2 generations of affected individuals and an equal number of affected females as males. This family also had a few individuals where the affection status of the skeletal Class III trait was uncertain, however, with the information currently available, this family appears to have an autosomal dominant mode of inheritance. Urban Caucasian family #7 had 2 generations with only females affected with the trait. Urban Family #8 in the, had 2 generations of affected individuals with twice as many females affected as males, revealing an autosomal dominant mode of inheritance.

The relevant regions obtained using J48 are between markers D1S2865 to D1S435, D1S435 to D1S206, D3S3725 to D3S3041, and D12S368 to D12S83 yielded on chromosomes 1, 3 and 12. Furthermore, single nucleotide polymorphism chip technology (SNP) has been used to identify regions associated with skeletal Class III malocclusion in the urban families.

The Affymetrix GeneChip Mapping Assay (Affymetrix, Inc., Santa Clara, CA) has been used to analyze SNP's in both affected and unaffected individuals. We anticipate the results from the urban linkage analysis will reveal other genes in addition to those found from the results of the rural analysis.

Table 1 Linkage Analysis

Markers	Chromosome
D1S2865 to D1S435	1
D1S435 to D1S206	1
D3S3725 to D3S3041	3
D12S368 to D12S83	12

Although the findings of this study are suggestive of linkage at chromosome 1 for the skeletal Class III trait, the answers regarding the genetic etiology of true mandibular prognathism might be found in the IGF system which would suggest that the mutation could be on chromosome 15. If the somatomedin hypothesis introduced by Daughaday almost a half-century ago is correct, the regulation of the growth of the skeleton would be the key to the entire genetic study on skeletal Class III malocclusion.



Fig. 1 Diagrammatic Representation of 23 Chromosomes and Relative Location of Markers D1S2865 – D1S435 using Parametric Linkage Analysis for rural cohort

5 Conclusions

A visual inspection of the pedigrees suggests an autosomal dominant mode of inheritance of skeletal Class III malocclusion. Results from the linkage analysis in this study suggest that chromosomes 1, 3 and 12 are suggestive of linkage to the skeletal Class III trait. In light of other genetic studies currently being done with improved technology, these results are not consistent with some of the other previous studies which would suggest that the IGF-1 gene located on chromosome 15, is involved in the regulation of growth hormone and hence the development of the skeleton.

Once the skeletal Class III trait is phenotypically characterized according to type, i.e., maxillary hypoplasia, mandibular prognathism, or a combination of both, it may be possible to utilize the candidate genes identified for other syndromes that have a skeletal Class III component in the identification of the genes involved in the development of this trait.

Acknowledgments. This work has been supported by the Spanish JCyL project SA225A11-2.

References

- [1] Bui, C., King, T., Proffit, W., Frazier-Bowers, S.: Phenotypic Characterization of Class III Patients, A Necessary Background for Genetic Analysis. *Angle Orthodontist*, 564–569 (2006)
- [2] El-Gheriani, A.A., et al.: Segregation Analysis of Mandibular Prognathism in Libya. *J. Dent. Res.* 82(7), 523–527 (2003)
- [3] Horowitz, et al.: A cephalometric study of craniofacial variation in adult twins. *Angle Orthodontist* 30, 1–5 (1960)
- [4] Huang, C.M., et al.: Mandibular prognathism in the rabbit: discrimination between singlelocus and multifactorial models of inheritance. *The Journal of Heredity* 72(4), 296–298 (1981)
- [5] Jena, A., et al.: Class III malocclusion: Genetics or environment? A twins study. *J. of Indian Society of Pedodontics and Preventive Dentistry* 23(1), 27–30 (2005)
- [6] Litton, S.F., Ackeramn, L.V., Isaacson, R.J., Shapiro, B.L.: A genetic study of Class III malocclusion. *Am. J. Orthod.* 58, 565–577 (1970)
- [7] Mossey, P.A., et al.: The Heritability of Malocclusion: Part 1. Genetics, Principles and Terminology. *British Journal of Orthodontics* 26(3), 103–113 (1999)
- [8] Mossey, P.A., et al.: The Heritability of Malocclusion: Part 2. The Influence of Genetics in Malocclusion. *British Journal of Orthodontics* 26(2), 195–203 (1999)
- [9] Pereira, A., et al.: Association of GH and IGF-1 polymorphisms with growth traits in a synthetic beef cattle breed. *Genetics and Molecular Biology* 28(2) (2005)
- [10] Preising, M., et al.: On the effect of mutations of the fibroblast growth factor receptors as exemplified by three cases of craniosynostoses. *Klin. Monatsbl Augenheilkd* 220(10), 669–681 (2003)
- [11] Machicek, S.L., et al.: Mandible Size and Prognathism of MEK1 Transgenic Achondroplastic Mice. Presentation at AADR Meeting (March 2007)
- [12] Nakasima, A., Ichinose, M., et al.: Hereditary factors in the craniofacial morphology of Angle's Class II and Class III malocclusions. *Am. J. Orthod.* 82(2), 150–156 (1982)
- [13] Ravassipour, D.B., Powell, C.M., et al.: Variation in dental and skeletal open bite malocclusion in humans with amelogenesis imperfecta. *Archives of Oral Biology* 50, 611–623 (2005)
- [14] Singh, G.D., et al.: Craniofacial heterogeneity of prepubertal Korean and European-American subjects with Class III malocclusions: Procrustes, EDMA and cephalometric analyses. *Int. J. Adult Orthod. Orthog. Surg.* 13, 227–240 (1998a)
- [15] Sutter, N., et al.: A single IGF1 allele is a major determinant of small size in dogs. *Science* 316(5821), 112–115 (2007)
- [16] Tamura, Y.: A kindred of familial acromegaly without evidence for linkage to MEN-1 locus. *Endocr J.* 49(4), 425–431 (2002)
- [17] Townsend, G.C., Richards, L.C.: Twin and twinning, dentists and dentistry. *Aust. Dent. J.* 35, 317–327 (1990)
- [18] Yagi, T., et al.: Surgical Orthodontic Correction of Acromegaly with Mandibular Prognathism. *Angle Orthodontist* 74(1), 125–131 (2004)

- [19] Yamaguchi, T., et al.: Genome-wide Linkage Analysis of Mandibular Prognathism in Korean and Japanese Patients. *J. Dent. Res.* 84(3), 255–259 (2005)
- [20] Pinzón, C.I., De Paz, J.F., Navarro, M., Bajo, J., Julián, V., Corchado, J.M.: Real-time CBR-agent with a mixture of experts in the reuse stage to classify and detect DoS attacks. *Applied Soft. Computing* 11(7), 4384–4398 (2011)
- [21] Saravanan, N., Cholairajana, S., Ramachandran, K.I.: Vibration-based fault diagnosis of spur bevel gear box using fuzzy technique. *Expert Systems with Applications* 36(2-2), 3119–3135 (2009)
- [22] Holte, R.C.: Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 63–91 (1993)
- [23] Frank, E., Witten, I.H.: Generating accurate rule sets without global optimization, pp. 144–151. Morgan Kaufmann, San Mateo (1998)
- [24] Vogiatzis, D., Tsapatsoulis, N.: Active learning for microarray data. *International Journal of Approximate Reasoning* 47(1), 85–96 (2008)