Single nucleotide polymorphisms of intervertebral disc collagens and prospects for their correction

Eugene G. Pedachenko 1, Iryna G. Vasilyeva 2

1 Department of Minimvasive and Laser Spinal Neurosurgery, Romodanov Neurosurgery Institute, Kyiv, Ukraine
2 Department of Neurobiochemistry, Romodanov Neurosurgery Institute, Kyiv, Ukraine

Received: 09 September 2022
Accepted: 17 November 2022

Address for correspondence:
Iryna G. Vasilyeva, Department of Neurobiochemistry, Romodanov Neurosurgery Institute, 32 Platona Maiborody st., Kyiv, 04050, Ukraine, e-mail: vigvasileva@gmail.com

The main functions of the intervertebral disc (IVD) are ensured by the reliable integration of three structures of the IVD: the annulus fibrosus (AF), the hydrated nucleus pulposus (NP) and the two cartilaginous end plates (CEP). All molecular components are involved in the integration of the three anatomical structures of the IVD, however, the most important biomechanical properties - resistance to rupture/stretching/shift, resistance to static axial loads are mostly determined by collagens.

The unique properties of collagens depend on the amino acid sequence of the three alpha (α) chains, which, after spiralization and condensation, form the collagen molecule - tropocollagen. The amino acid sequence contains all the necessary information for spiralization, modification, secretion of tropocollagen, its processing, condensation into fibrils and fibers according to the self-assembly principle. Changes in the primary amino acid sequence, depending on the substitution itself and its localization, lead to disruption of the stages of tropocollagen formation, its extracellular processing, and condensation.

Currently, most of the research is devoted to the study of polymorphisms in the genes of IVD collagen types I, II, IX and XI. Algorithms for using information about genetic polymorphisms of collagen genes are only being formed. Data on genetic variation are often conflicting. An important aspect is the homogeneity of the study group by age, ethnicity, gender, as well as by the type of degenerative changes. There is also insufficient data on the effect of polymorphism on the properties of the collagen molecule, which greatly complicates the creation of standards for therapeutic correction.

This literature review is devoted to the consideration of new data on collagen genes polymorphisms, the impact of these polymorphisms on integrative relationships in IVD structures, as well as the prospects for the correction of genetic abnormalities.

Key words: intervertebral disc; collagen type I; collagen type II; collagen type IX; collagen type XI.

Aging of intervertebral disc tissue (IVD) begins earlier than aging of other tissues. Its main manifestation is the inability of the IVD to withstand axial and torsional loads, to perform the cushioning function in motion. The main functions of the IVD are provided by the integration of three structures: the annulus fibrosus (AF), the nucleus pulposus (NP) and the two cartilaginous end plates (CEPs). Each of these structures performs different functions and, according to these functions, has a different composition and extracellular matrix structure (ECM). The most important biomechanical properties - resistance to tearing, stretching, displacement and resistance to static axial loads are largely determined by collagens [1, 2].

In AF collagen type I (COL1) provides tensile strength, in NP collagen type II (COL2) - depreciation properties, in CEP the combination of COL1 and COL2 forms a strong structure which ensures the diffusion of plastic substances. Collagen type IX (COL9) and collagen type XI (COL11), which regulate fiber thickness, collagen type VI (COL6) and collagen type IV (COL4), which provide spatial communication between the components of the IVD disc, and also determine the micromechanical properties of the pericellular space [3,4].

Collagen molecules are formed as a result of the interaction of three α-chains to form a right-handed spiral. After stages of modification involving specific enzymes, the collagen monomer - tropocollagen is formed, which is secreted into the extracellular space. Tropocollagen molecules interact with each other, forming fibrils that assemble into collagen fibers and into complexes with other ECM elements. The amino acid sequence of α-chains of collagen contains all the information necessary to provide the processes of collagen fiber formation according to the principle of self-assembly. Changes in the primary amino acids sequence result in disruption of the stages of tropocollagen formation, its extracellular processing and condensation [5]. Collagen fiber configuration depends equally on structural collagens COL1 and COL2 and regulatory collagens COL9.
and COL11, which determine the lateral increase of the fiber, its integrative and functional properties.

Degenerative changes are thought to be a slow process of accumulation of micro-changes under physiological stress. Genetic factors are the trigger for the acceleration of degenerative changes of IVD. These can be point mutations (missense, nonsense, deletions, insertions, insertion-deletions, mutations leading to a frameshift mutations) and mutations with complex rearrangements. Mutations in the collagen molecule have different significance for modifying tissue phenotype and tissue properties. Some misassembled collagen molecules accumulate in the cell, leading to apoptosis. Other mutations do not prevent triple helix formation and tropocollagen secretion, but significantly change the physical and chemical properties of ECM collagen fibers [6–9]. Special attention is paid to genetic modifications of collagens integrating IVD into a single structure.

Depending on the localization of the substitution in the α-chain, clinical manifestations are classified as mild, moderate, and severe. There are also «silent» and lethal ones that are not compatible with the development of the embryo, so their clinical manifestations are unknown. Mutations affecting positions closer to the C-terminus of the α-chain molecule affect the initiation of the formation of the three-chain molecule. Such mutations correspond to more pronounced clinical phenotypes [7]. Mutations affecting positions closer to the N-terminus of the molecule correspond to milder clinical manifestations. Changes in other chain localizations can directly affect the ability of molecules to form adequate supramolecular structures, interact with ECM components and bioactive molecules, which also leads to pathology [10, 11].

The literature review is devoted to the consideration of new data on polymorphisms of the IVD collagen genes COL1, COL2, COL9 and COL11, the influence of these polymorphisms on the integrative bonds in the structures of the IVD, as well as the prospects for the correction of genetic abnormalities.

**Type I collagen**

COL1 is a fibrillar heterotrimer consisting of two COL1A1 chains and one COL1A2 chain. The COL1A1 and COL1A2 protein genes are located in positions 17q21.31–q22 and 7q22.1. The primary amino acid sequence of COL1 α-chains is important for ensuring the strength of AF IVD. The COL1A1rs1800012 polymorphism in the Sp1-binding site of the first intron (G>T substitution at position +1245) has been shown to result in increased mRNA expression level and, correspondingly in COL1A1 protein. An increase in the content of COL1A1 chains disrupts the balance of the 2:1 COL1A1 and COL1A2 ratio, resulting in the formation of homotrimers from three COL1A1 chains [12, 13]. Homotrimers have a more rigid structure determined by the primary amino acid sequence of COL1A1 α-helices. Morphologically, this is manifested in their spear-shaped form. Homotrimers and heterotrimers can assemble into heterofibrils, which affects the biomechanical properties of the tissue [14, 15].

Increased resistance of COL1A1 homotrimer to the action of proteases was established [16]. It is assumed that the heterofibrils formed as a result of the condensation of homofibrils and heterofibrils, are less strong and less stable, which is the reason for the accelerated IVD degeneration (DIVD) [12, 17].

Assumption of the significance of the COL1A1rs1800012 gene polymorphism in DIVD has been confirmed in population-based studies. A genetic analysis performed to determine the correlation between DIVD and the presence of COL1A1 TT and GT genotypes among Greek military personnel diagnosed with early lumbar disc degeneration showed that 33.3% of this contingent were carriers of the TT genotype, and, conversely, in the control group, the genotype TT COL1A1 Sp1 was not detected [18].

Dutch researchers showed that in a group of 966 patients aged over 65 years, carriers of the TT genotype had a 3.6 times higher incidence of DIVD compared to GT or GG genotype carriers [19].

It was also shown that patients with the TT genotype have a higher stage of Pfirman degeneration than patients with COL1A1 GT genotype. A higher stage of DIVD compared to the control group of patients was also found among GT genotype carriers. These studies indicate that the COL1A1 gene polymorphism in the Sp1 site is associated not only with an increased risk of DIVD, but also with a more severe form of degeneration [8].

In studies involving twins in Finland, a specific DIVD phenotype (intense magnetic resonance imaging (MRI) signal) was correlated with the GT COL1A1rs2075555 genotype. Studies also showed that 66.7% of individuals with DIVD had the GT genotype, while in the control group – 41.7% [20].

**Collagen type II**

The COL2 protein, a fibrillar homotrimer, is the structural basis of the NP. The COL2A1 gene is localized in the region 12q13.11–q13.2. COL2 accounts for more than 85% in the NP tissue. This collagen is characterized by a wide range of intermolecular interactions, which ensure the reticular structure of the NP and its hydration. COL2 is covalently bound to COL9 in NP. A study of interprotein interactions has shown that COL2 is key in the process of IVD NP degeneration. Mutations in the COL2A1 gene cause a spectrum of phenotypic manifestations, particularly cartilage and bone. In experiments using COL2A1-0 transgenic mice, there is no NP [21].

The association of DIVD with the COL2A1 gene polymorphism has been demonstrated in several population-based studies. Thus, a case-control association study of single nucleotide polymorphism COL2A1rs1793937C>G/intron, rs1793953G>A,C,T/sp1 intron, rs2276454 (2295 C>T, p.Gly76) among individuals of the Chinese Han population showed that the frequency of COL2A1rs1793953 and rs2276454 alleles in the group, the case after back injury incidents differs significantly in the degree of degeneration of the IVD according to Pfirman (stage III, IV and V) from the control group (stage I and II). The COL2A1rs2276454 polymorphism is also an independent risk factor for DIVD [22]. A study of the presence of the COL2A1rs2276454 polymorphism among DIVD patients in a hospital in China found a correlation between its frequency and the number of cases of degenerative changes. The COL2A1rs2276454 polymorphism correlates with the degree of disc degeneration according to Schneiderman, the number of degenerated segments, the type of herniation, and the
number of herniated segments, whereas the presence of rs2070739 polymorphism correlates with the degree of degeneration according to Schneiderman in men [23].

The COL2A1 gene polymorphism was studied among Korean patients with degenerative lumbar scoliosis that progressed after 50–60 years of age. Genotyping revealed a correlation of the presence of the COL2A1rs2276454 allele with an increased risk of degenerative lumbar scoliosis.

**Collagen type IX**

COL9 is a non-fibrillar heterotrimer, its chains COL9A1, COL9A2, COL9A3 are encoded by the genes COL9A1 (chromosome 6q), COL9A2 (chromosome 1p), COL9A3 (chromosome 20q), respectively. COL9 consists of three collagenous coiled, fibrillar domains and four non-collagenous (globular) domains. The globular domain of COL9 at the N-terminus is immersed in the interfibrillar space of COL2 collagen fibers, and the fibrillar domains interact with other ECM components and cell membrane proteins. Thus, COL9 modulates the surface properties of fibrils, regulates their linear and lateral growth. COL9 and COL2 form a reticular structure that ensures the preservation of the spherical shape [24]. Mutations or polymorphisms lead to COL9 dysfunction, which may be the cause of accelerated DIVD. Transgenic mice with a mutation in the COL9A1 gene show multiple degenerative changes in IVD [25].

Among the clinical variants, the correlation between accelerated DIVD and polymorphisms of the COL9A2 and COL9A3 genes is considered. A clinically significant polymorphism of the COL9A2 gene (rs137853213, Trp2 allele) leads to the substitution of the neutral amino acid glutamine for the aromatic hydrophobic amino acid tryptophan at the 326th position. Experiments have shown the effect of the substitution on the mechanical properties of the tissue: the swelling pressure of the disc tissue with the presence of the Trp2 allele is significantly lower than that of the tissue without the polymorphism [26].

Experimental data have been confirmed by clinical studies. In the Finnish population with DIVD, the Trp2 allele occurred three times more often than in the control group. Another study showed that the frequency of Trp2 polymorphism is higher in patients with radial ruptures of lumbar IVD [27]. The results of studies in a group of Japanese patients under 40 years of age indicate that the presence of the Trp2 allele is accompanied by more severe damage to the IVD tissue than in patients of the control group [28]. However, another study found that the Trp2 allele among individuals in the Japanese population is common and not associated with DIVD [29]. Subgroup meta-analysis also showed that the Trp2 allele is common in the Asian population. Although Z. Zhang et al. in a study of a large group of Chinese patients found a 2.4 increased risk of DIVD among individuals with Trp2 aged 30–39 years. The difference in results indicates the importance of forming similar groups by age, ethnicity and gender [30]. A meta-analysis showed an association of degenerative processes in IVD with the COL9A2rs137853213 polymorphism (Trp2 allele) for the Caucasian population [31]. A meta-analysis of the association of COL9A2rs12077781, rs12722877 and rs7533552 polymorphisms with DIVD did not reveal such an association.

The result of the sequence variation COL9A3rs61734651 (Trp3 allele) is the substitution of arginine, a positively charged amino acid, with tryptophan at position 103, which reduces the solubility of COL9, modifies the intercellular interactions of ECM, affects the mechanical properties, disrupts the conformation of the collagen triple helix, affects the formation of COL2 and COL9 heterofibers, and interaction with lysyl oxidase, which catalyzes the formation of hydroxylysine, which, after glycosylation is involved in the formation of interchain bonds [32]. The results of clinical studies indicate the correlation of the Trp3 allele with spinal stenosis and spondylolisthesis - types of DIVD [33]. In some studies, the correlation between the polymorphism and degenerative changes of IVD was found only in men. Other studies have found an association of the Trp3 allele and DIVD depending on ethnicity [34]. Meta-analysis also showed a divergent frequency of the presence of Trp3 allele in Caucasian and Asian populations. In contrast to the Trp2 allele, the Trp3 allele is more common in Asian population and is associated with DIVD [35].

**Collagen type XI**

COL11 is a fibrillar heterotrimer consisting of three chains COL11A1, COL11A2 and COL11A3. The two α-chains are encoded by the COL11A1 and COL11A2 genes (loci 1p21, 6p21.3, respectively). COL11A3 is a hyperglycosylated COL2A1 and is therefore expressed involving the COL2A1 gene (12q13.11–q13.2). COL11 is a regulator of the lateral increase of COL2 fibrils. It forms complexes with COL2 and COL9 and thus regulates fibrils diameter [36]. COL11 also binds to cell surface proteoglycans, which ensures tissue cohesion and integrity [2, 37].

COL11A1 gene aberrations have been shown to correlate with degenerative changes of IVD. The polymorphism COL11A1rs1676486, 4603 T>C results in the substitution of the uncharged amino acid serine at position 1535 with the nonpolar aliphatic amino acid proline. Proline is an amino acid involved in the stabilization of the α-helix and interchain bonds. The substitution affects coiling and, consequently, the conformation of tropocollagen and fibrils, and the transcripts have reduced stability. The association between the COL11A1rs1676486 polymorphism and lumbar IVD herniations was found in patients of the Japanese population. The frequency of polymorphism in the case group was 1.5 times higher than in the control group [38]. The association of COL11A1 genetic variants with the risk of lumbar IVD herniation was investigated involving 647 Chinese patients and 532 healthy donors. The COL11A1rs1676486 risk allele was detected more frequently in patients than in the control group. Patients also had significantly higher levels of DIVD [39]. The association of 4603 C/T with cervical DIVD in Japanese wrestlers has been shown. CT+TT variants prevailed in the group with DIVD of the cervical region [40]. A meta-analysis confirmed the association of the rs1676486 polymorphism of the COL11A1 gene with the risk of degenerative changes in IVD in homozygous, dominant, and recessive patterns of heredity [41]. The COL11A1rs1337185 allele has been shown to be a risk
factor for lumbar degenerative changes, such as hernias, stenosis, spondylolisthesis [42]. In a study involving 588 Finnish men, variations in intron 9 of the COL11A2 gene were found to be associated with MRI-diagnosed disc bulges and MRI signal changes [43]. It is believed that polymorphism can be important for the formation of unstable transcripts, which is associated with aberrations in the quantitative ratio of proteins and consequently ECM structure [20].

A strong correlation of degenerative changes of lumbar discs with the COL11A2rs2071025 polymorphism was detected. Male and female carriers of the COL11A2rs2071025 allele (Chinese population) have an increased risk of degeneration. The COL11A2rs986522 gene polymorphism is also associated with a significantly increased risk of lumbar disc degeneration in women [44].

**Current trends of therapy of degenerative processes in the IVD in genetic collagen abnormalities**

The goal of future therapy is the complete restoration of the biomechanical properties of the spine. Different areas of IVD tissue repair are currently being developed. The focus of injection biomolecular therapy is the transforming growth factor and bone morphogenic protein superfamilies. The use of post-translational regulators of protein expression is also considered: platelet-rich plasma, non-coding RNAs, miRNAs, long-chain non-coding RNAs, circular RNAs. Studies are conducted primarily *in vitro*. Clinical trials of the efficacy of intradiscal 2-time administration of recombinant human growth/differentiation factor rhGDF-5 showed good tolerability and safety [45]. Multicenter clinical trials of the efficacy of synthetic factor κB AMG0103 (AMG0103 in Subjects with Chronic Discogenic Lumbar Back Pain ClinicalTrials.gov Identifier: NCT03263611) involving patients with discogenic lumbar pain are ongoing.

Therapy using stem cell-derived exosomes is being studied. Exosomes originate from cellular multivesicular endosomes. Exosomes are secreted by various cell types and contain bioactive molecules that can be transported from cell to cell and thus deliver more than 4,000 types of bioactive molecules (proteins, mRNA, miRNA). Due to a wide range of bioactive molecules, exosomes have a powerful regenerative potential. The interest in these formations is due to the fact that they retain properties under extreme conditions, and demonstrate resistance to oxidative stress [46].

In genetic abnormalities underlying the degenerative changes of IVD, the use of bioactive drugs may be ineffective. Injections of factors directly into the disc improve the condition of the disc in experimental models, but their effectiveness is possible only in the presence of intact cellular composition without genetic abnormalities. Disc cells with genetic abnormalities, when stimulated by factors, will synthesize genetically altered components that are unable to ensure the full functioning of the IVD.

Significant limitations also exist for the use of a gene therapy strategy, when the target gene must be transfected into IVD cells, expressed and restore the structure of the disc tissue. With the current development of the method, transfection of target genes into IVD cells is not efficient enough, and transfected gene constructs based on adenovirus, adeno-associated virus, baculovirus, and lentivirus can damage the genome or cause oncotransformation.

The promising CRISPR/Cas9 genome editing method involves inserting the desired gene into a specific locus of the genome and triggering expression of the useful gene. But using this method can lead to errors. Guide RNAs should be precisely "targeted" to the desired parts of the genome, and then, using the Cas9 nuclease, deletion or insertion of the necessary parts of the genome should be carried out. This method has technical challenges that must be overcome. For example, guide RNA may not complex with target sites. There is a problem with the number of target cells and their quality. The method is believed to be promising, but the technology is still imperfect and it is unlikely that CRISPR/Cas9 can be used to correct the genetic code in DIVD in the near future.

For patients with DIVD associated with genetic disorders, cell therapy is the optimal strategy. Theoretically, restoration of the biomechanical properties of the IVD tissue in the presence of genetic abnormalities can be achieved by repopulating IVD tissue with cells synthesizing the "correct" ECM components. Thus, allogeneic histocompatible cells with the "correct" genotype will be more effective.

The most studies are devoted to mesenchymal stem cells (MSCs), hematopoietic stem cells, adipose tissue stem cells, discogenic cells, and fibroblasts. MSCs are adult stem cells found in many tissues: bone marrow, adipose tissue, liver, intestines, muscle tissue, skin, umbilical cord. MSCs can differentiate into different cell types, particularly chondrocytes and fibroblasts. A useful feature of MSCs is their ability to migrate to the sites of damage and low immunogenicity, due to which immunosuppressive agents are not required for allogeneic transplantation. MSCs secrete cytokines, anti-inflammatory molecules that positively affects the environment during transplantation [47].

Despite a large number of clinical trials, a standardized method of DIVD therapy based on cell therapy has not yet been created.

Recently, several clinical trials of intradiscal injection of autologous MSCs have been conducted, which proved the tolerability and safety of the procedure. Autologous MSCs obtained by aspiration from the bone marrow or iliac bone, when injected into the intradiscal space, contributed to a significant improvement of patients’ condition and reduction of pain syndrome [48,49]. These studies also showed the importance of the quantitative aspect: patients who received >2000 colony-forming units per milliliter of bone marrow aspirate demonstrated better results in terms of pain reduction. Good results of pain elimination have also been demonstrated in transplantation of autologous cells obtained after discectomy [50].

In addition to MSCs, the potential of other stem cells is being studied for DIVD therapy. Hematopoietic stem cells have a good potential for proliferation and differentiation, but the hypoxic environment of DIVD is unacceptable for them [51]. Adipose stem cells, when used in experimental models, have demonstrated the ability to increase the expression of proteoglycans, aggrecan and generally stimulate matrix anabolism and minimal ossification, but their clinical trials have not been conducted so far [52]. Positive results were
obtained when using adipose tissue stem cells together with platelet-rich plasma [53].

Allogeneic MSCs are the most promising for solving the problems of genetic abnormalities of IVD. In 2017, clinical trials involving 24 patients were completed, which showed the safety, tolerability and efficacy of intradiscal injection of allogeneic bone marrow MSCs [54]. Clinical trials of intradiscal injection of allogeneic mesenchymal progenitor cells with the participation of patients with pain syndromes in the lumbar region are ongoing. It introduction of conditioned medium obtained during MSCs cultivation is also promising [55].

Therefore, the complete restoration of the biomechanical properties of the IVD in the case of genetic abnormalities is a task for the future. Currently, science is accumulating data on the significance of genetic polymorphisms in IVD functioning disorders. At the current stage, information about the genetic features of the ECM structure is used for recommendations on lifestyle management capable of preventing, mitigating or delaying degenerative processes.

**Information disclosure**

**Conflict of interest**
The authors declare no conflict of interest.

**Ethical approval**
This article is a literature review, therefore no ethics committee approval was required.

**Funding**
The study was conducted without sponsorship.

**References**


http://theunj.org


