Genetic Determinism of Primary Early-Onset Osteoarthritis

Juliette Aury-Landas, Christian Marcelli, Sylvain Leclercq, Karim Boumédiene, Catherine Baugé

Title

Genetic determinism to primary early-onset osteoarthritis

Authorship

Juliette AURY-LANDAS[1,2], Christian MARCELLI[1,3,4,5], Sylvain LECLERCQ[1,2,6], Karim BOUMEDIENE[1,2], Catherine BAUGE[1,2]

[1]Normandie Univ, France

[2]UNICAEN, EA4652 MILPAT, BIOCONNECT team, F-14032 Caen, France

[3]UNICAEN, COMETE, Caen, 14032, France

[4]Inserm, U1075 COMETE, Caen, 14032, France

[5]CHU de Caen, Department of Rheumatology, Caen, 14000, France

[6]Service de chirurgie orthopédique, Clinique Saint-Martin, Caen, France

Corresponding author: Juliette AURY-LANDAS (juliette.aury-landas@unicaen.fr) and Catherine BAUGE (catherine.bauge@unicaen.fr)

The authors declare no competing financial interests.

Keywords

Early-onset osteoarthritis, genetics, mutation

Abstract

Osteoarthritis (OA) is the most common joint disease observed worldwide. A minority of cases correspond to familial presentation characterized by early-onset forms and are genetically heterogeneous.

This review brings a new point of view of the molecular basis of OA by focusing on gene mutations responsible for early-onset OA determinism. Recently, thanks to whole-exome sequencing, a gain of function mutation in *TNFRSF11B* gene was identified in two distant family members with EO-OA, opening new therapeutic perspectives for OA. Indeed, unravelling the molecular basis of rare Mendelian OA forms will help improve the understanding of molecular processes involved in OA pathogenesis and to ameliorate patient diagnosis, management and therapy.

Primary osteoarthritis, a genetic disease?

Osteoarthritis (OA, see glossary) is the most common joint disease observed worldwide and affects nearly 27 million US adults [1]. It is the main source of morbidity in developed countries. Symptoms, when present, include pain and reduced mobility causing partial or complete inability to work, difficulties with activities of daily living and impaired quality of life, leading to a very important social-economic cost. OA is mainly characterized by the gradual loss of articular cartilage, particularly affecting the knee, hip, spine, hand and foot (Figure 1). However, OA is not only a disease of cartilage. It affects all the tissues of the joint, including synovium, subchondral bone, capsule, ligaments, periarticular muscles and the sensory nerves whose termini lie within these tissues [2]. Currently, the majority of OA cases are considered as primary, defined as OA lacking underlying factors and structural abnormalities. By contrast, secondary OA results from various factors such as traumatic injury, specific anatomical deformities in the joint or specific abnormalities of the cartilage extracellular matrix (ECM) [3]. Despite the high prevalence of OA and its substantial public health impact, genetic basis of primary OA pathogenesis are not fully elucidated. No efficient treatments are currently available, since they only slow the progression of the disease and reduce pain. A better understanding of the molecular mechanisms and pathways involved in primary OA genesis will help provide improved patient diagnosis, early management, and the development of innovative therapies.

Premature OA is rarely present without other clinical symptoms and is mainly observed as a part of **syndromes**. In these cases, **early-onset OA** (EO-OA) mostly develops secondary to inflammatory disease or to biomechanical defects, such as **osteochondrodysplasia**. Non-syndromic EO-OA families without dysplasia or other causative pathology are extremely rare. The phenotype of these rare OA families resembles common OA at later ages in the population

except for the early age of onset (20–50 years) [4]. However, identification and screening of such families is highly important and should rely on strict inclusion criteria (see Box 2 for criteria that we propose for future genetic studies). In fact, they may reveal causal genes responsible for the genetic determinism to EO-OA and lead to a better understanding of the physiopathological mechanisms involved in OA development.

First, this review describes the evidence for a genetic component in OA. Second, it focuses only on causal genetic alterations for the determinism of EO-OA, and aims to clearly distinguish causal mutations for EO-OA and those responsible for multiple phenotypes. In fact, most of the pathogenic mutations detected in EO-OA have been identified in patients harboring a syndromic presentation, including EO-OA, but also presenting other prominent secondary features such as ocular, auditory or aortic abnormalities. The presence of multiple clinical features simultaneously explains the difficulty to clearly conclude on the causal role of mutations for the genetic determinism of EO-OA. However, thanks to the use of strict inclusion criteria for cases and controls, and to whole-exome sequencing strategy, a gain of function mutation in *TNFRSF11B* gene was recently identified, leading to innovative therapy perspectives for OA [5].

Genetic basis for primary osteoarthritis

Primary OA is a complex disease, caused by a combination of three main risk factors: genetics, environmental factors and age [3]. Evidence for a genetic component in OA comes from family studies, twin studies and animal models. The first evidence for a strong genetic contribution to OA was first reported in 1941 by Stecher *et al.* who demonstrated that **Heberden's nodes** were three times more common in sisters of OA patients than in the general population [6]. Subsequently this was confirmed in families with generalized OA with the observation that first

degree relatives were twice as likely to present radiographically generalized OA as well [7]. In fact, susceptibility to common OA, which occurs at **late-onset**, is modulated by genetic risk factors, with single nucleotide polymorphisms more frequently observed in patients than in the general population. In 1996, Spector *et al.* estimated the influence of genetic factors between 39 and 65% in twins with hand and knee OA [8]. Genetic risk factors associated with primary OA susceptibility (common OA) have been extensively reviewed [9–14]. To date, 21 independent OA susceptibility loci were established based on genome-wide significance in GWAS ($p \leq 5\times10^{-8}$) and/or proven functional involvement in OA by follow-up studies [15]. By contrast, rare EO-OA forms represent Mendelian diseases with an autosomal dominant inheritance pattern, as reported in families with EO-OA [16–19]. Genetic determinism to EO-OA forms are caused by single gene mutations, which are only carried by patients. In 1993, Nakata *et al.* reported that transgenic mice with a heterozygous central deletion of *COL9A1*, developed OA without **chondrodysplasia** [20]. Since then, several **animal models** have been developed to explore and confirm the causal role of mutations identified in patients, at the genetic level [21]. Indeed, the role of genetics in idiopathic EO-OA pathogenesis has now been clearly established. By contrast, findings from mouse models give new insight into OA pathogenesis and could be further explored in human. As example, a deficiency of mitogen-inducible gene-6 (Mig-6) in mice leads to the development of an EO-OA-like disorder in multiple synovial joints. Recently, this model provides insight into the critical role of chondrocytes in developing the OA-like disorder in the knees, and into the requirement of other cell types for full development of the Mig-6–deficient joint phenotype [22].

Several approaches have been used to unravel the molecular basis of EO-OA. Thanks to technological improvements, the scale and resolution of screened genomic regions have been enhanced. Target studies of a single candidate genes (e.g. using CSGE or *conformation sensitive*

gel electrophoresis, RFLP or *restriction fragment length polymorphisms* or Sanger sequencing) have been replaced by high-throughput screening (e.g. genome-wide linkage analysis or whole exome sequencing). Mutations are generally identified in families harboring a severe phenotype and subsequently validated in replication cohorts of cases with the same phenotype. In patients with EO-OA, these strategies have led to the identification of mutations in genes encoding components of the ECM, the TGF-β pathway, vesicular transport or bone remodeling (Figure 2, Key Figure). Distinction between mutated genes identified in idiopathic EO-OA patients (Table 2) and those identified in patients having syndromic forms of EO-OA (Table 3) is clearly established.

Mutations in genes encoding collagen proteins

The collagen type II, alpha 1 gene (*COL2A1*), encoding one of the three chains of type II collagen (Box 1), has been targeted for extensive analysis in degenerative diseases of the joint due to the abundance of type II collagen in **epiphyseal** and articular cartilage tissue. Mutations in *COL2A1* gene have been identified in patients with various clinical features and are known as type II collagenopathies [23] (Table 1), for a phenotypic spectrum description see [24]. All of these syndromes are characterized by premature OA. As non-syndromic EO-OA (sporadic or familial) are infrequent, *COL2A1* mutations have been rarely screened or identified in these patients. In 1989, a genetic linkage between *COL2A1* gene and familial EO-OA was described in two unrelated Finnish families presenting autosomal dominant inheritance of OA with a mean age of symptoms of 38 years [16]. Patients had radiographically verified polyarticular OA, without evidence of chondrodysplasia. Furthermore, **spondyloepiphyseal dysplasia**, **Stickler syndrome** (STL), or any other congenital cartilage abnormalities were excluded. This linkage was also confirmed in an unrelated family presenting OA with mild chondrodysplasia [25]. Despite exhibiting chondrodysplasia of the metatarsal head and spine, osteoarthritic

changes developed progressively in many joints without any epiphyseal deformities or other evidence of chondrodysplasia. In addition, common causes of secondary OA were excluded. This suggested that generalized OA was the main trait in this family and that OA was probably non-syndromic. A missense mutation (c.2155C>T, p.Arg719Cys) was subsequently found in this family, cosegregating with phenotype and absent in 57 unrelated individuals [26]. In 1995, *COL2A1* mutations were detected in one of 45 patients with familial EO-OA, and in more than 20% of 36 patients with moderate to severe chondrodysplasias [27]. Therefore, *COL2A1* mutations are not exclusive to non-syndromic EO-OA. However, several patients with *COL2A1* mutations, primarily manifested as generalized OA in late childhood or early adulthood have been reported [19,28], confirming that *COL2A1* mutations are responsible for EO-OA (for review of *COL2A1* mutations -- as a predominant feature -- in OA patients, see [23]).

The majority of mutations found in *COL2A1* are missense mutations where glycine is usually replaced by a bulkier amino acid. Among the non-glycine missense mutations, the arginine-to-cysteine substitutions predominate. The α-chain of type II collagen consists of repeat tripeptide sequences, with glycine strictly in place every third amino acid. These tripeptides are essential for collagen triple helix assembly. Thus, mutations resulting in amino acid changes can impair protein stability and thus, helical structure and proper type II collagen function [29]. However, no clear genotype-to-phenotype profiles have been established in type II collagenopathies [24,30].

Experimental mouse models do favor a role of *COL2A1* alterations in OA pathogenesis. *Del1(+/-)* mice harboring a small deletion in *Col2a1* have been found to develop early OA-like lesions in their knee joints [31]. Heterozygous *Sedc*/+ mouse, mutant for *Col2a1*, appears phenotypically normal, yet develop premature OA [32]. Homozygous (sedc/sedc) mouse expresses dwarfism in addition to the OA-like effect on articular cartilage [33]. Based on immunohistochemistry and electron microscopy analyses, the authors conclude that collagen in

the mutant's articular cartilage (both heterozygote and homozygote) fails to provide the normal meshwork required for matrix integrity and overall cartilage stability.

Mutations in minor collagen types have also been identified in patients presenting EO-OA. The three alpha chains of type IX collagen are encoded by the collagen type IX, alpha 1, 2 and 3 genes (*COL9A1*, *COL9A2* and *COL9A3*), respectively. Type IX collagen mutations are responsible for **multiple epiphyseal dysplasia** (MED) [34–36]. Further support for the possible role of type IX collagen in OA has been obtained from animal studies. Mutant mice for *Col9a1* develop a severe degenerative joint disease resembling human OA [20,37]. To date, only one study has screened mutations in the three type IX collagen genes within 72 Finnish non-syndromic EO-OA patients. One unique exonic and five unique intronic variations were found in *COL9A1*, *COL9A2* and *COL9A3*, but none of them altered known splicing consensus sequences or resulted in amino acid substitutions [19]. Recently, a novel missense *COL9A3* mutation were identified in 2 Korean related patients, the proband presented with MED and his father with EO-OA [38]. This supports the hypothesis that *COL9A3* mutations could also be causal for EO-OA.

The type XI collagen fibrils are encoded by the collagen type XI, alpha 1 and 2 genes (*COL11A1* and *COL11A2*), respectively. *COL11A1* has been implicated in OA development in one mouse model. Heterozygosity for a loss-of-function mutation in *Col11a1* resulted in OA development in the knee and temporomandibular joints of *cho/+* mice [39]. This suggests that *COL11A1* mutations in humans could be causal for OA. Nevertheless, no causal *COL11A1* mutations have been identified to date (one Dutch family, 72 Finnish probands and one US family with primary EO-OA) [17,19]. This may be due to a limited number of OA cases studied.

Presently, only one study has identified mutations in the *COL11A2* gene in non-syndromic EO-OA [19]. All other studies have been conducted in patients presenting syndromic OA, **otospondylomegaepiphyseal dysplasia** (OSMED) and other overlapping phenotypes such as **cleft palate**, **Robin sequence**, and **non-ophthalmic STL**. In the former study, *COL11A2* mutations have been identified in two families presenting EO-OA. Despites great efforts to define strict selection criteria for probands and to establish OA diagnosis on radiological data, related patients have unclear clinical definition. In fact, the first mutation is a single nucleotide deletion affecting a consensus splice site and leading to an in-frame deletion of 36 amino-acids of the exon 43 (c.3151-2delA, p.Gly1051_Lys1086del) This splice variant was identified in a small US family with EO-OA, segregating with the OA phenotype in the two affected cases. One of them has experienced knee pain since age 17 after participating in athletic activities, suggesting that OA development is secondary to excessive biomechanical constraints. This splice mutation leads to exon 43 skipping and thus to an in-frame deletion of 36 amino-acids. The authors suggest that this deletion leads to only partially folded molecules of type XI collagen. The second mutation is a missense variant (c.1615C>T, p.Arg539Trp) identified in a Finnish proband with primary bilateral hip EO-OA. This mutation segregates with the OA phenotype in the four EO-OA cases of the family. One member presenting late-onset OA does not share the mutation, suggesting that the mutation is specific to EO-OA, and responsible for the early-onset OA in these patients. However, one mutation carrier has a loose cartilage fragment caused by **osteochondritis dissecans** (OD) (which could explain secondary OA development), and one other mutation carrier is asymptomatic at 45 years of age. This missense mutation leads to the substitution of an arginine to a tryptophan in the major triple-helix and has been predicted to be probably damaging by **Polyphen *in silico* analysis**. The authors assume that this tryptophan substitution (the most hydrophobic amino acid), has an effect on the collagen triple-helix conformation or affects the interactions with other matrix molecules.

Mutations in genes encoding other components of the cartilage extracellular matrix

Proteoglycans represent another major protein class in the ECM. The aggrecan gene (*ACAN*) encodes aggrecan, the most abundant proteoglycan in cartilage. Its main function is to resist compression, protecting bones and joints. Heterozygous mutations in *ACAN* have been identified in two families affected by **spondyloepiphyseal dysplasia type Kimberley** (SEDK) [40] and OD [41] respectively. These two mutations affect the C-type lectin G3 domain of aggrecan, which mediates interactions with other proteins in the ECM [42]. However, a homozygous missense mutation of *ACAN* has been identified in a family presenting a recessive form of spondyloepimetaphyseal dysplasia without manifestation of OA phenotype [43]. This mutation affects also the C-type lectin G3 domain of aggrecan. Thus, disruption of this domain does not inevitably lead to an EO-OA phenotype. Yet, a 7-bp deletion in the *ACAN* gene has been found to cause certain forms of chondrodysplasia in *cmd* mouse [44]. So, mutations in *ACAN* seems not to be directly causal for primary EO-OA but rather to skeletal disorders leading to secondary EO-OA.

Genes that encode other non-collagenous proteins of the ECM have also been screened for mutations in EO-OA patients. The cartilage oligomeric matrix protein gene (*COMP*) encodes a pentameric ECM glycoprotein that catalyzes collagen assembly and promotes the formation of well-defined fibrils [45]. Thus, it may play a role in the structural integrity of cartilage. *COMP* gene mutations were initially identified in familial and sporadic **pseudoachondroplasia** (PSACH) cases [46,47] and in sporadic MED cases [47] and were subsequently validated by numerous studies [48–52]. In vivo, the tetracycline-inducible D469del-COMP mouse model most closely mimics the PSACH pathology [53]. Genotype-to-phenotype correlations of *COMP* gene mutations in these chondrodysplasias were studied for 300 *COMP* mutations [54].

In 2011, a missense *COMP* mutation (c.2152C>T, p.Arg718Trp) was identified in a large extended six-generation Taiwanese kindred with familial EO-OA, negative for *COL2A1* mutations [55]. Twenty-five relatives had radiologically diagnoses OA with a mean age onset of 30 years, and most of the secondary causes of OA were ruled out by physical and laboratory findings. On 26 mutation carriers, 21 had EO-OA and five were asymptomatic with a mean age of 18 years, possibly explaining the absence of OA. Four cases with clinical and radiological EO-OA did not carry the mutation, suggesting a multifactorial inheritance of OA. The mutation has been previously described in MED patients [49,56,57]. Thus, the authors conclude that the mutation causes a mild form MED mimicking the EO-OA phenotype.

Matrilin-3 gene (*MATN3*) mutations have been identified in patients with MED [58–60]. All are missense mutations and mainly affect conserved residues within the beta-sheet of the single A-domain of matrilin-3. These mutations cause protein misfolding, and preventing its secretion from the rough endoplasmic reticulum, both *in vitro* and *in vivo* [61–63]. In mammalian cells overexpressing MATN3 mutant (CHO-B2 cell line or primary bovine articular chondrocytes), mutant protein were retained and accumulated within the cells, in particular within the rough endoplasmic reticulum due to unfolded conformation [61,62]. These observations were also found in a murine model of epiphyseal dysplasia generated by knocking-in a *Matn3* mutation. The phenotype of mice homozygous for the mutation is consistent to human multiple epiphyseal dysplasia. In contrast, a missense variation affecting the epidermal growth factor–like domain of matrilin-3 cosegregated with hand OA in several families [64]. This variation does not appear to interfere with secretion but rather with extracellular assembly of matrix structures [65]. The variation frequency is slightly more than 2% in patients with hand OA in the Icelandic population [64]. It appears to contribute to the susceptibility to adult-onset OA affecting hand joints preferentially, and is not associated with an early-onset or unusual form of OA [64]. The

phenotype of transgenic mouse models are in contradiction. For instance, in one study, *Matn3* null mice appeared normal, were fertile, and showed no obvious skeletal malformations [66]. This suggested functional redundancy among matrilins, indicating that MED disorder phenotypes might not be caused by the absence of matrilin-3 in ECM [66]. However, in a second study, aged *Matn3* null mice were much more predisposed to develop severe OA than their wild-type littermates. Also, the lack of *Matn3* did not lead to postnatal chondrodysplasia but yet accounted for higher incidence of late-onset OA [67].

Mutations in the *SMAD3* gene: involvement in the TGF-β pathway

Members of the **transforming growth factor-beta** (TGF-β) superfamily represent a major signaling pathway in joints and are involved in numerous cellular processes in cartilage [2]. Mice harboring mutations in members of the TGF-β signaling pathway display phenotypes similar to human osteoarthritis. For instance, mice overexpressing a dominant-negative form of Tgfbr2 [68] and Smad3$^{-/-}$ mice [69] have shown severe progressive osteoarthritis-like disease. To date, only the *SMAD3* gene has been screened for mutations in human OA patients. This gene encodes a transcriptional modulator activated by transforming growth factor-beta (TGF-β) that plays a role in the anabolic pathway in cartilage [70].

Based on the phenotype observed in an animal model, *SMAD3* exons were screened in 32 patients with sporadic knee OA [71]. It is important to note that this cohort was not evaluated for other anomalies. A missense mutation (c.590A>T, p.Asn197Ile) was identified in one patient, with higher expression of pro-MMP-2 and pro-MMP-9 in serum relative to controls and to non-carrier OA patients. This suggests that this mutation may lead to upregulation of MMPs and eventually to the development of OA. However, (i) this mutation is located in the linker region of SMAD3 which is not well conserved, (ii) only one sporadic OA case carried a *SMAD3* mutation (no replication cohorts were screened to identify a *SMAD3* mutation in

additional cases), and (iii) neither the prematurity nor the age of onset of OA were specified. Subsequently, *SMAD3* mutations were identified in patients with **aneurysms-osteoarthritis syndrome** (AOS)[72]. The evaluation of the clinical phenotype of 45 patients from eight different AOS families with eight different *SMAD3* mutations showed that radiologically validated OA in 96% of the 26 examined patients, with 75% having two or more affected joints [73]. Early-onset joint abnormalities, including OA, intervertebral disc degeneration, OD and meniscal anomalies were present in almost all patients with AOS, in contrast to other forms of thoracic aortic aneurysms and dissections in which these anomalies are rarely described. This establishes early-onset joint abnormalities as a key feature of this new syndrome. The causal role of heterozygous loss of function mutations in the *SMAD3* gene were further validated by independent studies in human [74,75]. Recently, the clinical spectrum of *SMAD3* mutations were completed on 50 mutation carriers [74]. Osteoarticular manifestations were recorded in all patients. Joint involvement could be severe, requiring surgery in young patients (at a mean age of 33), with unusual localization such as tarsus, shoulder and carpus, or mimicking chondrocalcinosis. In contrast, other cohorts did not report OA, joint pain or deformities in some *SMAD3* mutation carriers [76,77]. However, either the individuals did not undergo imaging, with the diagnosis of OA and joint pain based on medical records/interviews to assess joint pain and potentially leading to the underestimation of OA cases [76], or, mutation carriers were young, which could explain the absence of radiographic signs of OA [77]. In addition, as mutations were mostly described in the context of cardiac disease, the apparent absence of OA could be explained by incomplete clinical investigation or description (X-rays are not systematic). Recently, the vascular phenotypes of *Smad3* knockout mice were investigated, showing that this mouse model also replicates clinical vascular aspects of AOS, marked by the progressive development of aneurysms [78,79].

Mutations in genes encoding proteins involved in other functions and pathways

The trafficking protein particle complex 2 gene (*TRAPPC2*) encodes the sedlin protein that plays a role in vesicular transport from endoplasmic reticulum to Golgi. *TRAPPC2* mutations are causal for **spondyloepiphyseal dysplasia tarda** (SEDT)[80]. Distinctive radiological signs include mild to moderate epiphyseal dysplasia which usually leads to severe premature OA of large joints frequently necessitating hip joint replacement in the third or fourth decade [81]. As a majority of mutations identified in *TRAPPC2* lead to protein truncation, and based on the speculation that less severe missense mutations may have different phenotypic effects such as EO-OA alone, a cohort of 37 male patients with early end-stage primary hip or knee OA were screened for *TRAPPC2* mutations [82]. Secondary OA risk factors were excluded. No *TRAPPC2* mutations were identified in the entire coding region indicating that *TRAPPC2* mutations are not a common cause of early primary OA in men. However, neither non-coding regions nor introns of *TRAPPC2* were screened, and other known causal genes in EO-OA were not investigated.

In the last few years, whole-exome sequencing has been successfully applied to identify the molecular basis of numerous Mendelian diseases. In 2015, Ramos *et al.* used this approach for the first time in patients with EO-OA and particularly in two distant family members with dominantly inherited primary EO-OA at multiple joint sites with radiographic **chondrocalcinosis** [5]. The age of onset of OA in the family varies between 30 years and 50 years, and OA arises in the absence of mild or severe chondrodysplasia. Thanks to this high-throughput technique, the causal mutation for the familial EO-OA phenotype was identified in the *TNFRSF11B* gene (tumor necrosis factor receptor superfamily, member 11b gene) that encodes the osteoprotegerin (OPG) protein, a member of the tumor necrosis factor receptor superfamily. This protein plays an important role in bone remodeling. The power of this method

is highlighted by the fact that candidate gene screening and genetic linkage analysis previously failed to reveal the molecular basis of OA development in this large family. The heterozygous stop loss mutation (c.1205A>T, p.Ter402Leu) cosegregates with the phenotype of OA in the family and is absent in 744 controls. Functional analyses showed that this mutation establishes a gain of function of the encoding protein OPG. The authors hypothesize that the mutation acts *via* an unfavorable interplay between subchondral bone and cartilage towards enhanced matrix mineralization, which is a major hallmark of OA disease process [83].

Concluding Remarks

Primary EO-OA is a genetically heterogeneous disease. Several mutations in genes encoding components of the ECM or involved in distinct biological pathways (TGF-β, vesicular transport or bone remodeling) have been identified in idiopathic EO-OA patients or in patients having syndromic forms of EO-OA. Mutations in these genes have only been reported in a few studies in human due to the rarity of EO-OA cases. In addition, there is little genetic overlap between early- and late-onset OA, with only *COL11A1* and *SMAD3* genes that harbor mutation or polymorphism for both types of OA [84,85]. This suggests that EO-OA and common late-onset OA are two very different types of OA from a molecular genetics perspective.

The genetic heterogeneity observed in primary EO-OA can be explained in part by various clinical features associated with EO-OA, mainly observed in syndromic diseases. In the latter, OA is not systematically the key feature studied, leading to a lack of phenotypic information. In addition, the range of clinical and radiological definitions of OA used in genetic studies results in difficulties when interpreting the causal role of mutations in OA determinism, thus complicating a clear genotype-to-phenotype description.

Further studies are required to clarify the molecular basis of EO-OA (Outstanding Questions). First, only *COL2A1* mutations are clearly involved in EO-OA determinism. The causal role of other genes needs to be confirmed by independent studies in idiopathic EO-OA patients. Second, assessment of the prevalence of gene mutations in idiopathic EO-OA is needed to determine their respective share in EO-OA development. Third, establishment of a genotype-to-phenotype correlation in large cohorts of both idiopathic and syndromic forms of EO-OA may give clues as to which classes of mutations or relevant protein domains are specific to a particular phenotype. For this purpose, the use of an international recognized definition for EO-OA is important as well as a complete description of the patients' phenotypes. Of note, imaging is absolutely necessary in establishing OA diagnosis, as the disease can be asymptomatic, particularly in young individuals.

To highlight a new molecular basis for EO-OA, future studies should focus on patients with severe phenotypes, as these are the most informative cases. Clear molecular basis, prevalence data, and genotype-to-phenotype correlations are essential to improve EO-OA diagnosis, management and therapy. For this end, the ExoRhum project (NCT01999166) aims to identify new gene mutations involved in EO-OA pathogenesis, and to estimate the prevalence of mutations in known causal genes (https://clinicaltrials.gov/ct2/show/NCT01999166). This genetic study will be performed on patients with idiopathic EO-OA, based on a strict definition of OA and excluding known causes of secondary OA. It will rely on whole exome sequencing, an approach that has been successfully applied to identify the molecular basis of numerous Mendelian diseases, and which has recently enabled to reveal a gain of function mutation in the *TNFRSF11B* gene in two distant family members with idiopathic EO-OA [5]. The use of exome sequencing should identify candidate genes harboring nucleotidic variants or copy number variations that have never been studied in the context of EO-OA.

In conclusion, to-date studies have identified twelve genes potentially causal for EO-OA, that could be targeted for OA treatment. Thus, a gain of function mutation was recently identified in *TNFRSF11B*, thanks to a strict selection of cases and controls and to whole-exome sequencing. Enhanced function of osteoprotegerin, encoded by *TNFRSF11B*, could be the direct underlying cause of OA and hence, a potential therapeutic target for OA.

Acknowledgements

JAL is a recipient of a fellowship from Région Basse-Normandie. Research by the authors is supported by French Minister of Health (PHRC ExoRhum), Agence Nationale de la Recherche (ANR), Région Basse-Normandie and Société Française de Rhumatologie (SFR).

References

1 Lawrence, R.C. *et al.* (2008) Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum.* 58, 26–35

2 Baugé, C. *et al.* (2014) Regulation and Role of TGFβ Signaling Pathway in Aging and Osteoarthritis Joints. *Aging Dis.* 5, 394–405

3 Sandell, L.J. (2012) Etiology of osteoarthritis: genetics and synovial joint development. *Nat. Rev. Rheumatol.* 8, 77–89

4 Min, J.L. *et al.* (2007) Mutation analysis of candidate genes within the 2q33.3 linkage area for familial early-onset generalised osteoarthritis. *Eur. J. Hum. Genet. EJHG* 15, 791–799

5 Ramos, Y.F.M. *et al.* (2015) A gain of function mutation in TNFRSF11B encoding osteoprotegerin causes osteoarthritis with chondrocalcinosis. *Ann. Rheum. Dis.* 74, 1756–1762

6 Stecher, R. (1941) , Heberden's nodes: heredity in hypertrophic arthritis of the finger joints. , *Am J Med Sci*

7 Kellgren, J.H. *et al.* (1963) Genetic Factors in Generalized Osteo-Arthrosis. *Ann. Rheum. Dis.* 22, 237–255

8 Spector, T.D. *et al.* (1996) Genetic influences on osteoarthritis in women: a twin study. *BMJ* 312, 940–943

9 Tsezou, A. (2014) Osteoarthritis year in review 2014: genetics and genomics. *Osteoarthr. Cartil. OARS Osteoarthr. Res. Soc.* 22, 2017–2024

10 Reynard, L.N. and Loughlin, J. (2013) Insights from human genetic studies into the pathways involved in osteoarthritis. *Nat. Rev. Rheumatol.* 9, 573–583

11 Reynard, L.N. and Loughlin, J. (2013) The genetics and functional analysis of primary osteoarthritis susceptibility. *Expert Rev. Mol. Med.* 15, e2

12 Chapman, K. and Valdes, A.M. (2012) Genetic factors in OA pathogenesis. *Bone* 51, 258–264

13 Loughlin, J. (2015) Genetic contribution to osteoarthritis development: current state of evidence. *Curr. Opin. Rheumatol.* 27, 284–288

14 Panoutsopoulou, K. and Zeggini, E. (2013) Advances in osteoarthritis genetics. *J. Med. Genet.* 50, 715–724

15 Ramos, Y.F.M. and Meulenbelt, I. (2015) Implementation of Functional Genomics for Bench-to-Bedside Transition in Osteoarthritis. *Curr. Rheumatol. Rep.* 17, 53

16 Palotie, A. *et al.* (1989) Predisposition to familial osteoarthrosis linked to type II collagen gene. *Lancet Lond. Engl.* 1, 924–927

17 Meulenbelt, I. *et al.* (1997) Genetic linkage analysis of 14 candidate gene loci in a family with autosomal dominant osteoarthritis without dysplasia. *J. Med. Genet.* 34, 1024–1027

18 Ingvarsson, T. *et al.* (2001) A large Icelandic family with early osteoarthritis of the hip associated with a susceptibility locus on chromosome 16p. *Arthritis Rheum.* 44, 2548–2555

19 Jakkula, E. *et al.* (2005) The role of sequence variations within the genes encoding collagen II, IX and XI in non-syndromic, early-onset osteoarthritis. *Osteoarthritis Cartilage* 13, 497–507

20 Nakata, K. *et al.* (1993) Osteoarthritis associated with mild chondrodysplasia in transgenic mice expressing alpha 1(IX) collagen chains with a central deletion. *Proc. Natl. Acad. Sci. U. S. A.* 90, 2870–2874

21 Fang, H. and Beier, F. (2014) Mouse models of osteoarthritis: modelling risk factors and assessing outcomes. *Nat. Rev. Rheumatol.* 10, 413–421

22 Staal, B. *et al.* (2014) Cartilage-specific deletion of Mig-6 results in osteoarthritis-like disorder with excessive articular chondrocyte proliferation. *Proc. Natl. Acad. Sci. U. S. A.* 111, 2590–2595

23 Husar-Memmer, E. *et al.* (2013) Premature osteoarthritis as presenting sign of type II collagenopathy: a case report and literature review. *Semin. Arthritis Rheum.* 42, 355–360

24 Nishimura, G. *et al.* (2005) The phenotypic spectrum of COL2A1 mutations. *Hum. Mutat.* 26, 36–43

25 Knowlton, R.G. *et al.* (1990) Genetic Linkage of a Polymorphism in the Type II Procollagen Gene (COL2A1) to Primary Osteoarthritis Associated with Mild Chondrodysplasia. *N. Engl. J. Med.* 322, 526–530

26 Ala-Kokko, L. *et al.* (1990) Single base mutation in the type II procollagen gene (COL2A1) as a cause of primary osteoarthritis associated with a mild chondrodysplasia. *Proc. Natl. Acad. Sci. U. S. A.* 87, 6565–6568

27 Ritvaniemi, P. *et al.* (1995) Identification of COL2A1 gene mutations in patients with chondrodysplasias and familial osteoarthritis. *Arthritis Rheum.* 38, 999–1004

28 Lankisch, P. *et al.* (2014) COL2A1 mutation as a cause of premature osteoarthritis in a 13-year-old child. *Joint Bone Spine* 81, 83–85

29 Horton, W.A. *et al.* (1992) Characterization of a type II collagen gene (COL2A1) mutation identified in cultured chondrocytes from human hypochondrogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 89, 4583–4587

30 Kannu, P. *et al.* (2012) Clinical phenotypes associated with type II collagen mutations. *J. Paediatr. Child Health* 48, E38–43

31 Säämänen, A.-M.K. *et al.* (2000) Osteoarthritis-like lesions in transgenic mice harboring a small deletion mutation in type II collagen gene. *Osteoarthritis Cartilage* 8, 248–257

32 Holt, D.W. *et al.* (2012) Osteoarthritis-like changes in the heterozygous sedc mouse associated with the HtrA1-Ddr2-Mmp-13 degradative pathway: a new model of osteoarthritis. *Osteoarthr. Cartil. OARS Osteoarthr. Res. Soc.* 20, 430–439

33 Macdonald, D.W. *et al.* (2013) Structural variations in articular cartilage matrix are associated with early-onset osteoarthritis in the spondyloepiphyseal dysplasia congenita (sedc) mouse. *Int. J. Mol. Sci.* 14, 16515–16531

34 Muragaki, Y. *et al.* (1996) A mutation in the gene encoding the alpha 2 chain of the fibril-associated collagen IX, COL9A2, causes multiple epiphyseal dysplasia (EDM2). *Nat. Genet.* 12, 103–105

35 Paassilta, P. *et al.* (1999) COL9A3: A third locus for multiple epiphyseal dysplasia. *Am. J. Hum. Genet.* 64, 1036–1044

36 Czarny-Ratajczak, M. *et al.* (2001) A mutation in COL9A1 causes multiple epiphyseal dysplasia: further evidence for locus heterogeneity. *Am. J. Hum. Genet.* 69, 969–980

37 Fässler, R. *et al.* (1994) Mice lacking alpha 1 (IX) collagen develop noninflammatory degenerative joint disease. *Proc. Natl. Acad. Sci. U. S. A.* 91, 5070–5074

38 Jeong, C. *et al.* (2014) Novel COL9A3 mutation in a family diagnosed with multiple epiphyseal dysplasia: a case report. *BMC Musculoskelet. Disord.* 15, 371

39 Xu, L. *et al.* (2003) Osteoarthritis-like changes and decreased mechanical function of articular cartilage in the joints of mice with the chondrodysplasia gene (cho). *Arthritis Rheum.* 48, 2509–2518

40 Gleghorn, L. *et al.* (2005) A Mutation in the Variable Repeat Region of the Aggrecan Gene (AGC1) Causes a Form of Spondyloepiphyseal Dysplasia Associated with Severe, Premature Osteoarthritis. *Am. J. Hum. Genet.* 77, 484–490

41 Stattin, E.-L. *et al.* (2010) A missense mutation in the aggrecan C-type lectin domain disrupts extracellular matrix interactions and causes dominant familial osteochondritis dissecans. *Am. J. Hum. Genet.* 86, 126–137

42 Aspberg, A. (2012) The different roles of aggrecan interaction domains. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* 60, 987–996

43 Tompson, S.W. *et al.* (2009) A recessive skeletal dysplasia, SEMD aggrecan type, results from a missense mutation affecting the C-type lectin domain of aggrecan. *Am. J. Hum. Genet.* 84, 72–79

44 Watanabe, H. *et al.* (1994) Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in the aggrecan gene. *Nat. Genet.* 7, 154–157

45 Acharya, C. *et al.* (2014) Cartilage oligomeric matrix protein and its binding partners in the cartilage extracellular matrix: Interaction, regulation and role in chondrogenesis. *Matrix Biol.* 37, 102–111

46 Hecht, J.T. *et al.* (1995) Mutations in exon 17B of cartilage oligomeric matrix protein (COMP) cause pseudoachondroplasia. *Nat. Genet.* 10, 325–329

47 Briggs, M.D. *et al.* (1995) Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. *Nat. Genet.* 10, 330–336

48 Jackson, G.C. *et al.* (2012) Pseudoachondroplasia and Multiple Epiphyseal Dysplasia: A 7-Year Comprehensive Analysis of the Known Disease Genes Identify Novel and Recurrent Mutations and Provides an Accurate Assessment of Their Relative Contribution. *Hum. Mutat.* 33, 144–157

49 Kennedy, J. *et al.* (2005) COMP mutation screening as an aid for the clinical diagnosis and counselling of patients with a suspected diagnosis of pseudoachondroplasia or multiple epiphyseal dysplasia. *Eur. J. Hum. Genet. EJHG* 13, 547–555

50 Briggs, M.D. and Chapman, K.L. (2002) Pseudoachondroplasia and multiple epiphyseal dysplasia: mutation review, molecular interactions, and genotype to phenotype correlations. *Hum. Mutat.* 19, 465–478

51 Unger, S. and Hecht, J.T. (2001) Pseudoachondroplasia and multiple epiphyseal dysplasia: New etiologic developments. *Am. J. Med. Genet.* 106, 244–250

52 Briggs, M.D. *et al.* (1998) Diverse mutations in the gene for cartilage oligomeric matrix protein in the pseudoachondroplasia-multiple epiphyseal dysplasia disease spectrum. *Am. J. Hum. Genet.* 62, 311–319

53 Posey, K.L. *et al.* (2014) Pseudoachondroplasia/COMP — translating from the bench to the bedside. *Matrix Biol. J. Int. Soc. Matrix Biol.* 0, 167–173

524 54 Briggs, M.D. *et al.* (2014) Genotype to phenotype correlations in cartilage oligomeric
525 matrix protein associated chondrodysplasias. *Eur. J. Hum. Genet.* 22, 1278–1282
526 55 Mu, S.-C. *et al.* (2009) A large kindred of early-onset osteoarthritis of the knee and
527 hip: excluding the link to COL2A1 gene. *Rheumatology* 48, 371–374
528 56 Jakkula, E. *et al.* (2003) A recurrent R718W mutation in COMP results in multiple
529 epiphyseal dysplasia with mild myopathy: clinical and pathogenetic overlap with collagen IX
530 mutations. *J. Med. Genet.* 40, 942–948
531 57 Mabuchi, A. *et al.* (2003) Novel types of COMP mutations and genotype-phenotype
532 association in pseudoachondroplasia and multiple epiphyseal dysplasia. *Hum. Genet.* 112, 84–
533 90
534 58 Chapman, K.L. *et al.* (2001) Mutations in the region encoding the von Willebrand
535 factor A domain of matrilin-3 are associated with multiple epiphyseal dysplasia. *Nat. Genet.*
536 28, 393–396
537 59 Mabuchi, A. *et al.* (2004) Novel and recurrent mutations clustered in the von
538 Willebrand factor A domain of MATN3 in multiple epiphyseal dysplasia. *Hum. Mutat.* 24,
539 439–440
540 60 Mäkitie, O. *et al.* (2004) Clinical and radiographic findings in multiple epiphyseal
541 dysplasia caused by MATN3 mutations: description of 12 patients. *Am. J. Med. Genet. A.*
542 125A, 278–284
543 61 Cotterill, S.L. *et al.* (2005) Multiple epiphyseal dysplasia mutations in MATN3 cause
544 misfolding of the A-domain and prevent secretion of mutant matrilin-3. *Hum. Mutat.* 26, 557–
545 565
546 62 Otten, C. *et al.* (2005) Matrilin-3 mutations that cause chondrodysplasias interfere
547 with protein trafficking while a mutation associated with hand osteoarthritis does not. *J. Med.*
548 *Genet.* 42, 774–779

549 63 Leighton, M.P. *et al.* (2007) Decreased chondrocyte proliferation and dysregulated
550 apoptosis in the cartilage growth plate are key features of a murine model of epiphyseal
551 dysplasia caused by a matn3 mutation. *Hum. Mol. Genet.* 16, 1728–1741
552 64 Stefánsson, S.E. *et al.* (2003) Genomewide Scan for Hand Osteoarthritis: A Novel
553 Mutation in Matrilin-3. *Am. J. Hum. Genet.* 72, 1448–1459
554 65 Otten, C. *et al.* (2010) A matrilin-3 mutation associated with osteoarthritis does not
555 affect collagen affinity but promotes the formation of wider cartilage collagen fibrils. *Hum.*
556 *Mutat.* 31, 254–263
557 66 Ko, Y. *et al.* (2004) Matrilin-3 is dispensable for mouse skeletal growth and
558 development. *Mol. Cell. Biol.* 24, 1691–1699
559 67 van der Weyden, L. *et al.* (2006) Functional knockout of the matrilin-3 gene causes
560 premature chondrocyte maturation to hypertrophy and increases bone mineral density and
561 osteoarthritis. *Am. J. Pathol.* 169, 515–527
562 68 Serra, R. *et al.* (1997) Expression of a Truncated, Kinase-Defective TGF-β Type II
563 Receptor in Mouse Skeletal Tissue Promotes Terminal Chondrocyte Differentiation and
564 Osteoarthritis. *J. Cell Biol.* 139, 541–552
565 69 Yang, X. *et al.* (2001) TGF-beta/Smad3 signals repress chondrocyte hypertrophic
566 differentiation and are required for maintaining articular cartilage. *J. Cell Biol.* 153, 35–46
567 70 Zhu, Y. *et al.* (2015) Transforming growth factor‐β1 induces type II collagen and
568 aggrecan expression via activation of extracellular signal‐regulated kinase 1/2 and Smad2/3
569 signaling pathways. *Mol. Med. Rep.* 12, 5573–5579
570 71 Yao, J.-Y. *et al.* (2003) Mutation analysis of the Smad3 gene in human osteoarthritis.
571 *Eur. J. Hum. Genet.* 11, 714–717
572 72 van de Laar, I.M.B.H. *et al.* (2011) Mutations in SMAD3 cause a syndromic form of
573 aortic aneurysms and dissections with early-onset osteoarthritis. *Nat. Genet.* 43, 121–126

73 van de Laar, I.M.B.H. *et al.* (2012) Phenotypic spectrum of the SMAD3-related aneurysms–osteoarthritis syndrome. *J. Med. Genet.* 49, 47–57

74 Aubart, M. *et al.* (2014) Early-Onset Osteoarthritis, Charcot-Marie-Tooth Like Neuropathy, Autoimmune Features, Multiple Arterial Aneurysms and Dissections: An Unrecognized and Life Threatening Condition. *PLoS ONE* 9,

75 Zhang, W. *et al.* (2015) A Novel Mutation of SMAD3 Identified in a Chinese Family with Aneurysms-Osteoarthritis Syndrome. *BioMed Res. Int.* 2015,

76 Regalado, E.S. *et al.* (2011) Exome Sequencing Identifies SMAD3 Mutations as a Cause of Familial Thoracic Aortic Aneurysm and Dissection with Intracranial and Other Arterial Aneurysms. *Circ. Res.* 109, 680–686

77 Wischmeijer, A. *et al.* (2013) Thoracic Aortic Aneurysm in Infancy in Aneurysms–Osteoarthritis Syndrome Due to a Novel SMAD3 Mutation: Further Delineation of the Phenotype. *Am. J. Med. Genet. A.* 161, 1028–1035

78 Ye, P. *et al.* (2013) GM-CSF contributes to aortic aneurysms resulting from SMAD3 deficiency. *J. Clin. Invest.* 123, 2317–2331

79 Tan, C.K. *et al.* (2013) SMAD3 Deficiency Promotes Inflammatory Aortic Aneurysms in Angiotensin II–Infused Mice Via Activation of iNOS. *J. Am. Heart Assoc. Cardiovasc. Cerebrovasc. Dis.* 2,

80 Gedeon, A.K. *et al.* (1999) Identification of the gene (SEDL) causing X-linked spondyloepiphyseal dysplasia tarda. *Nat. Genet.* 22, 400–404

81 Bannerman, R.M. *et al.* (1971) X-linked spondyloepiphyseal dysplasia tarda: clinical and linkage data. *J. Med. Genet.* 8, 291–301

82 Fiedler, J. *et al.* (2002) Mutations in the X-linked spondyloepiphyseal dysplasia tarda (SEDL) coding sequence are not a common cause of early primary osteoarthritis in men. *Clin. Genet.* 62, 94–95

83 Goldring, M.B. and Goldring, S.R. (2010) Articular cartilage and subchondral bone in the pathogenesis of osteoarthritis. *Ann. N. Y. Acad. Sci.* 1192, 230–237

84 Rodriguez-Fontenla, C. *et al.* (2014) Assessment of osteoarthritis candidate genes in a meta-analysis of nine genome-wide association studies. *Arthritis Rheumatol. Hoboken NJ* 66, 940–949

85 Valdes, A.M. *et al.* (2010) Genetic variation in the SMAD3 gene is associated with hip and knee osteoarthritis. *Arthritis Rheum.* 62, 2347–2352

86 Moon, S. *et al.* (2015) A genome-wide association study of copy-number variation identifies putative loci associated with osteoarthritis in Koreans. *BMC Musculoskelet. Disord.* 16, 76

87 Panoutsopoulou, K. *et al.* (2011) Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study. *Ann. Rheum. Dis.* 70, 864–867

88 Mu, S.-C. *et al.* (2011) A mutation in cartilage oligomeric matrix protein (COMP) causes early-onset osteoarthritis in a large kindred study. *Ann. Hum. Genet.* 75, 575–583

89 Myllyharju, J. (2014) Extracellular Matrix and Developing Growth Plate. *Curr. Osteoporos. Rep.* 12, 439–445

Figure legends

Figure 1: **Radiographic phenotype of osteoarthritis in humans**

(A) Normal knee joint. (B) Osteoarthritis of the knee located in the medial femoro-tibial compartment: joint space narrowing, tibial and femoral **osteophytosis**, **subchondral bone geode** of the medial tibial plateau.

Figure 2, Key Figure: **Proteins involved in early-onset osteoarthritis pathogenesis**

(A) *Articular cartilage: composition and metabolism*. Chondrocytes are surrounded by ECM which confers stiffness and elasticity to cartilage. The TGF-β pathway has been implicated in promoting cartilage synthesis. Briefly, TGF-β binds to TGFBR2 leading to the recruitment of TGFBR1, which subsequently phosphorylates SMAD2 or SMAD3. Activated SMAD2 or SMAD3 complexes with SMAD4, translocating into the nucleus. The Smad complex binds to the promoter regions of TGF-β target genes, interactiing with other transcription factors and co-factors, and inducing the transcription of target genes encoding EMC components. Components of ECM are secreted by chondrocytes by vesicle-mediated transport. Sedlin is a part of the TRAPP complex, required for the endoplasmic reticulum export of procollagen. ECM consists of collagens (type II, IX and XI), proteoglycans (aggrecan, biglycan, decorin, fibromodulin), other non-collagenous proteins (link protein, fibronectin, COMP, matrilin-3, integrin) and hyaluronan.

(B) *Bone resorption*. OPG is a soluble decoy receptor which inhibits **osteoclastogenesis** by competing with RANK (expressed on the membrane of pre-**osteoclasts**) in binding its ligand, RANKL. Gain of function mutations of OPG lead to a more efficient antagonism of osteoclastogenesis.

Proteins encoded by genes involved in EO-OA pathogenesis are shown in red.

Abbreviations: co-F, cofactors; COMP, cartilage oligomeric matrix protein; ECM, extracellular matrix; OPG, osteoprotegerin; RANK, receptor activator of the nuclear factor-κB; RANKL, RANK ligand; TF, transcription factor; TGF-β, transforming growth factor-β; TGFBR1/2, TGF- β receptor type-1/2; TRAPP, trafficking protein particle.

Glossary

Aneurysms-osteoarthritis syndrome (AOS): syndromic form of autosomal dominant thoracic aortic aneurysms and dissections characterised by the presence of arterial aneurysms and tortuosity, mild craniofacial, skeletal and cutaneous anomalies and EO-OA.

Animal model: laboratory animals used to study disease mechanisms. Results can be extrapolated to human disorders.

Chondrocalcinosis: deposition of calcium salts in the cartilage of joints.

Chondrocyte: cell found in cartilage, that produces extra-cellular matrix components.

Chondrodysplasia: bone disease characterized by abnormal growth of the epiphyses.

Cleft palate: opening in the roof of the mouth.

Early-onset osteoarthritis (EO-OA): osteoarthritis that occurs before 50 years old, rare and mainly familial.

Epiphyse: end of a long bone.

Heberden's nodes: bony enlargement of the terminal joint of a finger.

Late-onset osteoarthritis: osteoarthritis that occurs after 50 years old, frequent.

Multiple epiphyseal dysplasia (MED): relatively mild and clinically variable osteochondrodysplasia, characterized by delayed and irregular ossification of the epiphyses and EO-OA.

Non-ophthalmic Stickler syndrome: clinically variable disorder characterized by auditory, skeletal and orofacial abnormalities.

Osteoarthritis (OA): degenerative joint disease, mainly characterized by the gradual loss of articular cartilage, particularly affecting the knee, hip, spine, hand and foot.

Osteochondritis dissecans (OD): skeletal disorder defined as a separation of cartilage and subchondral bone from the surrounding tissue which primarily affects the knee, ankle and elbow joints.

Osteochondrodysplasia: abnormal growth or development of cartilage and bone.

Osteoclast: cell involved in bone resorption.

Osteoclastogenesis: development of osteoclasts.

Osteophytosis: presence of one or more osteophytes wich are bony growths that develop on bone extremities.

Otospondylomegaepiphyseal dysplasia (OSMED): characterized by sensorineural hearing loss, enlarged epiphyses, disproportionate shortness of the limbs, abnormalities in vertebral bodies and typical facial features.

Platyspondyly: radiographic feature refering to flattened vertebral bodies throughout the axial skeleton.

Polyphen *in silico* analysis: tool which predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations.

Pseudoachondroplasia (PSACH): osteochondrodysplasia characterized by disproportionate short stature, deformity of the lower limbs, brachydactyly, loose joints and ligamentous laxity.

Robin sequence: set of abnormalities affecting the head and face (micrognathia, glossoptosis and cleft palate).

Spondyloepiphyseal dysplasia: disorder of bone growth that results in short stature, skeletal abnormalities and problems with vision and hearing.

Spondyloepiphyseal dysplasia, Kimberley type (SEDK): characterized by short stature and premature degenerative arthropathy.

Spondyloepiphyseal dysplasia tarda (SEDT): rare X-linked recessive osteochondrodysplasia, characterized by disproportionate short stature.

Stickler syndrome (STL): clinically variable disorder characterized by ocular, auditory, skeletal and orofacial abnormalities.

Subchondral bone geode: lytic lesion in the periarticular surfaces.

Syndrome: group of signs and symptoms that occur together and characterize a particular abnormality.

Transforming growth factor-beta (TGF-β) pathway: major signaling pathway in joints, involved in numerous cellular processes in cartilage.

Table 1: Pathologies associated with secondary early-onset osteoarthritis development

Pathology	**Phenotype MIM number**	**Clinical features**
OA secondary to excessive biomechanical constraints		
Post-traumatic OA		
Osteochondritis dissecans (OD)	#165800	skeletal disorder defined as a separation of cartilage and subchondral bone from the surrounding tissue which primarily affects the knee, ankle, and elbow joints. Familial OD is characterized by multiple osteochondritic lesions in knees and/or hips and/or elbows,

		disproportionate short stature, and early-onset osteoarthritis.
Pseudoachondroplasia (PSACH)	#177170	osteochondrodysplasia characterized by disproportionate short stature, deformity of the lower limbs, brachydactyly, loose joints, and ligamentous laxity. Vertebral anomalies, present in childhood, usually resolve with age, but osteoarthritis is progressive and severe. Exhibits a similar, but more severe phenotype than MED.
Spondyloepiphyseal dysplasia, Kimberley type (SEDK)	#608361	skeletal disorder with proportionate short stature and early-onset progressive osteoarthropathy
Spondyloepimetaphyseal dysplasia (SEMD)	#612813	skeletal dysplasia with a predominant epiphyseal component leading to premature osteoarthritis of the weight-bearing large joints
OA secondary to weakening of the cartilage, induced by metabolic or inflammatory pathologies		

Alkaptonuria (AKU)	#203500	metabolic disorder characterized by accumulation of homogentisic acid, leading to darkened urine, pigmentation of connective tissue (ochronosis), joint and spine arthritis, and destruction of the cardiac valves
Ankylosing spondylitis (AS)	#106300	common inflammatory rheumatic disease that affects the axial skeleton. AS is the major subtype and a main outcome of an inter-related group of rheumatic diseases now named spondyloarthropathies .
Chondrocalcinosis (CCAL)	#600668	cartilage calcification which is a common condition that usually results from deposition of crystals of calcium pyrophosphate dihydrate (CPPD) in articular hyaline and fibro-cartilage
Rheumatoid arthritis (RA)	#180300	inflammatory disease, primarily of the joints, with autoimmune features
OA secondary to joint deformities caused by collagen defect		

Avascular necrosis of the femoral head (ANFH) in adults	#608805	debilitating disease that usually leads to destruction of the hip joint in the third to fifth decade of life
Legg-Calve-Perthes disease (LCPD) in children	#150600	form of ANFH in growing children
Aneurysms-osteoarthritis syndrome (AOS) also known as Loeys-Dietz syndrome type III	#613795	aortic aneurysm syndrome with widespread systemic involvement, which is associated with early-onset osteoarthritis
Czech dysplasia	#609162	characterized by normal height, early-onset osteoarthritis, **platyspondyly**, short metatarsals, but with absence of ophthalmological complications or cleft palate
Kniest dysplasia	#156550	disorder of bone growth characterized by short stature (dwarfism) with other skeletal abnormalities and problems with vision and hearing.
Multiple epiphyseal dysplasia (MED)	#132400 #600204 #600969 #607078	skeletal disorder characterized by short stature and early-onset osteoarthritis.

	#614135	Similar, but milder phenotype than PSACH
Otospondylomegaepiphyseal dysplasia (OSMED)	#215150	characterized by sensorineural hearing loss, enlarged epiphyses, disproportionate shortness of the limbs, abnormalities in vertebral bodies, and typical facial features
Spondyloepiphyseal dysplasia tarda (SEDT)	#313400	skeletal disorder associated with early-onset osteoarthritis
Spondyloperipheral dysplasia	#271700	disorder that impairs bone growth
Stickler syndrome (STL)	#108300 #609508 #604841 #184840 #614134 #614284	clinically variable disorder characterized by ocular, auditory, skeletal, and orofacial abnormalities. Additional findings may include mild spondyloepiphyseal dysplasia, and early-onset osteoarthritis.

704

705 **Table 2: Mutated genes identified in idiopathic early-onset osteoarthritis**

Gene symbol	Approved name	Locus	Mutation[a]	Phenotype	Reference
Genes encoding collagen proteins					
COL2A1	collagen, type II, alpha 1	12q12 -q13.2	NM_001844.4:c.823C >T (p.Arg275Cys)	- Familial EO-OA - AO: 12-40y - Bilateral joint sites simultaneously: hip and/or knee	[28]
			NM_001844.4:c.2155 C>T (p.Arg719Cys)	- Familial generalized EO-OA - AO: 2nd-3rd decade - Multiple joint sites simultaneously: hip, knee, shoulder, wrist, hand	[26]
			NM_001844.4:c.2659 G>A (p.=)[b]	- Familial EO-OA - AO: 30-50y - Multiple joint sites	[19]

				simultaneously: hip, knee, hand	
COL11A2	collagen, type XI, alpha 2	6p21.3	NM_080680.2:c.1615 C>T (p.Arg539Trp)	- Familial EO-OA - AO: 38y (proband), <50y (family members) - Multiple joint sites simultaneously: hip, knee, hand, spine	[19]
			NM_080680.2:c.3151-2delA (p.?)	- Familial EO-OA - AO: 17-20y - Multiple joint sites simultaneously: knee, hip, spine	[19]
Gene encoding other component of the cartilage extracellular matrix					
COMP	cartilage oligomeric matrix protein	19p13.1	NM_000095.2:c.2152 C>T (p.Arg718Trp)	- Familial EO-OA - AO: 29.7 ± 16.1y - Classical OA joint sites involved	[88]

Gene encoding protein involved in bone remodeling					
TNFRSF11B	tumor necrosis factor receptor superfamily, member 11b	8q24	NM_002546.3:c.1205A>T (p.Ter402Leu)	- Familial generalized EO-OA with chondrocalcinosis - AO: 30-50 y - Multiple joint sites simultaneously: hand, knee, hip, spine	[5]

[a]Mutation at cDNA and protein level according to HGVS nomenclature.

[b]Splice mutation associated with decreased mRNA level in carrier.

AO, age of onset; EO, early-onset; OA, osteoarthritis

Table 3: Causal genes identified in early-onset osteoarthritis forms associated with other clinical features

Gene symbol	Approved name	Locus	Associated disease presenting EO-OA[a]
Genes encoding collagen proteins			
COL2A1	collagen, type II, alpha 1	12q12-q13.2	- Familial EO-OA - Avascular necrosis of femoral head - Legg-Calvé-Perthes disease - Czech dysplasia - Kniest dysplasia - Otospondylomegaepiphyseal dysplasia - Spondyloperipheral dysplasia - Stickler syndrome, type I
COL9A1	collagen, type IX, alpha 1	6q13	- Multiple epiphyseal dysplasia, type VI - Stickler syndrome, type IV
COL9A2	collagen, type IX, alpha 2	1p33-p32	- Multiple epiphyseal dysplasia, type II - Stickler syndrome, type V
COL9A3	collagen, type IX, alpha 3	20q13.3	- Multiple epiphyseal dysplasia, type III - Stickler syndrome
COL11A1	collagen, type XI, alpha 1	1p21	- Stickler syndrome, type II
COL11A2	collagen, type XI, alpha 2	6p21.3	- Familial EO-OA - Ostopondylomegaepiphyseal dysplasia - Stickler syndrome, type III
Genes encoding other components of the cartilage extracellular matrix			

ACAN	aggrecan	15q26.1	- Osteochondritis dissecans - Spondyloepimetaphyseal dysplasia - Spondyloepiphyseal dysplasia, Kimberley type
COMP	cartilage oligomeric matrix protein	19p13.1	- Familial EO-OA - Multiple epiphyseal dysplasia, type I - Pseudoachondroplasia
MATN3	matrilin 3	2p24-p23	- Multiple epiphyseal dysplasia, type V
Gene encoding protein involved in the TGF-β pathway			
SMAD3	SMAD family member 3	15q21-q22	- Aneurysms-osteoarthritis syndrome
Gene encoding protein involved in vesicular transport			
TRAPPC2	trafficking protein particle complex 2	Xp22	- Spondyloepiphyseal dysplasia tarda
Gene encoding protein involved in bone remodeling			
TNFRSF11B	tumor necrosis factor receptor superfamily, member 11b	8q24	- Familial EO-OA

712

Box 1: Description of the collagen types

Collagens are structural components of the ECM secreted by **chondrocytes** [89]. They bring traction resistance to cartilage. The major collagen protein is the fibril-forming type II collagen, other collagen types being the type XI and IX collagens. Type IX collagen fibrils are typically found within the core of type II collagen fibrils which help maintain the spacing and diameter of type II collagen fibrils, while type IX collagen resides at the surface of type II/XI fibrils. So mutations in genes encoding collagens lead to loss of resistance to traction, absorption and distribution of forces in cartilage which affect the structure of the cartilage.

Box 2: Strict inclusion criteria for primary EO-OA genetic studies

1. No known causes of secondary OA development (e.g. negative trauma history)
2. Body mass index (BMI) ≤ 30
3. Radiologically proven OA (Kellgren and Lawrence grade ≥ 2)
4. Age at onset ≤ 40 years and ≥ 1 joint site involved

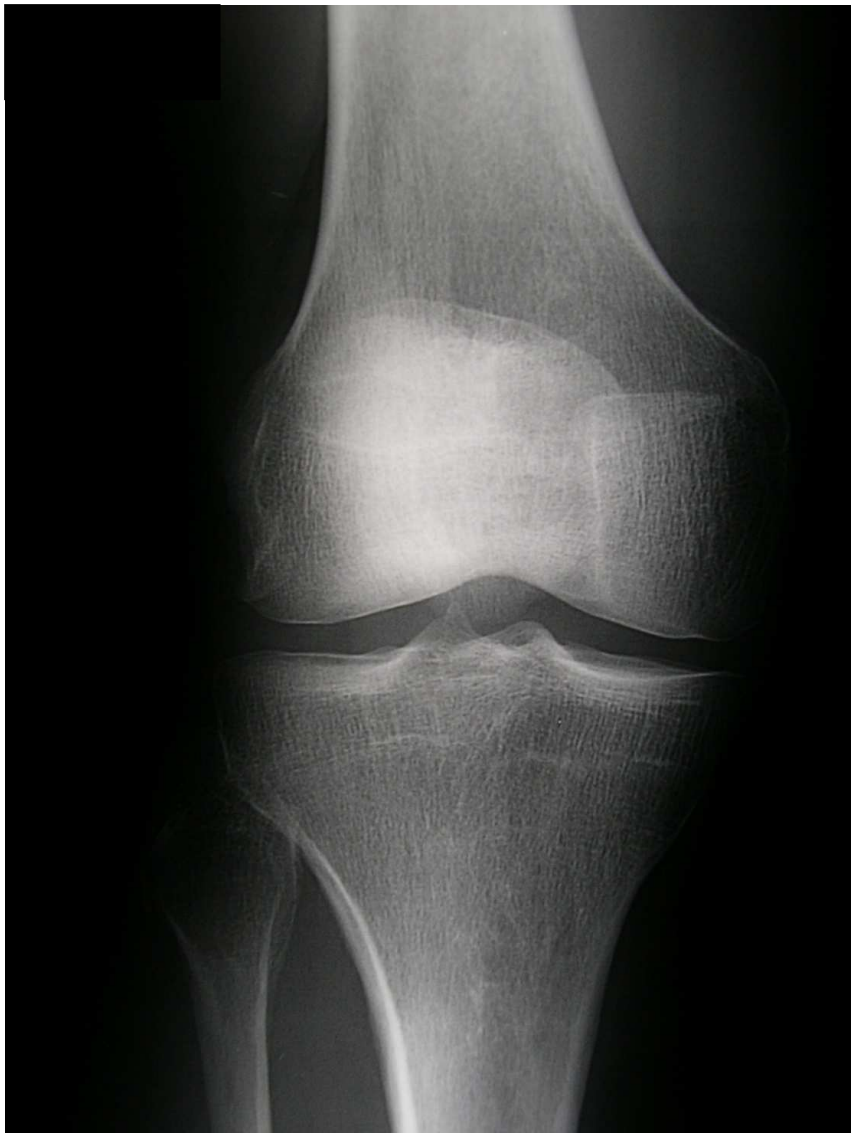
   or Age at onset ≤ 50 years **and** multiple joints sites involved (≥ 2 joints)

   or Age at onset ≤ 50 years **and** 1 joint site involved **and** positive family history

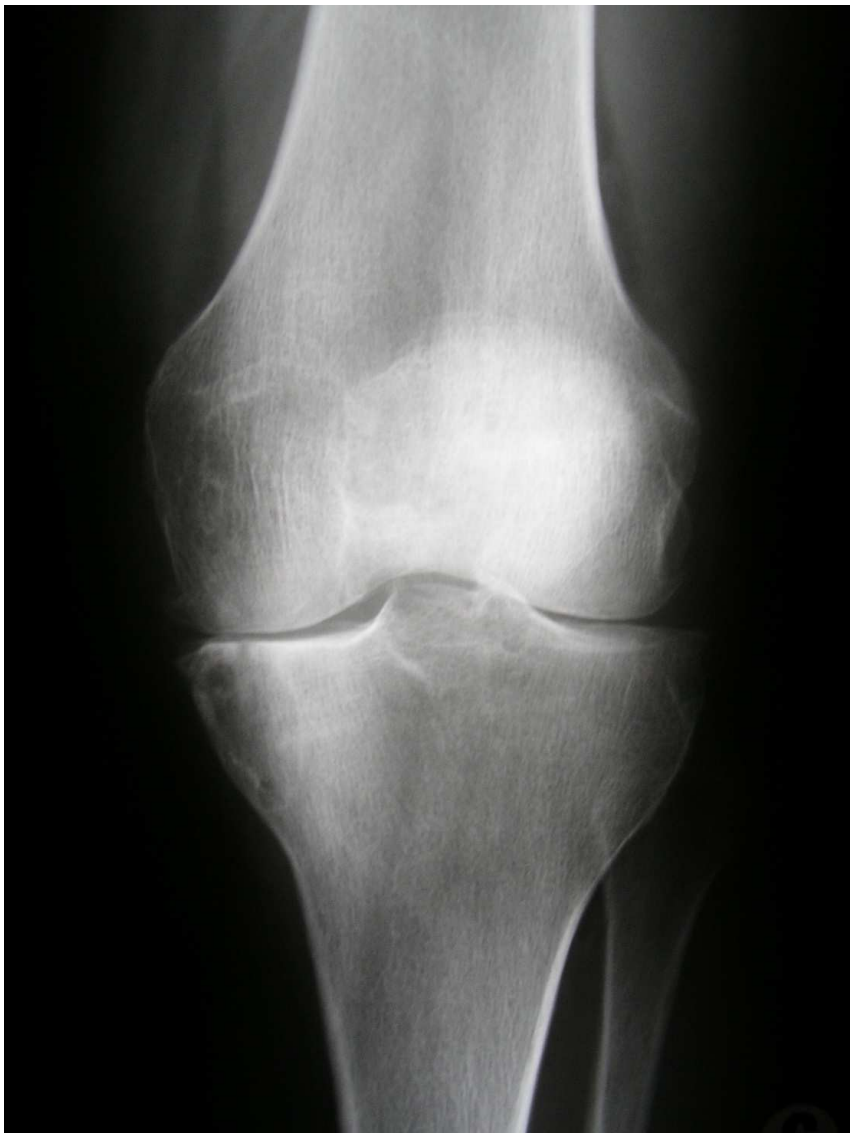
   or Age at onset ≤ 50 years **and** 1 unusual joint site involved (shoulder, elbow, wrist, ankle)

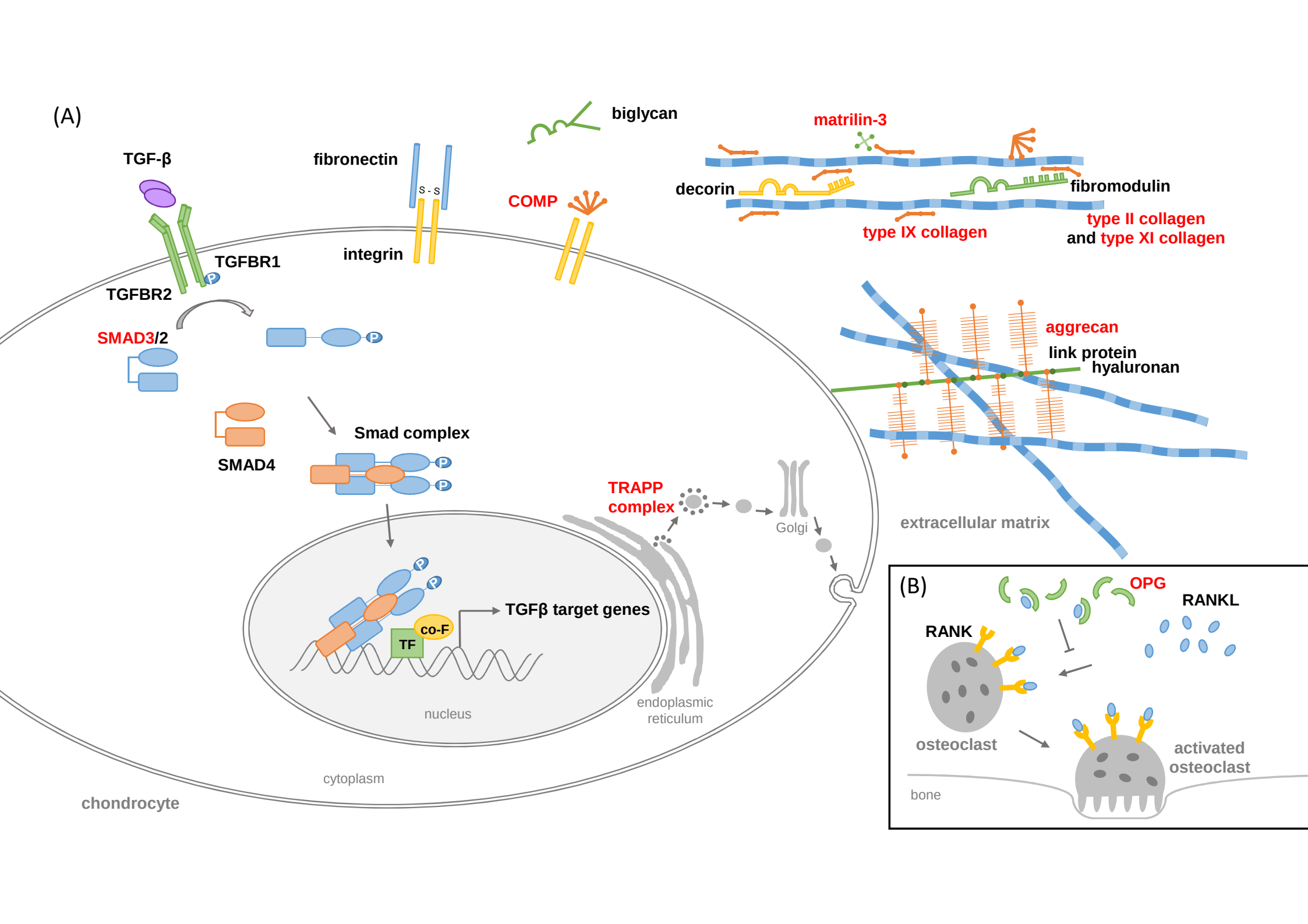
Criteria adapted from [16]

(A)



(B)




(A)
TGF-β
TGFBR1
TGFBR2
SMAD3/2
SMAD4
Smad complex
fibronectin
S - S
integrin
biglycan
COMP
matrilin-3
decorin
fibromodulin
type IX collagen
type II collagen
and type XI collagen
aggrecan
link protein
hyaluronan
extracellular matrix
TRAPP
complex
Golgi
TGFβ target genes
co-F
TF
nucleus
endoplasmic
reticulum
cytoplasm
chondrocyte
(B)
OPG
RANKL
RANK
osteoclast
activated
osteoclast
bone

Outstanding questions box

- What is the prevalence of mutations in known causal genes for EO-OA cases? Determining the respective share of each causal gene in EO-OA development will be essential to improve EO-OA diagnosis, in particular to prioritize gene mutations screening.
- Are there genotype-to-phenotype correlations in patients with EO-OA? These correlations will improve EO-OA diagnosis as well as management.
- Are others genes and pathways involved in EO-OA pathogenesis than those yet established? Identification of new genes or pathways could open new therapeutic perspectives as highlighted with the recent description of a gain of function mutation in *TNFRSF11B* gene encoding osteoprotegerin which leads the authors to advocate that agents counteracting the function of osteoprotegerin could be a promizing therapy for OA [5].
    - What is the contribution of copy number variations in the genetic determinism to EO-OA? Rare copy number variations are involved in the pathogenesis of various diseases such as neurological diseases or cancers. Recently, a genome-wide association study identified putative loci associated with osteoarthritis in Koreans [86]. However, their contribution to OA EO-OA determinism has never been studied and will lead to a better understanding of the pathogenesis of OA.
- Are some idiopathic EO-OA cases oligogenic? Genome-wide association scan data suggests that common osteoarthritis is a highly polygenic disease with multiple risk variants conferring small effects[87]. So we can hypothesis that some idiopathic EO-OA cases harbor few genes mutations and that their combination is causal for the observed phenotype.

Trends box

- By contrast to late-onset OA, idiopathic EO-OA are rare Mendelian forms, mainly including other clinical features such as ocular, auditory or aortic abnormalities.
- Genetic heterogeneity is observed in EO-OA patients. To date, only *COL2A1* mutations are clearly causal for EO-OA. Eleven additional genes need to be further validated, including components of the ECM or proteins involved in the TGF-β pathway, vesicular transport or bone remodeling.
- The rise of high-throughput sequencing technologies give new opportunities, at a reasonable cost, to identify genes or pathways that could be new therapeutic target for OA. Thus, in 2015, a gain of function mutation was identified in *TNFRSF11B* which segregates with a familial form of EO-OA. Enhanced function of osteoprotegerin, encoded by *TNFRSF11B*, could be the direct underlying cause of OA and hence, a potential therapeutic target for OA.