Implementation of Functional Genomics for Bench-to-Bedside Transition in Osteoarthritis

Yolande F. M. Ramos & Ingrid Meulenbelt

Abstract Collaborative work has led to the identification of robust osteoarthritis susceptibility genes. The translation of these gene deviations towards underlying biological mechanisms remains challenging yet necessary to allow the transition of these results in drug development and to applications in the clinic. Besides the availability of disease-relevant tissues which will contribute to a deeper insight into ongoing pathophysiological processes as result of the gene deviations, availability of respective cells will be key to establish in vitro models to further characterize genetic variants at a transcriptional and functional level. To allow exploration of effective use of drugs and biologicals such findings should eventually be further tested in in vivo models. In this review, we discuss possibilities for the implementation of an effective functional genomic pipeline for bench-to-bedside translation of identified susceptibility loci.

Keywords Functional genomics · Susceptibility loci · GWAS · Cell models · Osteoarthritis

Introduction

Since the number of published genetic studies that lack experimental data showing the biological relevance of identified genetic variation is still large, we advocate that enhanced implementation of functional genomics is needed to substantially augment translation to drug development and disease management. Clear insight into the etiology of OA would greatly contribute to early diagnosis, as well as to the development of urgently needed disease-modifying treatments. Consequently, major efforts should be made to unravel the functional role of genetic factors in predisposition to OA.

At present, based on genome-wide significance in GWAS \( p \leq 5 \times 10^{-8} \) and/or proven functional involvement in OA by follow-up studies, 21 independent OA susceptibility loci were established from candidate gene studies, linkage studies, and GWAS (Table 1). This has resulted in a variety of compelling OA candidate genes. However, GWAS also identified loci within gene deserts or within regions with multiple genes that have not been implicated previously in OA etiology [1, 2]. Multiple single-nucleotide polymorphisms (SNPs) in a region can be in strong linkage disequilibrium (LD) with each other, making it often challenging to determine the causal variant. This illustrates the typical problem in mapping complex traits like OA and exposes the existing gap between genetic evidence and the molecular mechanism underlying OA. In order to overcome these problems, multiple molecular determinants (genomic, transcriptomic, epigenomic, and proteomic) will have to be integrated on a more regular basis [3, 4, 5].

In this review, we discuss the current state of genetic research in OA. Although the number of published genetic studies lacking experimental data that functionally characterize the biological relevance of identified genetic variation is still large, in the field of OA research, implementation of the step from genomics to biological function, also named ‘functional
### Table 1  Established OA susceptibility loci

<table>
<thead>
<tr>
<th>SNP</th>
<th>CHR</th>
<th>POS</th>
<th>Annotated gene</th>
<th>Gene symbol</th>
<th>EA</th>
<th>EAF</th>
<th>OR (95 % CI)</th>
<th>Joint Sex</th>
<th>Sex</th>
<th>Ref</th>
<th>FF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Candidate gene study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12901499</td>
<td>15</td>
<td>67370695</td>
<td>Mothers against decapentaplegic homolog 3</td>
<td>SMAD3</td>
<td>G</td>
<td>0.57</td>
<td>1.22 (1.12–1.34)</td>
<td>7.5E-6</td>
<td></td>
<td></td>
<td>[55]</td>
</tr>
<tr>
<td>rs143383</td>
<td>20</td>
<td>34025983</td>
<td>Growth differentiation factor 5</td>
<td>GDF5</td>
<td>T</td>
<td>0.74</td>
<td>1.79 (1.53–2.09)</td>
<td>2.0E-13</td>
<td></td>
<td></td>
<td>[51], [9•, 23, 26, 57]</td>
</tr>
<tr>
<td>rs7775</td>
<td>2</td>
<td>183699584</td>
<td>Frizzled motif associated with bone development</td>
<td>FRZB</td>
<td>G</td>
<td>0.11</td>
<td>4.1 (1.6–10.7)</td>
<td>4.0E-03</td>
<td></td>
<td></td>
<td>[58], [38, 39, 59]</td>
</tr>
<tr>
<td>rs225014</td>
<td>14</td>
<td>80669580</td>
<td>Deiodenase iodothyronine type 2</td>
<td>DIO2</td>
<td>C</td>
<td>0.34</td>
<td>1.79 (1.37–2.34)</td>
<td>2.0E-05</td>
<td></td>
<td></td>
<td>[60], [6•, 13, 59]</td>
</tr>
<tr>
<td>rs2615977</td>
<td>1</td>
<td>103452642</td>
<td>Collagen-type 11A1</td>
<td>COL11A1</td>
<td>C</td>
<td>0.22</td>
<td>1.10 (1.05–1.16)</td>
<td>3.2E-05</td>
<td></td>
<td></td>
<td>[61], [62] [63]</td>
</tr>
<tr>
<td>rs7639618</td>
<td>3</td>
<td>15216429</td>
<td>Double Von Willebrand Factor A</td>
<td>DVWA</td>
<td>G</td>
<td>0.63</td>
<td>1.43 (1.28–1.59)</td>
<td>7.3E-11</td>
<td></td>
<td></td>
<td>[64]</td>
</tr>
<tr>
<td>rs6976</td>
<td>3</td>
<td>52728804</td>
<td>3p21.1 locus: e.g., glycosyltransferase 8 domain containing 1 and guanine nucleotide binding protein-like 3</td>
<td>GLT8D1; GNL3</td>
<td>T</td>
<td>0.37</td>
<td>1.12 (1.08–1.16)</td>
<td>7.2E-11</td>
<td></td>
<td></td>
<td>[65]</td>
</tr>
<tr>
<td>rs12107036</td>
<td>3</td>
<td>189600160</td>
<td>Tumor protein p63</td>
<td>TP63</td>
<td>G</td>
<td>0.52</td>
<td>1.21 (1.13–1.29)</td>
<td>6.7E-08</td>
<td></td>
<td></td>
<td>[2•]</td>
</tr>
<tr>
<td>rs10947262</td>
<td>6</td>
<td>32373312</td>
<td>Butyrophilin-like 2 and major histocompatibility complex, class II, DQ beta 1</td>
<td>BTN2L; HLA-DQB1</td>
<td>C</td>
<td>0.58</td>
<td>1.31 (1.20–1.44)</td>
<td>5.1E-09</td>
<td></td>
<td></td>
<td>[66]</td>
</tr>
<tr>
<td>rs10948172</td>
<td>6</td>
<td>44777691</td>
<td>Suppressor of Tr 3 homolog and CD5 molecule-like</td>
<td>SUPT3H; CDC5L</td>
<td>G</td>
<td>0.29</td>
<td>1.14 (1.09–1.20)</td>
<td>7.9E-08</td>
<td></td>
<td></td>
<td>[2•]</td>
</tr>
<tr>
<td>rs9350591</td>
<td>6</td>
<td>76241527</td>
<td>Filamin A interacting protein 1 and SUMO1/sentrin-specific peptidase 6</td>
<td>MILIP1; SENP6</td>
<td>T</td>
<td>0.11</td>
<td>1.18 (1.12–1.25)</td>
<td>2.4E-09</td>
<td></td>
<td></td>
<td>[2•]</td>
</tr>
<tr>
<td>rs3815148</td>
<td>7</td>
<td>106934820</td>
<td>7q22 locus: e.g., component of oligomeric golgi complex 5 and HMG-box transcription factor 1</td>
<td>COG5; HBP1</td>
<td>C</td>
<td>0.23</td>
<td>1.14 (1.09–1.19)</td>
<td>8.0E-08</td>
<td></td>
<td></td>
<td>[1], [67•]</td>
</tr>
<tr>
<td>rs4730250</td>
<td>7</td>
<td>107207695</td>
<td>7q22 locus: e.g., dihydrouridine synthase 4-like</td>
<td>DUS4L</td>
<td>G</td>
<td>0.17</td>
<td>1.17 (1.11–1.24)</td>
<td>9.2E-09</td>
<td></td>
<td></td>
<td>[46]</td>
</tr>
<tr>
<td>rs4836732</td>
<td>9</td>
<td>119266695</td>
<td>Astrotactin 2 and pregnancy-associated plasma-protein A</td>
<td>ASTN2; PAPPA</td>
<td>C</td>
<td>0.47</td>
<td>1.2 (1.13–1.27)</td>
<td>6.1E-10</td>
<td></td>
<td></td>
<td>[2•]</td>
</tr>
<tr>
<td>rs10492367</td>
<td>12</td>
<td>28014970</td>
<td>Kelch domain containing 5 and parathyroid hormone-like hormone</td>
<td>KLHDC5; FTHILH</td>
<td>T</td>
<td>0.19</td>
<td>1.14 (1.09–1.20)</td>
<td>1.5E-08</td>
<td></td>
<td></td>
<td>[2•]</td>
</tr>
<tr>
<td>rs835487</td>
<td>12</td>
<td>105060767</td>
<td>Carbohydrate (chondroitin 4) sulfotransferase 11</td>
<td>CHST11</td>
<td>G</td>
<td>0.34</td>
<td>1.13 (1.09–1.18)</td>
<td>1.6E-08</td>
<td></td>
<td></td>
<td>[2•]</td>
</tr>
<tr>
<td>rs11842874</td>
<td>13</td>
<td>113694509</td>
<td>MCF2 cell line derived transforming sequence-like</td>
<td>MCF2L</td>
<td>A</td>
<td>0.93</td>
<td>1.17 (1.11–1.23)</td>
<td>2.1E-08</td>
<td></td>
<td></td>
<td>[68]</td>
</tr>
<tr>
<td>rs12907038</td>
<td>15</td>
<td>58244032</td>
<td>Aldehyde dehydrogenase 1 family, member A2</td>
<td>ALDH1A2</td>
<td>G</td>
<td>0.507</td>
<td>1.51</td>
<td>4.0E-10</td>
<td></td>
<td></td>
<td>[7•], [7•]</td>
</tr>
<tr>
<td>rs8044769</td>
<td>16</td>
<td>53839135</td>
<td>Fat mass and obesity associated</td>
<td>FTO</td>
<td>C</td>
<td>0.5</td>
<td>1.11 (1.07–1.15)</td>
<td>6.9E-08</td>
<td></td>
<td></td>
<td>[2•]</td>
</tr>
</tbody>
</table>
genomics’, has only recently started (Table 1, column FF: functional follow-up; [6••, 7••, 8••, 9••, 10]). To proceed and improve the approaches taken, we propose the implementation of a genomic road map (Fig. 1), proceeding from suggestions from Freedman et al. [11], Sunyaev [12], and from several OA studies.

The Importance of Biobanking

To achieve effective implementation of functional genomics in complex diseases such as OA, the assembly of a biobank is required with relevant tissues stored in a manner tailored to isolate RNA, DNA, and protein and suitable for a range of downstream analyses, including the generation of gene expression profiles and analysis of differential allelic expression, genotyping and eQTL analysis, determination of the epigenetic landscape, and proteomics and immunohistochemical analysis. The additional collection of relevant cell types (mesenchymal stem cells, chondrocytes, bone cells, etc.) would facilitate the setup of experimental in vitro cell models. Together, this will allow the analyses at multiple molecular levels, as was also done for the OA susceptibility gene deiodinase, iodothyronine, type II (DIO2) using tissues and cells collected within the Research Arthritis and Articular Cartilage (RAAK) study [6••, 13]. Human genome-wide gene expression datasets are available for OA cartilage from non-OA and OA-affected joints [14, 15] as well as from unaffected and affected cartilage areas of the same joints [16–18], OA-affected and OA-unaffected subchondral bone [19••], and OA-affected synovial membrane [20•]. However, besides the combined transcriptome and methylome analyses of unaffected and affected cartilage areas of the same joints [21••], most of these datasets were obtained from tissues collected in different studies. Ideally, they should be derived from the same study group and in combination with a variety of phenotypic data, including clinical questionnaires, X-ray, and MRI, which could provide detailed information on the dynamics of molecular changes during pathophysiological processes. Inclusion of blood and/or urine would also permit the identification of biomarkers for diagnosis in easily accessible tissues, while the availability of additional follow-up data would allow identification of prognostic biomarkers or biomarkers for monitoring ongoing disease.

Transcription Regulation

Since many of the identified signals are localized outside the transcriptome, a generally accepted concept for conferring OA susceptibility by genetic variation is related to altered gene expression in disease-relevant tissues. Genetic variants can interfere with transcription factor binding resulting in altered
gene expression (eQTL). In some cases, this interference is due to changes in epigenetic control mechanisms (e.g., mQTL and eQTM) [4••, 11]. The latter was nicely shown for the OA susceptibility locus of \textit{DIO2} [6••] and for \textit{GDF5} in which DNA methylation at the susceptibility locus rs143383 (Table 1) results in altered binding of transcription factors leading to differential allelic expression (DAE) of the gene [9••, 22]. Therefore, a first approach, which is especially interesting when applied to intronic or intergenic SNPs, is to investigate whether a genetic variant, localized in the coding
region and in high LD with the identified SNP, is associated with DAE as has been applied to GDF5 [23], DIO2 [13], and the susceptibility locus of ALDH1A2 [7••]. However, since chromosome folding can greatly affect transcriptional regulation [24], the gene that is transcriptionally affected by a locus identified in GWAS can be located at distances that are megabases away. In this respect, Chromosome Conformation Capture (3C), Circularized Chromosome Conformation Capture (4C), or Carbon-Copy Chromosome Conformation Capture (5C) can be applied to explore chromosome folding and to identify specific regions of the genome interacting with the site of interest. In fact, the use of 4C sequencing has been of valuable help in the identification of Iroquois-Class Homeodomain Protein 3 (IRX3) as the gene genetically targeted by the obesity-associated locus within the fat mass and obesity-associated gene (FTO) [25••]. As such, it might very well be that the SNP identified in association with OA (rs8044769; Table 1), which is in LD with the SNP associating with obesity (rs9930506; LD=0.67) and which is localized within the first intron of FTO, also results in susceptibility to OA by targeting IRX3.

**In Vitro Cell Models in OA**

In vitro cell models can be broadly applied and are especially suitable to investigate transcriptional regulation of a gene of interest and the consequences of genetic variation [6••, 9••, 26], which may interfere eventually with its regulation in a clinical setting. Different in vitro models can be employed to investigate the effects of strong candidate genes. In these studies, the choice of cells and the method of analysis are of utmost importance.

The use of established cell lines, which are easy to obtain and maintain, is the most straightforward and enables performance of studies within a relatively short period of time. For example, in OA research, the murine ATDC5 (chondrogenic) or MC3T3 (osteogenic) cells permit experiments of around 3 weeks in duration. These cell lines were used, for example, to functionally characterize the effects of altered IGFBP3 expression during in vitro chondrogenesis and osteogenesis, respectively [10]. In addition, findings were validated by over-expressing IGFBP3 in primary chondrocytes isolated from patients who had undergone knee replacement surgery. We particularly advocate the use of primary human cells on a regular basis, since established cell lines may have stable modifications that not only permit constant proliferative activity, but also modify basal gene expression. Furthermore, in cell lines the normal association between cell growth and differentiation may be altered and they have frequently accumulated mutations during the course of the extended proliferation. Thus, the altered cell responses may not necessarily reflect the responses to gene variations occurring in humans in vivo. This is not a problem in most studies investigating specific molecular pathways; however, caution must be exerted in interpreting experimental results in studies of cell differentiation and matrix deposition. It may be possible to circumvent the challenges of primary cells such as their phenotypic instability and limited lifespan by using induced pluripotent stem cells (iPS), which provide a stable source of cells. Moreover, iPS cells can also be differentiated towards multiple lineages, including chondrogenic and osteogenic progenitors [27]. Several characteristics of the different types of cells are summarized in Table 2.

Previous studies with human articular chondrocytes showed that monolayer cultures (2D) mainly result in a hypertrophic phenotype of the chondrocytes [28•], thereby hampering the investigation of the effects of variable expression of the gene of interest on matrix deposition. Similar results were obtained in our laboratory using ATDC5 cells in monolayer, while attempts to form 3D pellets of this cell line failed (YR and IM; unpublished data). Therefore, despite the substantial discrepancies observed between monolayer and 3D cultures, we strongly encourage exploiting 3D pellet cultures derived from human bone marrow-derived mesenchymal stem cells (BMSCs) or primary human chondrocytes to study in vitro chondrogenesis, rather than 2D cultures of immortalized cell lines. In addition to increased mechanistic insight in studying candidate gene and protein perturbations, in vitro cell models can be used to select drugs or biologicals with potential disease-modifying effects prior to in vivo studies. For example, a similar approach was applied to the DIO2 gene [6••]. Since DIO2 expression is increased in OA, first the effect of increased thyroid signaling on chondrogenesis and deposition of cartilage extracellular matrix was investigated in 3D pellet cultures of BMSCs. When lopanoic acid (IOP), a pharmacological inhibitor of deiodinases, was added to the cultures this was found to be beneficial: addition of IOP attenuated the upregulation of catabolic enzymes and stabilized the COL2A1/COL1A1 ratio, together associated with a more dense cartilage matrix structure and prolonged integrity of cartilage and its homeostasis.

Since the different tissues actively interact with each other during development and disease, cocultures may be even more suitable in vitro models. It is possible to distinguish indirect and direct cocultures to investigate effects on cell

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Comparison of different types of cell lines</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Immortalized cells</td>
</tr>
<tr>
<td>Lifespan</td>
<td>Unlimited</td>
</tr>
<tr>
<td>Immortalized</td>
<td>Yes</td>
</tr>
<tr>
<td>Phenotypically stable</td>
<td>Yes</td>
</tr>
<tr>
<td>3D pellet model doable</td>
<td>No</td>
</tr>
<tr>
<td>Experiment duration</td>
<td>3 weeks</td>
</tr>
</tbody>
</table>
function and response, extracellular matrix deposition, and cell survival. Indirect cocultures include (1) the transfer of conditioned medium from one type of cell culture to the other and (2) the separation of cell types in transwell chambers that allow the diffusion of secreted factors from one culture to the other. Direct cocultures can be performed by the simultaneous culturing of different combinations of mesenchymal stem cells, chondrocytes, osteoblasts, explants, or synoviocytes in the same dish. Frequently applied in studies of tissue regeneration and engineering of replacements suitable for implantation is the combination of mesenchymal stem cells and chondrocytes [29–32]. For example, He et al. [33] established cultures of osteoblasts, BMSCs, and fibroblasts on a hybrid silk scaffold to investigate whether the BMSCs cocultured in the presence of ligament and bone cells would differentiate into fibrocartilage. Leyh et al. [34] showed that the mechanical and biochemical properties of extracellular matrix deposited by BMSCs and chondrocytes is impaired in the presence of OA cartilage explants due to the altered fibrillar network in the host tissue. In the Tuan’s lab, recently a 3D ‘micro-tissue’ was developed suitable for study of OA pathogenesis [35]. This micro-tissue can also be used to evaluate the effects of different reagents and compounds on osteochondral health and the osteochondral interface [36•]. Taken together, these multilineage 3D in vitro models are preferred above single cell type models, before application in animal models in vivo [37].

In Vivo Animal Models in OA

As stated by Sunyaev et al. [12], genetic manipulation of animals such as the mouse, rat, rabbit, horse, and zebrafish provide a possibility to test phenotypic rather than molecular consequences of human genetic variants, considered by many as the ultimate proof of the causal effect of a susceptibility locus. Of the different types of laboratory animals, mice are used most frequently because they are relatively inexpensive and easy to handle and to breed. With respect to cartilage destruction, multiple different mouse strains have been generated for genes known to be involved in chondrogenesis or endochondral ossification. However, such studies are lacking for the known OA susceptibility genes despite their importance in the etiology underlying OA in humans. An exception is the generation of Frzb knockout mice and their use in established OA models [38, 39]. Gdf5 knockout mice were generated long before the identification of the GDF5 susceptibility locus and were used subsequently to study the development of OA in different models (reviewed by Cornelis et al. 2011 [40]). Also, Dio2 knockout mice were generated before the discovery of the association of the DIO2 susceptibility locus with OA [41]. Transgenic rats overexpressing DIO2 in articular cartilage are more prone to cartilage damage [42•], thereby confirming that the detected overexpression of DIO2 in OA cartilage [13] is not only a consequence of OA pathophysiology, but also probably underlies disease development. This is in line with the beneficial effects of Dio2 knockout in mice upon mechanical loading of the joints observed by molecular and microscopic approaches [43•]. Some mouse strains such as C57/BL6 mice and the STR/Ort spontaneously develop OA upon aging due to unknown cause [44], for that matter resembling human primary OA. However, to age the mice and wait until they spontaneously develop OA is an expensive and time-consuming process. Therefore, in most studies OA is induced using surgery (de-stabilization of the medial meniscus or DMM; cruciate ligament transection), enzymes or chemical substances (intra-articular collagenase injections), or mechanical loading. The recently generated cartilage-specific double knockout mice Nfatc1/Nfatc2−/− are extremely prone to develop OA shortly after birth [45]. According to the authors, this is highly favorable in OA research, since it reduces the time–cost ratio of the model and does not require further surgical or chemical intervention. The fact that these mice exhibit early onset, aggressive OA affecting multiple joints, however, suggests that it is a very specific mouse model, more resembling cases of early-onset OA during development, such as chondrodysplasias, in contrast to the commonly observed, slowly progressing primary OA at older ages in humans. When using these mice for specific studies in drug development, for example, this should be taken into account. We recommend increased awareness of the susceptibility loci identified in genetic studies and encourage generation of mice with these variants for functional characterization of the effects of natural variation similar to that occurring in humans. This will facilitate the subsequent development of drugs or identification of biomarkers for early detection of disease. Nevertheless, an important application of animal models is to investigate new treatment options and to test compounds selected because of their positive effects in vitro assays. The use of drugs with encouraging outcome in animal models, i.e., inhibited or reversed disease severity, may lead subsequently to their application in clinical trials.

Conclusions and Suggestions for Future Studies

Given the fact that scientists have been able to collect large numbers of samples by collaborations within large consortia, compelling susceptibility loci have been identified for complex diseases such as OA. The most recent large-scale analyses for OA were done in populations stratified according to joint involvement and resulted in additional, joint-specific, OA susceptibility loci [46, 47•]. The next step, which should now be considered seriously, is the functional characterization of the identified genetic variants. This is important for identifying druggable targets and clinical applications. Moreover, the ability to classify different types of OA and obtain insight
into their specific etiologies will significantly contribute to the optimization of personalized medicine.

We invite geneticists to collaborate on a more regular base with molecular and cell biologists to develop strategies for the implementation of a functional genomic pipeline such as that proposed in Fig. 1, in order to more efficiently translate the finding to drug development and the identification of biomarkers for complex diseases and to invest in the generation of transgenic mice carrying gene variants identified in genetic studies.

Acknowledgments We want to thank our colleagues for inspiring discussions and we are grateful to Dr. L. ‘t Hart, N. Bomer, and W. den Hollander for critically reading the manuscript and useful suggestions.

Compliance with Ethics Guidelines

Conflict of Interest Yolande F. M. Ramos and Ingrid Meulenbelt declare no conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:

• Of importance
• Of major importance


4. Westra HJ, Franke L. From genome to function by studying eQTLs. Biochim Biophys Acta. 2014. Excellent review clearly explaining the importance of eQTL analysis as a means to come from susceptibility loci to causal genes.


36. Lin H, Lozito TP, Alexander PG, Gottardi R, Tuan RS. Stem cell-based microphysiological osteochondral system to model tissue response to interleukin-1beta. Mol Pharm. 2014;11:2203–12. Sophisticated model system in which different joint tissues as well as their interaction and response to drugs can be studied.


