GENETIC ASPECTS OF IDIOPATHIC ESCOLIOSIS - A LITERATURE REVIEW

ABSTRACT

The idiopathic scoliosis (IS) is the most common form of spinal deformity. The pathogenesis of IS is still poorly understood. Several studies show evidence that the genetic component is determinant to the development of IS. In this setting, a crescent focus has been placed on the identification of genes, associated genetic polymorphisms, and multiple susceptibility loci. This review highlights the genes and genetic polymorphisms currently studied, identified as influential in the genesis of IS, such as MMP-3, IL-6, type I collagen, and vitamin D and estrogen receptors. We concluded that IS remains a complex disease with a polygenic background and that genetic polymorphisms are intrinsically related to this condition. Level of evidence III; Narrative Review.

Keywords: Adolescent; Scoliosis; Spine; Genetics; Polymorphism, Genetic.

INTRODUCTION

Idiopathic scoliosis (IS) is a three-dimensional spinal deformity. The Scoliosis Research Society define it clinically as spinal curvature with angulation greater than 10° in the coronal plane.1,2 It is the most common form of spinal deformity with a prevalence of approximately 2-3%, 10% of which will progress over time.3 A young age, a pattern of thoracic curvature, and skeletal immaturity increase the probability of progression,4 leading to changes in adulthood, such as cosmetic deformity, pain, and even limitation of pulmonary function.5,6

Four clinical subgroups are defined by age: infantile (up to 3 years of age), juvenile (from 4 to 10 years of age), adolescent (from 10 to 18 years of age), and adult (after 18 years of age). Adolescent IS is the most common and affects approximately 4% of the population, predominantly girls, with a gender ratio close to 8:1.7 Most cases are diagnosed between 12 and 15 years of age and one out of every six will have a progressive curve that requires active treatment.8

The etiology and pathogenesis of IS are poorly understood, largely due to the heterogeneity of related factors.9 Some of the suggested etiological factors are growth pattern deviation, neuromuscular or connective tissue changes, asymmetric growth of the limbs and trunk, changes in the sagittal plane of the spine, environmental factors, and genetic influence.10,11 Studies about the familial aggregation of IS and concordance in monozygotic twins provided the first evidence of a possible primary genetic cause of the disease.12,13 Family-based genetic linkage studies have identified multiple loci of susceptibility.14,15 Loci have been identified that implicate possible...
biological processes related to the pathogenesis of IS, such as axial modeling, cartilage formation, and growth asymmetries.¹

New information has revealed that the genetic component is a determinant in the development of IS and that genetic polymorphisms are intrinsically related to its progression.² In the last two decades, analyses of genomic association in the genomes of individuals with IS have been conducted in an attempt to identify genetic variations that predict the disease.¹⁴ Therefore, the objective of this study is to gather the main polymorphisms currently being studied through a narrative review of the literature, providing updates that can help elucidate the etiopathogenesis of IS and foster new diagnostic and therapeutic approaches to IS.

METHODS

The bibliographical research was conducted using the Medline, Scielo and Web of Science databases and systematic reviews of the Cochrane Library. Articles selected from three types of studies were included: familial aggregation studies with twins, case-control studies, and meta-analyses. Because the objective of our study was to retrieve advances in studies establishing causal relationships between IS and genetic factors, we did not determine a minimum publication date for article selection. We used “Adolescent”, “Deformity”, “Scoliosis”, “Spine”, “Genetics”, and “Polymorphism” as descriptors. Only works that addressed the correlation of genes, genetic variations, or genetic polymorphism related to structure, or deformities related to the genesis of IS, were included. Repeated articles, dissertations, theses, validation articles, and those that did not make the full text available were excluded. Works that described structural spinal changes directly related to candidate genetic mechanisms were also included. We also checked their respective references in search of additional articles. This study is a literature review not involving patients so it was not submitted to the Institutional Review Board for approval.

RESULTS

We found 27 works that met the evaluation criteria, including 8 relevant studies with twins and familial aggregation; 15 case-control studies addressing the most relevant currently associated genes - MMP-3, IL-6, the collagen and vitamin D receptor genes, and estrogen; and finally, 4 meta-analyses of the same.

DISCUSSION

Population-based association studies and genetic sequencing approaches have identified numerous loci associated with the disease, however, a lack of appropriate animal models has made understanding its biological origins difficult.¹⁷ Most of the genetic studies for adolescent IS prior to 2010 are based on candidate genes selected based on hypotheses generated from clinical observations. Most of these genes are unconfirmed or have widely varying published results, depending on the structuring of the study and/or populations tested.¹⁴ Initially, studies with twins were conducted that showed the preponderance of genetic factors in relation to the environmental factors of deformity genesis. Subsequently, case-control studies tried to identify the genes and polymorphisms involved in the genesis and progression of the disease.

Studies with twins and familial aggregation

Studies with twins and familial aggregation are usually the starting point for in-depth investigations of the genetic influence on a specific disease, since they are capable of identifying similarities between individuals with close genetic profiles (Table 1).

Wynne-Davies analyzed familial incidence of IS, reporting that 6.94% of first-degree relatives, 3.69% of second-degree relatives, and 1.55% of third-degree relatives were affected.¹⁸ In 1970, Cowell evaluated a database with 36 million names and constructed multi-generational family lineages, demonstrating that 97% of the patients with adolescent IS were related to other families with adolescent IS.¹⁹ Two years later, the same author evaluated 110 families and reported that 80% of the patients with adolescent IS had other affected family members.²⁰ In 1976, Kling analyzed a series of systematic reviews and reported an IS concordance rate of 73% in 37 pairs of monzygotic twins and 36% in 31 pairs of dizygotic twins.²¹

Czeizel and Riseborough also observed a higher occurrence of adolescent IS in first-degree relatives than in the general population.²²,²³ Data about IS in twins in the Swedish population in 2012 supported the claim that 38% of the cases that developed IS had a familial genetic component.²⁴

In Brazil, Wajchenberg, analyzed the genealogy of 100 Brazilian families that had at least one member with adolescent IS with a curve greater than or equal to 20° and reported that 33% had some family member with the same disease, 5.21% being first-degree relatives, 4.54% second-degree relatives, and 8.97% third-degree relatives.²⁵

Case-control studies – evaluations of polymorphisms

From the studies that confirmed the preponderance of genetic factors in the genesis of IS, case-control studies became necessary to analyze the possible associated polymorphisms. These polymorphisms were based on clinical observation of the patients and of the possible related structures. In 2010, K. Ward identified more than 300 single nucleotide polymorphisms (SNPs) that are related statistically to the progression of adolescent IS curves, concluding from the analysis of case-control studies that adolescent IS is a complex autosomal and polygenic disease with dominant, recessive, and codominant inheritance patterns.²⁶

In this review, we will cite the most-studied genetic polymorphisms. There are a large number of polymorphisms being studied, for which information is scarce of for which there are no systematized studies. We will review the type I collagen gene associated with the structuring of the intervertebral disc, the matrix metalloproteinase-3 gene (MMP-3) that produces enzymes capable of breaking down the disc, and genes related to bone structure – the vitamin D and estrogen receptor genes. (Table 2)

| Candidate gene | Work | Country |
|----------------|------|---------|
| MMP-3          |      |         |
|                | Aulisa et al., 2007²⁹ | Italy |
|                | Liu et al., 2010³³ | China |
|                | Mórocz et al., 2011³⁰ | Hungary |
|                | Zhao et al., 2016³¹ | China |
|                | Nikolova et al., 2016³² | Bulgaria |
|                | Su et al., 2017³³ | China |
| IL-6           |      |         |
|                | Aulisa et al., 2007³⁴ | Italy |
|                | Liu et al., 2010³⁵ | China |
|                | Mórocz et al., 2011³⁶ | Hungary |
|                | Nikolova et al., 2015³⁷ | Bulgaria |
|                | Nikolova et al., 2016³⁸ | Bulgaria |
| Collagen I     |      |         |
|                | Tótkai et al., 2015³⁹ | Turkey |
|                | Haller et al., 2016⁴⁰ | USA |
| VDR            |      |         |
|                | Suh et al., 2010⁴¹ | South Korea |
|                | Tótkai et al., 2015³⁹ | Turkey |
|                | Yin et al., 2017⁴² | China |
| Estrogen receptor |    |         |
|                | Wu et al., 2006⁴³ | China |
|                | Januszkiewicz et al., 2014⁴⁴ | Poland |
|                | Nikolova et al., 2015³⁸ | Bulgaria |
|                | Skibinska et al., 2016³⁹ | Poland |

In 1976, Kling analyzed a series of systematic reviews and reported an IS concordance rate of 73% in 37 pairs of monzygotic twins and 36% in 31 pairs of dizygotic twins.²¹

Czeizel and Riseborough also observed a higher occurrence of adolescent IS in first-degree relatives than in the general population.²²,²³ Data about IS in twins in the Swedish population in 2012 supported the claim that 38% of the cases that developed IS had a familial genetic component.²⁴

In Brazil, Wajchenberg, analyzed the genealogy of 100 Brazilian families that had at least one member with adolescent IS with a curve greater than or equal to 20° and reported that 33% had some family member with the same disease, 5.21% being first-degree relatives, 4.54% second-degree relatives, and 8.97% third-degree relatives.²⁵

Table 1. List of studies with twins and familial aggregation.

| Study | Country |
|-------|---------|
| Wynne-Davies, 1968 | USA |
| Cowell et al., 1970 | USA |
| Cowell et al., 1972 | USA |
| Riseborough et al., 1973 | USA |
| Keeling et al., 1997 | USA |
| Czeizel et al., 1978 | Hungary |
| Wajchenberg et al., 2005 | Brazil |
| Grauers et al., 2012 | Sweden |

Table 2. List of candidate genes and published works.
MMP-3
Matrix metalloproteinases (MMPs), such as MMP-3, are potent proteoglycan-degrading enzymes that play an important role in intervertebral disc degeneration (IVDD). Many studies point to a correlation between the degree of histological degeneration and MMP-3 that has increased significantly in the nuclei pulposi (NP), altering the constitution of their matrix. This change has been associated with the development of conditions that favor the development of IS.28,30

Genetic polymorphism in the MMP-3 promoter gene has been reported to be involved in MMP-3 expression. Gene regulation and certain alleles, such as 5A, have been indicated as risk factors for the acceleration of degenerative lumbar disc changes.12,23 Thus, the relationship between MMP-3 polymorphisms and IS has been one of the most investigated associations and the one presented in the greater number of studies. However, in spite of those studies reporting this possible action, there is no consensus about the already proven relationship between MMP-3 and the onset or progression of IS.32,33

IL-6
Several inflammation mediators, including interleukin 1 (IL-1) and interleukin 6 (IL-6) have been implicated in IS etiopathogenesis.34,36 IL-6 is a potent pro-inflammatory mediator and is involved in lumbar disc herniation. Despite this, the exact role of IL-6 in IS has not been fully elucidated.18 It has been shown that the NP of spondiotic discs respond to exogenous pro-inflammatory stimuli by secreting increased amounts of IL-6 and other pro-inflammatory cytokines.32,33

Studies documenting single nucleotide polymorphisms (SNP) in the gene of IL-6 have been widely performed, aimed at documenting what the real role of this polymorphism is in inflammatory imbalance and in its influence on the matrix changes in IVDD.30,32,36
The hypothesis is that this polymorphism leads to imbalance of the pro-inflammatory cytokines and, thus, accelerates inflammation, because the expression of IL-6 RNA-m is in fact highly correlated with the levels of IL-1 beta RNA-m, tumor necrosis factor alpha (TNF-α), IL-8, MMP-3, MMP-9, and MMP-12.37,38

Type 1 Collagen
Alterations in the codification genes of type I collagen would lead to the structural imbalance of IVDD, and in turn it would morphologically alter the distribution of loads, leading to abnormal curvatures.37,38 Studies have been developed with the intention of defining the degree of influence of collagen alterations, especially that of type I in the determination of the curvatures present in IS, particularly in the annulus fibrosis (AF) where collagen I creates a network of fibers that function to retain the NP which in turn has reticulated collagen IX fibers and collagen II fibers to provide stability.45

Vitamin D receptor
1,25-dihydroxyvitamin D3 plays a central role in skeletal metabolism, binding to its nuclear steroid receptor, the vitamin D receptor (VDR). About 40% of skeletal mass is acquired during puberty and vitamin D performs an important role in this gain during adolescence. In addition, it has been reported that girls with hypovitaminosis D run the risk of not achieving the maximum bone mass desired for that age, which suggests that the genes of any component of the vitamin D endocrine system is a candidate locus as gene influencing vertebral formation.41

Researchers have analyzed the association between polymorphisms of the VDR and DMO genes and their influence on abnormal curvatures, since modified receptors would alter the physiological metabolism of vitamin D.42 However, studies to demonstrate the relationship between bone mass in girls with IS and polymorphism of VDR are still lacking. What we have in the literature is the female predominance as a risk factor for the development EI as already presented.43

Estrogen receptor
One of the theories about the etiology of IS is associated with the potential influence of estrogens and estrogen receptors in the formation of IVDD, of vertebral structures, and in the cell function of the paraspinal musculature. Estrogens affect muscle tissue function in terms of adaptation to angiogenesis and myogenesis resistance training.44 Furthermore, the estrogens repress bone remodeling and the balance of control between bone formation and reabsorption.45

Expression of the estrogen receptor was detected in human osteoblasts and osteoclasts and mutation in the gene of codifier of the estrogen receptors caused bone loss and delay in skeletal growth in human beings. The association between polymorphisms of the estrogen receptor gene and bone density was reported by different investigators, but this relationship is not firmly established.46,47

Systematic reviews and meta-analyses
Through the systematic review, we discovered that IS has multifactorial hereditary characteristics and multiple genes are affected by the development of the disease. Genes discovered as being linked to IS include SNTG1 at 8q11.22, ESR1 at 6q25.1, and CHD7 at 8q12.1.3

Out of 50 studies, KF Gorman et al. identified 34 studies of candidate genes (6 linkages, 28 associations) and 16 complete genome studies (14 linkages based on linkage, 2 studies of complete genome association). The findings involved genes related to the structure of connective tissue, such as type I collagen, the formation/metabolism of MMP-3, inflammation signalization pathways, such as IL-6, and secondary pathways, such as melatonin and axonal guidance pathways. It should be noted that the variability among the results suggests ethnic and/or genetic heterogeneity.3

It was concluded that the main difficulty in IS research is phenotypical and genetic heterogeneity. Weak studies were over-represented in the genetics research. The use of biological models and restrictive clinical definitions may help to separate variations and increase the potency of studies to detect or confirm an effect.14

Using meta-analysis aimed at detecting whether the main genetic polymorphisms, those of MMP-3 and IL-6, are correlated with adolescent IS, CK Kepler concluded that, overall, there is no significant association between IL-6 genetic polymorphism and the risk of adolescent IS. A significant association was observed in the homozygotic model of genetic MMP-3 polymorphism. When stratified into Caucasian and Asiatic populations, a positive association was observed in the Caucasian population, while there was no significant association observed in the Asiatic population. This study concluded that certain genotypes of genetic MMP-3 polymorphism were associated with adolescent IS, especially in the Caucasian population. However, no significant association was detected between IL-6 genetic polymorphism and adolescent IS.31 Meta-analyses of the involvement of VDR and estrogen receptors lack sufficient statistical data to indicate any correlation.

Few studies to date have applied complete genome association testing in populations of individuals with IS. This is mainly related to the high cost of this type of testing in large populations, which is believed to be necessary for the detection of significant results that could be applied to the general population. This reflects the real complexity of IS. Although the disease aggregates within families, it does not segregate clearly like a Mendelian disease.45 Genetic and environmental interactions may be essential to the expression of IS. Individual alleles may vary from one affected subject to another and they may not be either necessary or sufficient to be caused by nature. Knowledge of the causal mechanisms of a complex genetic disease like IS is an essential step towards the development of efforts for focal/practical application of the Human Genome Project to clinical medicine.49

CONCLUSION
It is known that IS is a polygenic disease influenced by various genes that codify essential spinal support proteins. It is not possible to determine precisely all the genes involved, but certain polymorphisms are already being related, such as those of the metalloproteinases, collagen codifiers, and estrogen and
vitamin D receptors. In this way, the polymorphism of these genes and their expression in tissues still require further investigation in order to propose more assertive therapeutic targets that corroborate with the non-progression of the disease. The identification of diagnostically and genetically valuable molecular markers can be useful in early detection in children at risk for developing IS and for prognosis, with more certain means for predicting the risk of progression, for avoiding follow-ups, potentially unnecessary radiographs and treatments, and for the application of more effective treatment.

All authors declare no potential conflict of interest related to this article.