The handle http://hdl.handle.net/1887/33065 holds various files of this Leiden University dissertation

**Author:** Ramos, Y.F.M.
**Title:** Osteoarthritis, a degenerative disease of the articular joints : towards the implementation of functional genomics in OA
**Issue Date:** 2015-05-26
Introduction
OSTEOARTHRITIS, AN OVERVIEW

Osteoarthritis (OA) is a degenerative disease of the joints showing increased incidence with aging. While the incidence of the disease is steadily increasing due to an increasing population of elderly and rising numbers of people suffering from obesity, OA has become part of the top 5 most common diseases especially in Europe and the United States.[2-4] Of 291 conditions included in the Global Burden of Disease 2010 study (GBD 2010) OA ranked 11th highest in terms of so-called ‘years of life lived with disability’ (YLDs;[1]) and it is anticipated that OA will become a major problem for health systems globally with large socio-economic burden on society. Occurrence of OA is rare before the age of 40 years, with prevalence of knee OA peaking around 50 years of age while prevalence of hip OA continues to increase with age (Figure 1A). Similar to hip OA, prevalence of hand OA increases with age and percentages are joint-specific (Figure 1B).[5, 6]

Since OA is a chronic and painful disability and a major cause of morbidity, development of drugs to delay onset and/or progression is essential. This requires however more knowledge of the etiology, improved mechanistic insight into the disease, and sensitive biomarkers to diagnose OA at an early stage before irreversible damage becomes apparent in X-ray, manifested by the presence of osteophytes, joint space narrowing, and subchondral cyst formation of the bone. Up until now, treatments that could cure or stop disease progression are still poor and osteoarthritis is not diagnosed until late stages when damage is irreversible. Therapeutic approaches are aimed at pain relief, controlling symptoms, and minimizing disability, until in the end stage total joint arthroplasty is required.[7] Main impediments in OA management are the lack of biomarkers that detect OA at an early stage when damage is still reversible, or that allow stratification of subsets of OA patients with different disease modes. This considerably hampers development of effective therapies and requires better insight into the heterogeneous etiology of OA.

**Figure 1 Prevalence osteoarthritis, stratified by joint-site and sex.**

**A** Hip and knee OA (data for North American population[1]).

**B** Hand OA stratified for 3 joints (data from an American community-based, longitudinal study[5]).

Legend: red line: females; blue line: males; diamonds: knee OA; filled circles with dashed line: hip OA; asterisks with dashed line: hand OA 2nd DIP; triangles: hand OA 3rd PIP; squares with dotted line: hand OA 1st CMC.
Chapter 1

ONE NAME, MULTIPLE DISEASES

Even though the most common form of arthritis has a single name, osteoarthritis, in fact it has become manifest that it is a heterogenic disease. In general, two categories of OA are distinguished (reviewed by Johnson and Hunter[8]): primary or idiopathic OA which has no clear cause, and secondary OA which results from a known event (e.g. traumatic events[9]) or diseases such as acromegaly[10] or patellofemoral joint osteoarthritis[11]). Although OA is most prevalent at the joints of hand, hip, and knee, it can occur at any joint.[2]

For cases of primary OA, depending on the occurrence at a specific joint-site or the affection of 3 or more joints, respectively localized and generalized OA can be distinguished. Furthermore, different subtypes can be discriminated, for example: hypertrophic (when only osteophytosis is present), atrophic (only joint space narrowing), and classic or normotrophic OA (when both, joint space narrowing and osteophytosis, are present).[12]

Diagnosis of OA relies on the combination of clinical, radiographical, and physical findings. Clinical symptoms of osteoarthritis are pain, joint stiffness and crepitus, resulting in an increasing loss of mobility. Radiographically, OA becomes manifest by osteophytes, joint space narrowing (JSN), and subchondral cysts of the bone. However, the patient’s perception of pain is often not proportional to the clinical and radiographical stage of the disease thereby hampering consistent and analogous diagnoses. In addition, substantial discordance exists between clinical and radiographical OA.[13]

Osteoarthritis as whole joint disorder.

Originally, degradation of articular cartilage was emphasized most in OA pathology. However, based on their studies already around 1970 Radin and colleagues proposed for the first time that cartilage degeneration may in fact result from changes in the (subchondral) bone.[14, 15] At present, it is generally accepted that OA is a whole joint disease involving not only bone and cartilage but also ligaments, synovial tissues, and the joint musculature,[16, 17] although it is most likely that OA originates in bone and/or cartilage subsequently affecting the other tissues.

Epidemiological studies have shown that the risks for osteoarthritis and osteoporosis (OP) are inversely correlated, with higher systemic bone mass being a risk factor for OA which was recently confirmed in a study by Hardcastle and colleagues.[18] Unfavorable mechanical forces imposed on the cartilage during movement due to suboptimal (higher) bone mineral density (BMD) could result in the accumulation of microtraumas throughout life, eventually leading to OA. Despite the fact that bone and cartilage seem of solid, impermeable composition it has been shown that molecules can diffuse into the extracellular matrix of both, (calcified) cartilage and subchondral bone.[19] Catabolic enzymes or other molecules released in the bone following subchondral bone remodeling may diffuse to the cartilage and activate chondrocytes or vice versa.

GENETIC EPIDEMIOLOGY OF OSTEOARTHRITIS

The multifactorial origin of OA is reflected
Introduction

by the observation that it is transmitted in a non-Mendelian manner. Major risk factors for development and progression of OA are joint injury, aging, poor metabolic health (e.g. obesity), and genetic predisposition.[16, 20, 21] Importantly, genetic factors influencing OA can apply via metabolic health[22, 23] as well as via skeletal features: bone and cartilage structure and composition, and aspects of shape[24-26] (Figure 2). The heritability of osteoarthritis was already hypothesized in the beginning of last century and it was firstly confirmed in family-based studies showing that 39-60% of OA incidences can be attributed to heritability.[27, 28] In fact, Dieppe et al.[29] stated that development of OA is the result of the dynamics between genetics and environment, with a continuous distribution between the extremes of predominantly genetic and predominantly environmental influences.

Since clear insight into the etiology of OA would greatly contribute to early

Figure 2 Role of genetic variation in susceptibility to osteoarthritis and effects for joint tissues.
Chapter 1

diagnosis as well as to the development of disease modifying treatments that are still lacking and to the possibility to better predict and monitor OA progression, major efforts are made to unravel the genetic secrets of OA. As a complex disease, the genetic component of OA is highly polygenic, with multiple risk alleles conferring small effects. This adds up to the mentioned clinical and radiological heterogeneity among patients. Therefore, genetic studies in OA require careful consideration of the study design and patient selection to increase the chance to identify susceptibility genes. Whereas large-scale cohorts of random samples (population-based or case-control studies) are more likely to identify common variants with small effect sizes, family-based studies of more severe OA phenotypes are expected to result in the identification of rare variants, with larger effects.

Digging into the genetics underlying osteoarthritis. Genetic research in OA originally included mainly candidate gene association studies and genome wide linkage analyses. Especially the findings of the candidate gene studies appeared difficult to confirm in replication studies due to problems such as small sample size and the small numbers of polymorphic variants tested. The HapMap Project has shifted genetic epidemiology of complex inheritance away from linkage into genome wide association mapping of loci affecting disease and response to therapy. Although the resolution of genome wide association studies (GWAS) is limited, the use of single nucleotide polymorphisms (SNPs) instead of microsatellites which were used in linkage studies has substantially improved the resolution. Similar to genome wide linkage studies, GWA studies are hypothesis-free. However, GWAS require large sample sizes and are based on high throughput genotyping that typically ascertain at least 300,000 SNPs per sample nowadays. Before going on with the analysis of each of these variants for association with the trait of interest, genotyping is mostly followed by imputation of SNPs from the 1000 Genome Project (http://www.1000genomes.org/). The imputation methodology increases the genetic information per study subjects through the prediction of genotypes at genomic locations that have not actually been measured based on their known linkage disequilibrium (LD) with the genotyped SNPs.

A problem that scientists encounter in GWA studies is the so-called ‘missing heritability’. Although heritability estimates of OA are relatively high (39-60%), the common susceptibility loci identified thus far only explain a small part of the heritability. To some extend this could be explained by a possible overestimation of the heritability based on twin and family studies. On the other hand, it could suggest that many loci still remain to be identified. OA susceptibility loci identified by GWAS have small odds ratios. As explained by Manolio et al, this is due to a trade-off between higher risk allele frequencies and stronger genetic effects, typical for the identification of genetic variants predisposing to complex diseases. To overcome this problem and to increase power, researchers have tried to apply GWAS with ever increasing sample size: large scale genome wide initiatives were established to pile up substantial amounts of genome wide association data for meta-analyses such as the GWAS in arcOGEN and the joint-stratified meta-
analyses of GWAS performed within the Treat~OA consortium for knee[36] and for hip OA.[37] Together, the GWAS approach has resulted in the highest number of consistent OA susceptibility loci (Table 1A).

Next generation sequencing (NGS) has greatly augmented research possibilities and enhanced the pace of discovery of disease-associated genetic variants. Especially with the current increasing speed and decreasing costs, most likely NGS will shift its focus from sequencing of exomes to sequencing of the whole genome. The expectations that this may solve ‘the mysterious case of the missing heritability’ are high but to identify causal genetic variants by NGS, well-selected (severely affected) patients and their families are required as well as extensive computational tools to process the information. Although generalizability of results from mutations private to a particular family is uncertain, it is expected that more common variants in the same gene or in the same pathways in which the genes function will affect susceptibility in a larger population.[38] In OA this paradigm has been proven its validity for example for the $\text{SMAD3}$,[39, 40] $\text{COL11A1}$,[41-43] and $\text{ALDH1A2}$[44] genes.

Established osteoarthritis susceptibility loci. 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 1A and reviewed in references 31 and/or 45). As a comparison, in the first half of 2013 the number of loci known to predispose to other complex diseases such as type 2 diabetes and coronary heart disease were respectively 65[46] and 46.[47] For anthropometric traits such as body mass index (BMI), waist-hip ratio, and height the numbers of associated genetic loci identified in May 2013 were 36, 14, and 184, respectively.[48] Identified OA susceptibility loci resulted in a variety of compelling OA candidate genes, however, GWA studies also identified loci within gene deserts or within regions with multiple genes that have not yet been implicated in OA etiology.[35, 49] Multiple SNPs in a region can be in strong LD with each other (Table 1A), making it often challenging to determine the causal variant. This illustrates the typical problem in mapping complex traits and exposes the existing gap between genetic evidence and the molecular mechanism underlying the disease. In order to defeat these problems multiple molecular determinants (genetic, transcriptomic, proteomic, and epigenetic) will have to be integrated on a more regular base.[50, 51] Based on signals found to associate with hip OA Evangelou et al calculated that, together, the identified OA loci contribute only 2-3% to the heritability ($H^2$) of hip OA.[37] Using the same formula that was used by Evangelou and that was described by Park and colleagues,[52] the overall heritability is estimated to be around 13% based on the 17 independent OA susceptibility loci identified with GWAS (Table 1A below ‘Genome Wide Association Study’). This seemingly high percentage as compared to the $H^2$ of hip OA alone is mainly explained by a few loci with larger odds ratios such as rs12907038 annotated to the $\text{ALDH1A2}$ gene and associated with hand OA, and rs7639618 located within $\text{DVWA}$ or $\text{COL6A4P1}$ and associated with knee OA. However, as compared to the 39-60% estimated from family-based studies it suggests that still many of the loci involved in
Table 1. Established OA susceptibility loci (A) and pathways involved (B) (EA: effect allele; EAF: allele frequency of effect allele; OR: odds ratio; CI: confidence interval; F: female; M: male; FM: both female and male; FF: functional follow-up; *signal found in independent studies in Asians [53] and Europeans [54]; **SNPs in strong linkage disequilibrium; #SNPs in strong linkage disequilibrium; ‡SNPs in strong linkage disequilibrium; †signal decreased when adjusting for body mass index suggesting that effect on OA is applied via obesity, possibly targeting the iroquois homeobox 3 (IRX3) gene and not FTO [55]).

<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>p</th>
<th>Joint</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>
<td></td>
</tr>
<tr>
<td>Triple D repeat</td>
<td>9</td>
<td></td>
<td>Asporin</td>
<td>ASPN</td>
<td>2</td>
<td>1.4-2.8</td>
<td>6.60E-06</td>
<td>6.7E-03</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>Hip &amp; knee</td>
<td>FM</td>
<td>[39]</td>
<td>[61]</td>
</tr>
<tr>
<td>rs143385*</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>Hip</td>
<td>FM</td>
<td>[62]</td>
<td>[63-66]</td>
</tr>
<tr>
<td><strong>LINKAGE 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>
<td></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>Hip</td>
<td>F</td>
<td>[67]</td>
<td>[68-70]</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>Hip</td>
<td>F</td>
<td>[71]</td>
<td>[70-72, 73]</td>
</tr>
<tr>
<td><strong>GENOME WIDE ASSOCIATION 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>
<td></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>Hip</td>
<td>FM</td>
<td>[43, 74]</td>
<td>[75]</td>
</tr>
<tr>
<td>rs7639618</td>
<td>3</td>
<td>15216429</td>
<td>Double Von Wilkbrand 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>Knee</td>
<td>FM</td>
<td>[53]</td>
<td>[76]</td>
</tr>
<tr>
<td>rs11177**</td>
<td>3</td>
<td>52721305</td>
<td>3p21.1 locus: e.g. Guanine nucleotide binding protein-like 3</td>
<td>GNL3</td>
<td>A</td>
<td>0.38</td>
<td>1.12 (1.08-1.16)</td>
<td>1.3E-10</td>
<td>Hip &amp; knee</td>
<td>FM</td>
<td>[35]</td>
<td>[76]</td>
</tr>
<tr>
<td>rs6976**</td>
<td>3</td>
<td>52728804</td>
<td>3p21.1 locus: e.g. Glycosyltransferase 8 domain containing 1</td>
<td>GLT8D1</td>
<td>T</td>
<td>0.37</td>
<td>1.12 (1.08-1.16)</td>
<td>7.2E-11</td>
<td>Hip &amp; knee</td>
<td>FM</td>
<td>[35]</td>
<td>[76]</td>
</tr>
<tr>
<td>rs12107036</td>
<td>3</td>
<td>18960160</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>Knee</td>
<td>F</td>
<td>[35]</td>
<td>[76]</td>
</tr>
<tr>
<td>rs10947262#</td>
<td>6</td>
<td>32373312</td>
<td>Butyrophilin-like 2</td>
<td>BTNL2</td>
<td>C</td>
<td>0.58</td>
<td>1.31 (1.20-1.44)</td>
<td>5.1E-09</td>
<td>Knee</td>
<td>FM</td>
<td>[77]</td>
<td>[77]</td>
</tr>
<tr>
<td>rs7775228#</td>
<td>6</td>
<td>32658079</td>
<td>Major histocompatibility complex, class II, DQ beta 1</td>
<td>HLA-DQB1</td>
<td>T</td>
<td>0.62</td>
<td>1.34 (1.21-1.49)</td>
<td>2.4E-08</td>
<td>Knee</td>
<td>FM</td>
<td>[77]</td>
<td>[77]</td>
</tr>
<tr>
<td>rs10948172</td>
<td>6</td>
<td>44777691</td>
<td>Suppressor of Ty 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>Hip &amp; knee</td>
<td>M</td>
<td>[35]</td>
<td>[77]</td>
</tr>
<tr>
<td>rs9350591</td>
<td>6</td>
<td>7624152</td>
<td>Filamin A interacting protein 1 and SUMO1/sentrin specific peptidase 6</td>
<td>FILIP1; SENP6</td>
<td>T</td>
<td>0.11</td>
<td>1.18 (1.12-1.25)</td>
<td>2.4E-09</td>
<td>Hip</td>
<td>FM [35]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</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>rs3815148</td>
<td>7</td>
<td>67624152</td>
<td>7q22 locus: e.g. Component of oligomeric golgi complex 5 and HMG-box transcription factor 1</td>
<td>COG5; HBPI</td>
<td>C</td>
<td>0.23</td>
<td>1.14 (1.09-1.19)</td>
<td>8.0E-08</td>
<td>Knee &amp; hand</td>
<td>FM [49] [78]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs3815148‡</td>
<td>7</td>
<td>106938420</td>
<td>7q22 locus: e.g. Component of oligomeric golgi complex 5 and HMG-box transcription factor 1</td>
<td>COG5; HBPI</td>
<td>C</td>
<td>0.23</td>
<td>1.14 (1.09-1.19)</td>
<td>8.0E-08</td>
<td>Knee &amp; hand</td>
<td>FM [49] [78]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4730250</td>
<td>7</td>
<td>7624152</td>
<td>7q22 locus: e.g. Dihydouridine 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>Knee</td>
<td>FM [36]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4836732</td>
<td>9</td>
<td>119266695</td>
<td>Astrotactin 2 and Pregnancy-associated plasma protein A</td>
<td>ASTN2; KLHDC5; PPATHH</td>
<td>C</td>
<td>0.47</td>
<td>1.2 (1.13-1.27)</td>
<td>6.1E-10</td>
<td>Hip</td>
<td>F [35]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10492676</td>
<td>12</td>
<td>28014970</td>
<td>Kelch domain containing 5 and Parathyroid hormone-like hormone</td>
<td>KLHDC5; PTHLH</td>
<td>T</td>
<td>0.19</td>
<td>1.14 (1.09-1.20)</td>
<td>1.5E-08</td>
<td>Hip</td>
<td>FM [35]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs835487</td>
<td>12</td>
<td>107207695</td>
<td>7q22 locus: e.g. Component of oligomeric golgi complex 5 and HMG-box transcription factor 1</td>
<td>COG5; HBPI</td>
<td>C</td>
<td>0.23</td>
<td>1.14 (1.09-1.19)</td>
<td>8.0E-08</td>
<td>Knee</td>
<td>FM [36]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs11842874</td>
<td>13</td>
<td>113694509</td>
<td>MCF:2 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>Knee &amp; hip</td>
<td>FM [79]</td>
<td></td>
<td></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>Hand</td>
<td>FM [44]</td>
<td></td>
<td></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>Hip &amp; knee</td>
<td>F [35]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12982744</td>
<td>19</td>
<td>2177193</td>
<td>Dihydroxyacetone 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>Knee</td>
<td>FM [36]</td>
<td></td>
<td></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>NA</td>
<td>1.16 (1.11-1.22)</td>
<td>8.3E-09</td>
<td>Knee</td>
<td>FM [54]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6094710</td>
<td>20</td>
<td>46095649</td>
<td>Nuclear receptor coactivator 3</td>
<td>NCOA3</td>
<td>A</td>
<td>0.04</td>
<td>1.28 (1.18-1.39)</td>
<td>7.9E-09</td>
<td>Hip</td>
<td>FM [37]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0030326~embryonic limb morphogenesis</td>
<td>5</td>
<td>19.23</td>
<td>4.56E-06</td>
<td>1.19E-04</td>
<td>87</td>
<td>40.92</td>
<td>2.01E-03</td>
<td>2.01E-03</td>
<td>0.01</td>
</tr>
<tr>
<td>GO:0035113~embryonic appendage morphogenesis</td>
<td>5</td>
<td>19.23</td>
<td>4.56E-06</td>
<td>1.19E-04</td>
<td>87</td>
<td>40.92</td>
<td>2.01E-03</td>
<td>2.01E-03</td>
<td>0.01</td>
</tr>
<tr>
<td>GO:0035107~appendage morphogenesis</td>
<td>5</td>
<td>19.23</td>
<td>7.63E-06</td>
<td>1.98E-04</td>
<td>99</td>
<td>35.96</td>
<td>3.37E-03</td>
<td>1.69E-03</td>
<td>0.01</td>
</tr>
<tr>
<td>GO:0035108~limb morphogenesis</td>
<td>5</td>
<td>19.23</td>
<td>7.63E-06</td>
<td>1.98E-04</td>
<td>99</td>
<td>35.96</td>
<td>3.37E-03</td>
<td>1.69E-03</td>
<td>0.01</td>
</tr>
<tr>
<td>GO:0048736~appendage development</td>
<td>5</td>
<td>19.23</td>
<td>8.93E-06</td>
<td>2.32E-04</td>
<td>103</td>
<td>34.56</td>
<td>3.94E-03</td>
<td>1.32E-03</td>
<td>0.01</td>
</tr>
<tr>
<td>GO:0060173~limb development</td>
<td>5</td>
<td>19.23</td>
<td>8.93E-06</td>
<td>2.32E-04</td>
<td>103</td>
<td>34.56</td>
<td>3.94E-03</td>
<td>1.32E-03</td>
<td>0.01</td>
</tr>
<tr>
<td>GO:0048598~embryonic morphogenesis</td>
<td>6</td>
<td>23.08</td>
<td>3.92E-05</td>
<td>1.02E-03</td>
<td>307</td>
<td>13.92</td>
<td>1.72E-02</td>
<td>4.32E-03</td>
<td>0.06</td>
</tr>
</tbody>
</table>
development of OA remain to be identified.

Given the fact that 21 established susceptibility loci were identified, the number of publications reporting on their functional follow-up is highly disappointing. In fact, to the best of our knowledge as by August 2014 only 8 of the loci have been characterized in more detail following their identification (Table 1A). Some of the studies were aimed at determining the particular gene affected by the genetic variation.[75, 80] Other studies have determined the function of the gene expected to be targeted by the SNP,[37, 44, 81] and several studies intended to determine the mechanism underlying OA susceptibility via the identified locus: either at the molecular level or at an organismal level.

With respect to the latter, in particular Growth differentiation factor 5 (GDF5[63-66]), Frizzled motif associated with bone development (FRZB[68-70]), and Deiodenase iodothyronine type 2 (DIO2[70, 72, 73]) have received full attention.

Pathway analysis of the genes annotated to the OA susceptibility loci with the online Database for Annotation, Visualization and Integrated Discovery tool (DAVID[82]) shows enrichment for genes involved in limb development (Table 1B). Since several of the identified OA susceptibility genes encode for proteins involved in endochondral ossification during early development such as DIO2, FRZB, and GDF5 (see Figure 3), OA has been hypothesized to be an early developmental disease with late-life onset.[24] The developmental origin of OA is twofold: First, genes involved in endochondral ossification determine both composition of the bone and cartilage and the shape of the joint. While a suboptimal joint shape will make the tissues more prone to damage and thus to development of OA,[26] bone and cartilage strength determine the capacity to persist mechanical stresses. Secondly, the maturational arrested phenotype of articular chondrocytes is changed during OA and chondrocytes retrieve characteristics from hypertrophic growth plate chondrocytes.[83] Given this hypothesis, and thus the need to investigate early as well as late effects of genetic variation, the best would be to apply in vivo animal models in which developmental effects as well as late life effects can be studied, and where mechanical stresses can be applied in a controlled manner for example by using a running model. Alternatively, experiments can be performed in vitro to study genetic effects on chondrogenic potential of mesenchymal stem cells and matrix deposition of primary chondrocytes.

**Figure 3 OA susceptibility genes involved in endochondral ossification.**

**BIOBANKING IN OA RESEARCH**

To achieve the aims of the work presented in this thesis and in general for OA research, the assembly of a biobank with relevant tissues stored tailored to isolate RNA and DNA, possibly also to make protein lysates
for analysis by Western blotting, and to apply immunohistochemistry is required. In addition, the collection of relevant cell types (mesenchymal stem cells, chondrocytes, osteoblasts) would facilitate the setup of experimental cell models. Together, this will allow analyses at multiple molecular levels.

Human genome wide gene expression datasets are available for OA cartilage from non-OA and OA-affected joints\[84, 85\] as well as from unaffected and affected cartilage areas of the same joints\[86-88\], OA affected and unaffected subchondral bone\[89\] and OA synovial membrane\[90\]. However, these are collected in different studies. Ideally, the different datasets should be derived from the same study group which could provide detailed information on the dynamics of gene expression profiles during OA pathophysiological processes and also include expression profiles of blood which would permit detection of biomarkers in the blood that reflect ongoing OA. Besides gene expression, RNA can also be used to study differential allelic expression of genetic variants. The isolated DNA can be used for genotyping and to perform whole genome or exome sequencing and to establish epigenetic patterns of the different tissues.

For the studies performed in this thesis we used biomaterials of the ongoing RAAK study (Research Arthritis and Articular Cartilage), in which joint materials (cartilage of healthy and diseased (areas of the) joint, bone, and where available ligaments), mesenchymal stem cells (hip joints only), and primary chondrocytes of controls and patients undergoing joint replacement surgery in the Leiden University Medical Center and collaborating outpatient clinics in the Leiden area are collected. In addition to the RAAK study, studies were performed in the Genetics osteoArthritis and Progression (GARP) study. This study consists of sibling pairs with symptomatic and radiographic OA in at least two joints\[91\] and is tailored for molecular epidemiological research. Of all participants a complete collection of biomaterials such as DNA, RNA (blood), serum, plasma, and urine for biomarker detection is available in addition to extended clinical (X-rays, MRI) and demographic data in a follow-up design at 2 and 5 years. In parallel (including follow-up) data on 30 age-matched controls without OA in the indicated joints has been collected, the so-called NORREF study.

To summarize, genetic studies have identified many OA susceptibility loci but most of the heritability is still missing. Several issues warrant further investigation. In the first place, treatment of OA is substantially impaired by its late diagnosis, and therapies would significantly benefit from the availability of biomarkers that can detect the disease at an early stage, monitor disease progression, and distinguish between different OA subtypes. Secondly, further insight into genetic factors causal to OA will increase insight into the etiology of OA. With respect to GWA studies, more severe selection of samples to decrease heterogeneity is required. Current availability of next generation sequencing is likely to contribute to the identification of genetic variants predisposing to OA. However, this will require cutting-edge study design and careful selection of samples. Thirdly, functional genomic research to generate biological understanding of genetic susceptibility beyond that of knowing the genomic location is lacking way behind the mapping of OA loci.
Chapter 1

RATIONALE AND OUTLINE OF THIS THESIS

In this thesis, we set out to characterize in detail the ongoing pathophysiological processes in arthritic cartilage within the RAAK study (chapter 2). Genome wide gene expression profiles were generated from OA affected and macroscopically intact cartilage of the same joint and gene enrichment analysis was performed to study the major pathways involved. Also genome wide expression profiles were determined from blood of patients with generalized OA from the GARP study and compared to those of healthy controls (chapter 3). Results were used to identify genes that may serve as biomarkers for OA, and overlap between both databases (cartilage and blood) was investigated in order to learn to what extent ongoing OA processes may be reflected in the blood.

Although selection of biomarkers in OA is promising, the distinction between individual patients and controls is not yet possible due to relatively high inter-individual variation. Variation in basal levels of biomarkers may originate from genetic influences, and taking this variation into account by stratification for genotype or by using other biomarkers for persons with a specific genotype could improve the informativity of the biomarkers on OA onset and progression detection. Therefore, in chapter 4 we aimed at the identification of genetic loci affecting levels of the promising biomarkers serum COMP and urinary CTX-II. Subsequently, we explored a meta-analysis of GWAS to determine whether any of the identified loci were associated to OA.

To identify genetic loci associated with OA, 2 different approaches were taken. As an alternative to the ever increasing number of samples used in GWAS we investigated whether selection of OA patients with familial and relatively severe phenotype could be an alternative (chapter 5). This was done by investigating the GARP study, in which siblings were selected based on their relatively severe generalized OA phenotype. Furthermore, exome sequencing was done for 2 members of a family characterized by early-onset OA (chapter 6). By applying exome sequencing we aimed to identify a genetic variant with large effect that could hint towards general pathways involved in OA. Finally, in chapter 7, a 3-dimensional in vitro chondrogenesis method was set up that can be applied in functional genomic research. Here, it was used to functionally characterize the effects of variation in thyroid signaling during chondrogenesis as a model for variation in the OA susceptibility gene DIO2.

REFERENCES

Introduction


46 Yaghootkar H, Frayling TM. Recent progress in the use of genetics to understand links between type 2 diabetes and related metabolic traits. *Genome Biol* 2013;14:203.

Introduction


51 Westra HJ, Franke L. From genome to function by studying eQTLs. Biochim Biophys Acta 2014.


Chapter 1

One 2014;9:e86590.

66 Reynard LN, Bui C, Syddall CM, Loughlin J. CpG methylation regulates allelic expression of GDF5 by modulating binding of SP1 and SP3 repressor proteins to the osteoarthritis susceptibility SNP rs143383. Hum Genet 2014.


studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. Proc Natl Acad Sci U S A 2012;109:8218-23.


