The Genetics of Osteoarthritis: A Review

Sophie Catriona Warner 1,* and Ana Maria Valdes 1,2

1 Academic Rheumatology, University of Nottingham, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB, UK; ana.valdes@nottingham.ac.uk
2 Arthritis Research UK Pain Centre, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB, UK

* Correspondence: sophiecw Warner@gmail.com; Tel.: +44-115-82-31759; Fax: +44-115-82-31757

Abstract: Osteoarthritis (OA) is the most common form of arthritis and is a leading cause of pain and disability worldwide. There is thought to be an important genetic component to the development of OA. In this review article, the methods used to study the genetics of OA are discussed, along with the main findings from these analyses, with a particular focus on the results of genome-wide association scans (GWAS). Functional validation of the results from genetic studies of OA is also described to put the genetic results into a biological context and show their relevance to the mechanisms involved in the development and progression of OA. The clinical relevance of the genetic findings to date is also discussed, as well as the research, which is still needed in this field to further improve the understanding of OA pathogenesis.

Keywords: genetics; osteoarthritis; genome-wide association scan; pain

1. Introduction

Osteoarthritis (OA) is the most common form of arthritis and is a leading cause of pain and disability worldwide [1]. The total economic burden for arthritis is estimated to be 1%–2.5% of the gross national product in Western countries [2]. Symptoms include pain and stiffness, which may result in a loss of mobility and reduced quality of life [3]. The joint damage seen in OA can potentially occur in all tissues of the synovial joint, including cartilage and the underlying bone, muscles and nerves [4]. The increases in obesity and life expectancy in the UK and other populations worldwide [5–7] are leading to an increase in the prevalence of OA, making it a major public health concern [8–10].

Little progress on new therapies for OA has been made in the last decade compared to other rheumatological conditions, such as gout, rheumatoid arthritis and lupus [11]. Treatments for OA are still of only limited effectiveness in terms of addressing the changes caused by the disease [12].

So far, the identified risk factors do not fully explain an individual’s risk of developing OA and the progression of the disease [13]. The genetic contribution to OA has been supported by many studies, and it is thought that at least 30% of the risk of OA is genetically determined [1,14,15].

OA can affect one joint only or be more generalised, affecting multiple joint sites. The complexity of OA means there is not a single genetic variant responsible [16]. Instead, there are many genes that have been implicated in the incidence and progression of OA, each with a small effect on risk [17]. In addition to this, gene-environment interactions would eventually also need to be taken into account to improve the understanding of the mechanisms involved in OA.

There is a significant unmet need for better understanding of the molecular mechanisms underlying the development and progression of OA. Genetic studies have the potential to improve this knowledge and therefore identify and improve intervention and treatment targets at all stages of the disease.
2. Heritability of Osteoarthritis (OA)

If people with a relative affected by a disease have a higher risk of also developing that disease compared to the general population, it is possible that there is a heritable component to the disease. Heritability refers to the extent that genetic variation determines variation of a trait of interest, such as the development of OA [18].

The extent to which a trait runs in families is called “familial aggregation” [18]. The lambda sib statistic (\( \lambda_s \)) represents the sibling recurrence risk, which is the risk ratio of disease for a sibling of an affected individual [18,19]. In OA, the sibling recurrence risk has been estimated to be 2.08–2.31 for radiographic knee OA and 4.27–5.07 for radiographic hip OA [18]. This estimate ranges from 2.8–4.8 for the risk of total knee replacement (TKR) and 1.78–8.5 for total hip replacement (THR) [18]. These estimates are relatively low compared to the \( \lambda_s \) for other conditions, such as rheumatoid arthritis (\( \lambda_s \) ranges from 2–10 [20–24]), but in line with those seen for type 2 diabetes (\( \lambda_s \) ranges from 3–4 [25,26]). A limitation of this method is that the effects of a shared environment and similar lifestyle choices of individuals within the same family are not accounted for [18].

Classical twin studies compare identical and non-identical pairs, so environment can be controlled for, addressing the limitations described above [18]. In twin studies, genetically-identical twins are compared to non-identical twins to best estimate the extent of the contribution of genetic variation on a trait of interest (heritability) [14,27–30].

Variance component methods in pedigrees have also been used to study the heritability of OA [29].

Adjusting for confounding factors, such as age, sex and body mass index (BMI), which are known risk factors for OA, helps to strengthen these analyses. After adjusting for environmental risk factors and confounding factors, the heritability of OA has been estimated at between 30% and 65% [1,14,15].

Table 1. Heritability of large joint osteoarthritis-related traits in twin and family studies.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Heritability (h²)</th>
<th>Type of Study</th>
<th>Data from Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiographic knee OA</td>
<td>39%</td>
<td>Twins</td>
<td>[14]</td>
</tr>
<tr>
<td>Radiographic hip OA</td>
<td>60%</td>
<td>Twins</td>
<td>[27]</td>
</tr>
<tr>
<td>Generalised OA</td>
<td>42%</td>
<td>Spouse pairs,</td>
<td>[31,32]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Parent-child pairs,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sibling pairs</td>
<td></td>
</tr>
<tr>
<td>Knee pain reporting</td>
<td>44%–46%</td>
<td>Twins</td>
<td>[28,29]</td>
</tr>
<tr>
<td>Radiographic progression</td>
<td>69%</td>
<td>Twins</td>
<td>[30]</td>
</tr>
<tr>
<td>Cartilage volume</td>
<td>77%–85%</td>
<td>Sibling pairs</td>
<td>[33]</td>
</tr>
<tr>
<td>Changes lower limb in muscle strength</td>
<td>64%</td>
<td>Sibling pairs</td>
<td>[29]</td>
</tr>
</tbody>
</table>

3. Investigating the Genetic Risk Factors for OA

The aim of genetic studies is to identify which genetic variants at which genetic locations (loci) influence the risk of a disease. These variants can be identified using various approaches. One approach is to study candidate genes, i.e., to test variants in genes already hypothesised to be involved in OA. Another approach used in the past was genome-wide linkage scans, which consisted of testing ~400 genetic markers in the genome to assess differences in segregation in families with
members affected by OA. In the past few years, the approach of choice is that of genome-wide association scans (GWAS), which can be performed on a small or large scale depending on the number of variants used.

Candidate gene studies have been responsible for identifying several susceptibility loci for OA. A limitation of this analysis is that a priori knowledge is needed about disease aetiology. Another is that only very small regions of the genome can be investigated at a time, meaning that important genes may be overlooked using this method. Genes such as GDF5, ASPN, FRZB and PTGS2 have been identified this way. These genes continue to be the subject of functional studies and further genetic replication in independent populations [34–40]. However, a recent large study on 199 candidate genes found that only two variants (in the COL11A1 and VEGF genes) reached the significance level after adjustment for multiple tests (p < 1.58 × 10⁻⁵) [41]. The results of this study confirm that using existing knowledge of joint biology to identify variants in candidate genes is unlikely to contribute to the understanding of the risk of OA [41].

Hypothesis-free approaches, such as GWAS, have been suggested as the best approach when studying OA, due to the small amount of knowledge available so far about the genetic (and therefore molecular) mechanisms involved in OA pathogenesis [42]. At the time of writing, the above methods have identified 21 independent susceptibility loci for OA [43].

3.1. Traits and Outcomes Studied in the Genetics of OA

As OA is the common outcome of a group of disorders [17], there are different mechanisms and molecular pathways involved. These depend on an individual’s risk factors, and so, different genes and pathways can affect an individual’s risk of OA (see Figure 1). The heterogeneous nature of OA means that different presentations and symptoms are possible, leading to subphenotypes [13]. Further study is needed on symptoms, such as pain, the extent of cartilage loss, the presence of synovitis or generalised disease versus non-generalised disease, to clearly identify the different clinical phenotypes in OA, particularly in early OA [13]. Genetic factors contribute to different aspects of the disease, and the genetic contribution for some of these factors is only beginning to be addressed.

**Figure 1.** Schematic diagram of the different types of risk factors for osteoarthritis. BMI, body mass index.
It is extremely important to select an appropriate phenotype in genetic association studies [42]. Heterogeneity (diversity) in the phenotype can reduce the power to find true, significant and plausible results in genetic analysis [6,44].

3.2. Pain in OA

Pain, not tissue damage to the joint, is the main clinical outcome of OA and is a major unmet medical need [3,45]. One of the challenges in researching and treating OA is that there is only a weak association between radiographic changes of OA and the pain experienced by an individual after taking into account psychological and social factors and comorbidities [46,47].

On the whole, OA pain mechanisms are poorly understood, as is the molecular aetiology of pain in OA. Therefore, current treatments for pain in OA are not as effective as they could potentially be. Human genetic studies of OA pain will help discover the molecular pathogenesis of OA. An advantage of genetic studies is that they are non-invasive and can be used to test relevant clinical populations in large numbers.

Genetic variants implicated in pain sensitivity have been shown to be significantly different between asymptomatic radiographic cases of OA and symptomatic cases. These included amino acid change variants in COMT, SCN9A and TRPV1 [48–54]. These results suggest that the genetic risk of clinical OA is influenced by the genetic contribution to pain sensitivity. In addition to this, replication analysis of a variant in the PCSK6/PACE4 gene showed a consistently increased frequency of the minor allele in asymptomatic knee OA cases compared to symptomatic radiographic knee OA cases, and Pcsk6 knockout mice were significantly protected against pain [55].

4. Genome-Wide Association Studies

The completion of the Human Genome Sequencing Project in 2003 [56] has uncovered large numbers of polymorphisms that can be used as markers when studying genetic associations [57]. In addition to this, the International HapMap Project (whose aim was to describe the common patterns of human DNA sequence variation) has improved the understanding of the correlation between SNPs (linkage disequilibrium (LD)) [58]. This means the variation in the entire genome can be represented in the hundreds of thousands of SNPs used in a GWAS. GWAS uses these SNPs to test for an association with a disease or trait of interest [57]. These developments have made it easier to study the genetics of OA and other complex diseases, meaning in the future, the understanding of the causes of common diseases, such as OA, can be improved. This will improve the quality of life for many individuals. GWAS have become accepted as the way forward in the search for susceptibility loci for common diseases.

4.1. Genome-Wide Association Scans (GWAS) for OA

4.1.1. Background

Several GWAS have been carried out for the risk of OA and have been reported in the literature (see Table 2). However, a very large sample size (in the region of tens of thousands of individuals) or large genetic effects are needed to reach the genome-wide significance threshold used in GWAS ($p < 5 \times 10^{-8}$). This is to adjust for the huge number of tests run during a GWAS analysis.
Table 2. Genetic associations with large joint OA and related traits derived from genome-wide association scans (GWAS) studies with $p < 1 \times 10^{-7}$.

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Gene</th>
<th>Ethnic Group</th>
<th>Trait</th>
<th>Total Sample Size</th>
<th>$p$-Value</th>
<th>Putative or Known Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11718863</td>
<td>DVWA</td>
<td>Asians</td>
<td>Knee OA</td>
<td>982 cases, 1774 controls</td>
<td>$7 \times 10^{-11}$</td>
<td>Cartilage-specific tubulin binding</td>
<td>[8]</td>
</tr>
<tr>
<td>rs111777/rs6976 *</td>
<td>GLN3/GLT8D1</td>
<td>Caucasians</td>
<td>Hip or knee OA</td>
<td>14,883 cases, 53,947 controls</td>
<td>$7 \times 10^{-11}$</td>
<td>Cell cycle control, tumorigenesis and cellular senescence</td>
<td>[6]</td>
</tr>
<tr>
<td>rs4836732</td>
<td>ASTN2</td>
<td>Caucasians</td>
<td>Hip OA and total hip replacement (THR)</td>
<td>5813 cases, 53,947 controls</td>
<td>$6.1 \times 10^{-10}$</td>
<td>Glial neuronal migration</td>
<td>[6]</td>
</tr>
<tr>
<td>rs930591</td>
<td>FILIP1/SENP6</td>
<td>Caucasians</td>
<td>THR</td>
<td>5813 cases, 53,947 controls</td>
<td>$2 \times 10^{-9}$</td>
<td>Various genes</td>
<td>[6]</td>
</tr>
<tr>
<td>rs10947262</td>
<td>BTNL2</td>
<td>Asians</td>
<td>Knee OA</td>
<td>906 cases, 3396 controls</td>
<td>$5 \times 10^{-9}$</td>
<td>Various genes</td>
<td>[59]</td>
</tr>
<tr>
<td>rs4730250</td>
<td>COG5/GPR22/DUS4L/HBP1</td>
<td>Caucasians</td>
<td>Knee OA</td>
<td>6709 cases, 44,439 controls</td>
<td>$9 \times 10^{-9}$</td>
<td>Various genes (see [60])</td>
<td>[60]</td>
</tr>
<tr>
<td>rs11842874</td>
<td>MCT2L</td>
<td>Caucasians</td>
<td>Knee or hip OA</td>
<td>19,041 cases, 24,504 controls</td>
<td>$2 \times 10^{-8}$</td>
<td>Cell motility</td>
<td>[61]</td>
</tr>
<tr>
<td>rs10492367</td>
<td>PTHLH</td>
<td>Caucasians</td>
<td>Hip OA</td>
<td>6329 cases, 53,947 controls</td>
<td>$1.5 \times 10^{-8}$</td>
<td>Chondrogenic regulator</td>
<td>[6]</td>
</tr>
<tr>
<td>rs835487</td>
<td>CHST11</td>
<td>Caucasians</td>
<td>THR</td>
<td>5813 cases, 53,947 controls</td>
<td>$2 \times 10^{-8}$</td>
<td>Chondroitin sulfotransferase involved in cartilage metabolism</td>
<td>[6]</td>
</tr>
<tr>
<td>rs775228</td>
<td>HLA–DQB1</td>
<td>Asians</td>
<td>Knee OA</td>
<td>906 cases, 3396 controls</td>
<td>$2 \times 10^{-8}$</td>
<td>Immune response (antigen presentation)</td>
<td>[59]</td>
</tr>
<tr>
<td>rs12107036</td>
<td>TP63</td>
<td>Caucasians</td>
<td>Total knee replacement (TKR) in women</td>
<td>4085 cases, 33,587 controls</td>
<td>$7 \times 10^{-8}$</td>
<td>Member of the p53 family of transcription factors</td>
<td>[6]</td>
</tr>
<tr>
<td>rs8044769</td>
<td>Fat mass and obesity-associated (FTO)</td>
<td>Caucasians</td>
<td>TKR in women</td>
<td>4085 cases, 33,587 controls</td>
<td>$7 \times 10^{-8}$</td>
<td>Control of energy homeostasis</td>
<td>[6]</td>
</tr>
<tr>
<td>rs10948172</td>
<td>SUPT3H/RUNX2</td>
<td>Caucasians</td>
<td>OA (hip or knee) in men</td>
<td>5617 cases, 20,360 controls</td>
<td>$8 \times 10^{-8}$</td>
<td>Probable transcriptional activator</td>
<td>[6]</td>
</tr>
<tr>
<td>rs6094710</td>
<td>NCOA3</td>
<td>Caucasians</td>
<td>Hip OA</td>
<td>11,277 cases, 67,473 controls</td>
<td>$7.9 \times 10^{-9}$</td>
<td>Nuclear receptor</td>
<td>[62]</td>
</tr>
<tr>
<td>rs788748</td>
<td>IGBP3</td>
<td>Caucasians</td>
<td>Hip OA</td>
<td>3243 cases, 6891 controls</td>
<td>$2 \times 10^{-8}$</td>
<td>Cartilage catabolism and osteogenic differentiation</td>
<td>[63]</td>
</tr>
<tr>
<td>rs12982744</td>
<td>DOT1L</td>
<td>Caucasians</td>
<td>Hip OA</td>
<td>9789 cases, 31,873 controls</td>
<td>$8.1 \times 10^{-8}$</td>
<td>Wnt signalling</td>
<td>[64]</td>
</tr>
</tbody>
</table>

* these SNPs are in perfect linkage disequilibrium with each other.
4.1.2. Early Small-Scale GWAS Findings

Small-scale GWAS were the first studies of this kind, using approximately 100,000 SNPs to test for associations in small discovery populations and larger scale replication populations. The DVWA gene was identified in a study of this kind as being associated with OA in Asian populations, using 982 cases and 1774 controls [8]. A later UK GWAS of 357 cases and 285 controls identified a variant (rs4140564) between the PTGS2 and PLA2G4A genes [65]. Since then, several large-scale GWAS studies for OA and OA-related traits have been published, including studies from the Netherlands, the UK, Japan and the USA. These findings all help to contribute to the understanding of OA, a condition that affects millions of individuals worldwide.

4.1.3. Large-Scale GWAS Findings

Although not genome-wide significant, two SNPs within a small region of the human leukocyte antigen (HLA) locus on chromosome 6p were associated with knee OA in a Japanese study of 1879 cases and 4814 controls ($p = 7 \times 10^{-8}$) [59]. Surprisingly, this finding was not replicated in a Han Chinese population [66]. Large European cohorts totalling 5749 cases and 36,408 controls also conclusively showed no association with these two SNPs [66,67]. These results highlight the differences seen in genetic risk between different populations.

A GWAS from Rotterdam used a relatively small discovery cohort of 1341 OA cases and 3496 controls from the Netherlands. Replication analysis was performed on additional European and North Americans of European descent cohorts. This brought the total number of participants to 14,934 cases and 39,000 controls, providing a respectively-powered study. Sufficiently-powered studies are those that use an adequate sample size to detect the effects of a certain size at a given significance threshold ($p < 5 \times 10^{-8}$ for GWAS). A signal ($p < 8 \times 10^{-8}$) was identified in a region on chromosome 7q22. This region includes a large LD block extending over 500 kb associated with knee and hand OA. This signal was improved upon the addition of several more cohorts to the original study, increasing the evidence for this being a true signal. The LD block identified here contains six known genes, which are all equally good candidates for association with OA. Meta-analysis confirmed the genome-wide significant association of this signal with OA in European-descent samples with an odds ratio (OR) of 1.17 (95% confidence interval (CI) 1.11–1.24) and a $p$-value of $p = 9.2 \times 10^{-9}$, but not in Asian populations, where the OR was 1.03 (95% CI 0.85–1.25) and the $p$-value was not significant [60]. Odds ratios represent the difference in risk in an exposed group compared to an unexposed one. In the above results, an OR of 1.17 means that individuals with a certain genetic makeup (genotype) at the location of interest have a 17% increased risk of OA compared to individuals with a different genotype here.

The ARC Osteoarthritis Genetics (arcOGEN) study is a UK-based consortium based around seven collection centres. At the time of writing, this is still the most powerful GWAS on OA to date [42]. More recently, UK scientists have generated 1000 Genomes Project-based imputation [61] on the same data described above from the arcOGEN consortium. In this analysis, 3177 cases, some of whom were post-total joint replacement (TJR), and 4894 controls were used [61]. The use of these data allowed unidentified risk loci to be found, as well as improving the power of the study [61,68]. Robust associations with SNPs in the MCF2L gene have been discovered as the result of large-scale replication, using individuals with knee OA, hip OA and post-TJR [61]. This gene regulates neurotrophin-3-induced cell migration in Schwann cells (in the nervous system) [61]. Neurotrophin-3 is a member of the nerve growth factor (NGF) family and binds to two specific NGF receptors (tropomyosin receptor kinases B and C (TrkB and TrkC)) [61]. NGF inhibition has been found to improve joint pain and function in knee OA [69]. This exciting finding supports the relevance of molecular pain pathways in OA susceptibility. However, to date, no functional work has been carried out showing the role of the MCF2L gene in this context. The pain sensitivity of patients with OA is beyond the scope of this article; however, as discussed above in Section 3.2, there is evidence in the literature to support a genetic susceptibility to higher pain thresholds, specifically that conferred by the TPRVI gene, and a lower risk of reporting symptomatic OA. This finding also highlights the need for
more extensive coverage of the genome than that seen in medium-sized arrays (e.g., 600,000 SNPs), as this result was only discovered once many more variants were tested via imputation. This is something that should be considered when designing future studies. The top signal from this analysis on the arcOGEN data and imputation data was rs1184874 in the SPP13 gene, reaching a combined odds ratio of \( \hat{OR} = 1.17 \) (95% CI 1.11–1.23) and a \( p \)-value of \( 2.1 \times 10^{-8} \) [61].

A GWAS was performed by a large UK consortium on 7410 severe OA cases from the arcOGEN study (80% of whom were post-TJR) and 11,009 unrelated controls from the UK [6]. Replication was performed on the most promising signals using cases and controls from studies in Iceland, Estonia, the Netherlands and the UK. As discussed below, several new risk loci for large joint OA have been identified as a result of this study (see Table 2).

The most significant association was located on chromosome 3. Two SNPs in perfect LD with each other gave the next most significant results: rs11177 in the GNL3 gene and rs6976 in the GLT8D1 gene (\( \hat{OR} = 1.12 \) 95% CI 1.08–1.16; \( p = 7.24 \times 10^{-11} \)) [6].

The four other genome-wide significant (\( p < 5 \times 10^{-8} \)) results from this analysis relate to hip OA: (1) rs4836732 located in the ASTN2 gene (\( \hat{OR} = 1.18 \) 95% CI 1.12–1.25; \( p = 2.42 \times 10^{-9} \)); (2) rs9350591 between the FTO and SENP6 genes (\( \hat{OR} = 1.18 \) 95% CI 1.12–1.25; \( p = 2.42 \times 10^{-9} \)); (3) rs10492367 (\( \hat{OR} = 1.14 \) 95% CI 1.09–1.20; \( p = 1.48 \times 10^{-6} \)) between the KLHDC5 and PTHLH genes; and (4) rs835487 (\( \hat{OR} = 1.13 \) 95% CI 1.09–1.18; \( p = 1.64 \times 10^{-5} \), THR) in the CHST11 gene [6]. Three other signals approached genome-wide significance with different effects seen in males and females [6]. This sex-specific result is of interest, as OA is seen more commonly in women, and the potential differences in the risk factors, development and progression of OA between men and women require further study. The role of the molecules identified here remains mostly unclear with regards to their role in OA pathogenesis.

Although not genome-wide significant, rs8044769 in the fat mass and obesity-associated (FTO) gene was identified in this GWAS (\( \hat{OR} = 1.11 \) 95% CI 1.07–1.15; \( p = 6.85 \times 10^{-8} \)) in females. After genotyping in seven independent UK and Australian cohorts, this variant’s association with OA was investigated in case-control analyses, both with and without adjustment for BMI category [70]. A Mendelian randomisation approach was employed using the FTO variant to evaluate the role of being overweight on the risk of OA [70]. After meta-analysis, a significant association with knee OA was seen without adjusting for BMI (\( \hat{OR} = 1.08 \) 95% CI 1.02–1.14; \( p = 0.009 \)); however, the signal fully attenuated after adjusting for BMI (\( \hat{OR} = 0.99 \) 95% CI 0.93–1.05; \( p = 0.666 \)) [70]. This Mendelian randomisation approach confirms the causal role of being overweight in OA [70]. The increasing prevalence of both OA and being overweight and obesity in many populations [5–7] makes this a relevant and important finding.

A recent GWAS meta-analysis on 78,000 individuals identified a genome-wide significant variant (rs6094710) in the NCOA3 gene (\( \hat{OR} = 1.28 \) 95% CI 1.18–1.39; \( p = 7.9 \times 10^{-5} \)) [62]. This \( p \)-value was improved after combined analysis of the discovery (\( p = 5.6 \times 10^{-8} \)) and follow-up studies (\( p = 7.3 \times 10^{-4} \)). Two loci remained suggestively associated: rs5009270 at 7q31 (\( \hat{OR} = 1.10 \); \( p = 9.9 \times 10^{-7} \)) and rs3757837 (\( \hat{OR} = 1.27 \); \( p = 2.2 \times 10^{-6} \) in a male-specific analysis) [62].

A GWAS on hip OA, using data from the Osteoporotic Fractures in Men Study (MrOS) and the Study of Osteoporotic Fractures (SOF), was replicated in five independent studies [63]. The rs788748 SNP located near the IGFBP3 gene was genome-wide significant in this analysis and associated with a lower risk of hip OA (\( \hat{OR} = 0.71 \); \( p = 2.0 \times 10^{-5} \)). Although the association replicated in all five studies, the signal was weakened after replication, suggesting a possible false positive result (\( \hat{OR} = 0.92 \); \( p = 0.020 \)). Despite this, a role of this variant and gene in OA is suggested by the results of functional validation studies (see Section 4.2) [63]. Further replication is necessary, ideally with larger sample sizes, to make a more confident assertion of this variant’s role in OA [63].

A GWAS on cartilage thickness at the hip has been carried out using data from the Rotterdam study (RSI and RSII cohorts). An SNP in the DOT1L gene was strongly associated with minimum joint space width (mJSW) at the hip [71]. After replication in independent UK cohorts, an overall genetic
effect size (expressed as the regression coefficient $\beta$) of 0.09 mm/allele was achieved ($p = 1.1 \times 10^{-11}$ after meta-analysis) [64]. The risk allele for lower mJSW at this SNP was later associated with a 10% increased risk of hip OA ($p = 8.8 \times 10^{-8}$). This effect reached genome-wide significance in males (OR = 1.17 95% CI 1.11–1.23; $p = 7.8 \times 10^{-9}$), but was only nominally significant in women with a small effect size (OR = 1.05), consistent with the sexual dimorphism seen in some forms of hip OA [64]. The authors note that these findings suggest a role for DOT1L as a therapeutic target in hip OA [64].

The first GWAS to report genome-wide significant results for hand OA revealed an SNP (rs3204689) in the ALDH1A2 gene to be significantly associated with an increased risk of hand OA in an Icelandic discovery cohort (OR = 1.51, $p = 3.99 \times 10^{-10}$). This finding was replicated in cohorts from the UK and the Netherlands, showing an improved association and significantly increased risk of hand OA (OR = 1.46, $p = 1.10 \times 10^{-11}$). In silico replication found a significant association with this variant for knee OA, but not hip OA. Interestingly, this was a protective effect (OR = 0.95, $p = 0.044$), the opposite of the effect seen on hand OA. The authors note surprise at this finding, as the literature has previously suggested a close relationship between hand and knee OA [15]. This finding highlights the need for further study in the genetics of hand OA and how OA at different sites relates to OA at other sites.

Despite these important findings, more studies of this kind are needed. For example, there are currently no GWAS yet on the subject of hip morphology traits. This is due to sample size limitations. More individuals with both hip shape measures and genome-wide genotyping are required for a study of this kind.

4.2. What GWAS Has Taught Us about the Pathogenesis of OA

Good results have been seen from combining GWAS in meta-analyses. However, there are many factors that can lead to false positive results, which need to be considered in this kind of analysis. One major problem has been that of heterogeneity between the studies or populations. For example, population stratification due to differences in geographical location and ethnic background can affect genetic variation [25]. However, inconsistent ascertained criteria for inclusion into the study can also prove to be a problem. As a result, recommendations for OA genetic studies have been generated [72].

If the genes associated with OA in these analyses are expressed in joint tissues, differences in gene expression in OA-affected versus unaffected tissues can be looked for to provide functional validation and a better understanding of the biological relevance of the results of genetic analyses. Findings from functional validation studies could in theory identify ways to help overcome the genetic deficit seen in individuals with a genetic risk profile for OA at some genetic locations [73,74].

As the most robust susceptibility locus for OA to date, rs143383 in the GDF5 gene has been studied in this way. In a mouse model, GDF5 deficiency was associated with instability-driven joint damage, gait and subchondral bone changes [75]. The T susceptibility allele in rs143383 is also associated with lower gene expression levels [34,36]. An epigenetic study of human chondrocytes showed that GDF5 expression is influenced by CpG methylation at rs143383 [76]. Different effects on GDF5 expression have been seen in OA knee cartilage compared to OA hip cartilage and non-OA hip cartilage [76]. The joint-specific effects of this SNP on OA are supported by these findings [76].

Other OA susceptibility loci have been the subject of functional validation studies, including rs2615977 in the COL11A1 gene, rs11177 in the GNL3 gene (which is in perfect LD with rs6976 in the GLT8D1 gene), rs3815148 between the HBP1 and COG5 genes and rs6094710 in the NCOA3 gene. Gene expression or protein levels were found to be significantly different for these when comparing OA and non-OA human joint tissues [6,62,77–79]. The rs9350591 SNP between the FILIP1 and SENP6 genes was not significantly associated with SENP6 gene expression, but SENP6 expression was significantly lower in hip OA cartilage versus control cartilage [80].

Silencing of the DOT1L gene inhibited chondrogenesis in human chondrocytes in an in vitro study [71]. A mouse study has also supported the role of DOT1L in cartilage development [71]. IGFBP3 knockdown (reducing the gene’s expression) in chondrocytes prevented hypertrophy, a change
associated with OA pathogenesis [63]. Similarly, overexpression of IGFBP3 led to changes to cartilage consistent with those seen in OA pathogenesis [63].

Despite these findings, further investigation into the functional effects of these loci and genes is needed. Better understanding of the regulation of these genes will lead to better understanding of the effects of these OA susceptibility loci on OA pathogenesis and identify possible treatment targets [74].

4.3. Clinical Relevance

One of the criticisms raised against genetic studies is that they are far removed from clinical practice. However, the understanding of the pathogenesis of OA has been and is being informed by the results of genetic studies.

For example, categorising OA into idiopathic (of an unknown cause) and post-traumatic has traditionally been the case [81]. However, these forms of OA have been shown in a number of studies to differ substantially in their characteristics in terms of age, sex, BMI and radiographic characteristics. Despite these differences, a history of trauma was found to have the same, or even a slightly higher, contribution to the risk of OA when compared to individuals with idiopathic OA. It is suggested by these results that there is a clear genetic susceptibility to OA, which is exacerbated by exposure to trauma and not just by exposure to different molecular mechanisms [81].

Genetic studies have also suggested that different molecular mechanisms may underlie the generalised and non-generalised forms of large joint OA [15].

The potential of these genetic markers as a diagnostic tool has been explored in various radiographically-based OA cohorts. One study investigated the predictive value of nine of the markers listed in Table 2 ($p < 1 \times 10^{-7}$ after meta-analysis in Caucasian populations) as a risk factor for the development of OA [82]. Prediction models here shown poor predictive value [82]. The SNPs used in this study showed small effects on OA risk and did not predict the development of knee OA well in this population [82]. This suggests that although these markers may be prevalent in the population and, in the case of the rs12982744 SNP in the DOT1L gene, a potential therapeutic target, the markers identified so far are associated with only a small increase in risk, and therefore, it would not be clinically useful at present to test individuals for these markers [82].

As OA increases in prevalence, the health and economic burdens will also increase, and this has intensified the need for disease-modifying pharmacological treatments for OA. The heterogeneity seen in OA has made this an extremely challenging task. In addition to the different presentation of symptoms, differences in the response to treatment can be seen in people with OA.

The effective application of developed treatments depends on the ability for them to be applied in the early stages of the disease before irreversible or severe joint damage occurs. In order to do so, several requirements must be met. Firstly, the groups of people who are likely to respond to a given pharmacotherapy need to be identified (e.g., antiresorptives [83]). Next, the groups of people who are likely to progress within a reasonable time frame need to be identified, in order to reduce the cost of clinical development. Finally, the early stages of OA need to be identified and diagnosed. Genetic markers, in combination with imaging and biochemical markers, have the potential to assist in all three of the above tasks. However, in order for this to become a reality, a much larger number of genetic variants, able to explain a larger proportion of the genetic contribution to OA, are needed.

5. Conclusions

The genetics of OA are important in order to understand its initiation and progression. The increasing availability of genotype data from more cohorts means consistent signals are becoming increasingly possible, and in the near future, more genes will be discovered. This is already being reported at various international conferences.

In the future, it will become feasible to directly evaluate the role of the genetic variation in the entire genome in the risk of disease, as a result of next generation sequencing methods. Furthermore,
the biological and functional relevance of these genetic findings is essential to help put the research into a clinical context to benefit people with OA [43].

Acknowledgments: Sophie Catriona Warner is funded by a PhD studentship awarded by the University of Nottingham. This work was supported by a EULAR project grant to Ana Maria Valdes (Grant 108239) and by the Arthritis Research UK Pain Centre (Grant 18769). The authors did not receive funds and do not have funds for covering the costs to publish in open access.

Author Contributions: Sophie Catriona Warner and Ana Maria Valdes designed and wrote the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

OA Osteoarthritis  
GWAS Genome-wide association scan  
SNP Single nucleotide polymorphism  
LD Linkage disequilibrium

References


25. Syddall, C.M.; Reynard, L.N.; Young, D.A.; Loughlin, J. The identification of trans-acting factors that regulate the expression of GDF5 via the osteoarthritis susceptibility SNP rs143383. PLoS Genet. 2013, 9, e1003557. [CrossRef] [PubMed]


67. Zhang, W.; Dolan, M.E. Impact of the 1000 genomes project on the next wave of pharmacogenomic discovery. Pharmacogenomics 2010, 11, 249–256. [CrossRef] [PubMed]


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).