On the relation between genetic variation and osteoarthritis

Hollander, W. den

ON THE RELATION BETWEEN

GENETIC VARIATION AND OSTEOARTHRITIS

Wouter den Hollander, MSc

On the relation between genetic variation and osteoarthritis

Wouter den Hollander, MSc

This research was financially supported by the Reumafonds under grant agreement 10-1-402 and the Leids Universitair Medisch Centrum.

ISBN: 978-94-6233-922-4

Designed, copy-edited & printed by Gildeprint.

ON THE RELATION BETWEEN GENETIC VARIATION AND OSTEOARTHRITIS

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 29 maart 2018
klokke 11.15 uur

door

Wouter den Hollander
geboren te Noordwijk
in 1987

Promotores

Prof. dr. I. Meulenbelt

Prof. dr. P.E. Slagboom

Co-promotor

Dr. Y.F.M. Ramos

Leden promotiecommissie

Prof. dr. R.G.H.H. Nelissen

Dr. L.N. Reynard

Musculoskeletal Research Group, Newcastle University

Prof. dr. ir. M.J.T. Reinders

Pattern Recognition and Bioinformatics, Technische Universiteit Delft

Leiden Computational Biology Center, Leiden University Medical Center

Dr. J.J.B. van Meurs

Department of Internal Medicine, Erasmus Medical Center

CONTENTS:

1

GENERAL INTRODUCTION

OSTEOARTHRITIS

Osteoarthritis (OA) is a prevalent, degenerative muscoskeletal disease that affects all articular joints, although it is most prominent among hip, knee and the phalangeal joints (1). In 2011 it was estimated that over 1,1 million people in the Netherlands (7%) were visiting a general practitioner due to OA related complaints (2). While present across the entire population, OA predominantly affects the elderly, which is reflected by a disease prevalence under 65 years of age of 3.6% and 3.9% for men and women, whereas above 65 years of age these numbers increase to 20.9% and 36.2%, respectively (**Figure 1**) (2). Hence, in light of the increasing proportion of elderly in western civilization, OA's burden on our society is advancing and will likely keep doing so in the near future.

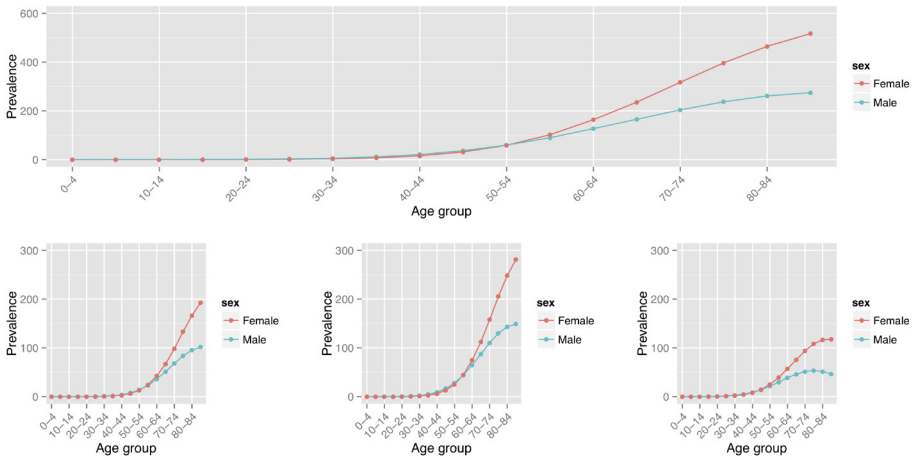


Figure 1. Estimated osteoarthritis prevalence per 1000 inhabitants among the Dutch population in 2011, stratified by gender. (**A**) Overall prevalence of osteoarthritis, measured in all joints. (**B**) Prevalence of hip osteoarthritis. (**C**) Prevalence of knee osteoarthritis. (**D**) Prevalence of hand osteoarthritis. Data from Nationaal Kompas Volksgezondheid 2011 (2).

The disease is principally portrayed by gradual degeneration of cartilage in articular joints, although, in recent years it has become apparent that multiple joint tissues, such as the subchondral bone and synovial membrane, are involved as well (3). Multiple risk factors have been described, and encompass, but are not limited to age, gender (4), body mass index (BMI) (5), joint injuries (6) and genetic predisposition with heritability estimates ranging from 40% to 60% (7). It manifests itself as stiff and painful joints due to joint space narrowing (8), calcified articular cartilage (9), formation of bony outgrowths called osteophytes (10) and remodelling of subchondral bone (11). Presently, no treatment other than pain relief exists and patients are ultimately required to undergo total joint replacement (TJR) surgery to guarantee proper functioning of the joint. While TJR is an effective treatment for end-stage patients, gaining a deeper understanding of the actual biological mechanisms that underlie and/or accompany OA pathophysiology will arguably aid future efforts in developing disease modifying treatments.

Although multiple joint tissues have been implicated in OA pathophysiology (12–14), articular cartilage is allegedly the pivotal tissue affected by the disease. While wear and tear plays a substantial role in cartilage breakdown and thus OA progression, degradation and calcification of the tissue is additionally actively mediated by chondrocytes (15). Chondrocytes, which are thought of to be the single cell type residing in articular cartilage (16), are responsible for maintenance of the extracellular matrix (ECM) and do so by actively breaking down, remodelling and repairing the ECM using a wide variety of both catabolic as well as anabolic proteins and enzymes when they are required to do so (17–21). In order to do so properly throughout life, it is crucial for chondrocytes that after tissue homeostasis is restored upon physical stresses and consequent microtrauma's, they return to their resting state to await reactivation when required. However, chondrocytes residing in OA affected articular cartilage seem to have escaped this perpetual but appropriate cycle of reactivation and resting (15,21,22). Specifically, as OA chondrocytes dedifferentiate, start dividing, form columnar structures and actively start calcifying the ECM. They appear to have lost their articular phenotype and have entered a process that, interestingly, resembles endochondral ossification during fetal development (15,21–26).

OSTEOARTHRITIS ASSOCIATED GENETIC VARIATION

Due to the substantial genetic component of OA, marked by high heritability estimates (7), in recent years, significant efforts have been made to elucidate the complex genetic architecture of the disease (27–46). However, the search for putative genetic variation that might predispose for or protect against the disease has emerged as a considerable challenge. The hereditary nature of OA is therefor still incompletely understood. Association analysis of genetic variation, traditionally performed among candidate genes and more recently by linkage analysis and genome wide association (GWA) studies, has identified a multitude of loci that are involved in OA pathophysiology. Specifically, the presence of common point mutations or single nucleotide polymorphisms (SNPs) in genes such as *GDF5* (31,39,41–43), *FRZB* (39,44,45), *ALDH1A2* (36), *DOT1L* (40,46), *GNL3* (47) and *DIO2* (32) has been associated with in- or decreased prevalence of OA among carriers. As a result, OA is designated a so-called genetically complex disease, i.e. there is no single, common genetic variant that determines whether respective carriers develop generalized OA (48–50). In that regard, uncommon genetic variants, often segregating by Mendelian nature within affected families, have been linked directly with the development of specific OA subtypes (e.g. chondrocalcinosic OA caused by a read-through mutation in the *TNFRSF11B* gene (51)). Whereas these monogenetic forms of OA are generally due to high-impact mutations, the associated effect sizes of SNPs associated with generalized forms of OA are relatively small (50), and as such leading to missed heritability for OA in general.

The nature by which these relatively small genetic effects arise seems to be due modest downstream effects of the associated genes. While these OA associated genetic variations rarely induce structural differences in the resulting protein product, alleles of common OA susceptibility SNPs do frequently mark transcriptional differences of genes in close proximity. Traditionally, SNPs that affect gene expression levels have been addressed by so-called expression quantitative trait locus (eQTL) analysis (52–54), which identified a magnitude of SNPs that affect gene expression of proximal genes marked by allele correlated expression levels. Albeit successful in doing so, eQTL studies require substantial sample sizes to reach statistical significance. In that regard, it has been shown that assessing the extent of imbalance of expressed alleles within heterozygous individuals was shown to be an effective alternative to identify genetic variation that affects gene expression levels using significantly less samples (36,42,55–59). Imbalanced expression of alleles of heterozygous carriers, generally referred to as an allelic imbalance (AI), allele specific expression (ASE) and/or allelic expression (AE), occurs when two alleles of a certain SNP are not expressed to equal extent. Because eQTLs frequently are tissue specific (60–62) and the fact that articular cartilage biobanks are generally of limited size, measuring AI of OA susceptibility SNPs in articular cartilage has gained increasing interest in recent years (36,42,55–59). Multiple common SNP alleles associated with OA have been reported to mark AI in articular cartilage and as such provided mechanistic insight into the putative biological mechanism that underlies the statistical genetic association. For example, it was shown that the risk allele C of rs225014, located in exon 3 of OA susceptibility gene *DIO2*, is expressed to higher extent compared to the wild type allele T in articular cartilage of heterozygous carriers of the SNP (57). Similar results have been reported for other robust OA susceptibility genes such as *GDF5* (42,56) and *ALDH1A2* (36), as well as for genes proximal to SNPs of which it is currently not completely certain which genes are respectively affected (*SPCS1* (55), *GNL3* (55), *COL11A1* (58)).

Nonetheless, as a consequence of these relative small effects governed by generalized OA associated SNPs, attempts to reduce missing heritability for OA is to some extent obstructed in even the largest GWA studies. In parallel, generalized OA arguably comprises a spectrum of subtypes (e.g. affected joint, osteophyte formation and/or presence of joint space narrowing) which might have just partly overlapping genetic determinants. As subtypes are generally combined in GWA studies, true association signals will be diluted into the noisy background. And although testing millions of alleles for association with larger sample sizes and/or deeper phenotypes will presumably yield novel OA associated loci, another worthwhile approach might be reducing the vast number of statistical tests, and thus lowering the multiple testing correction penalty. Selecting for SNPs that are more likely to confer susceptibility, for example by exclusion of SNPs that do not reside in proximity of genes that are involved in articular joint tissue homeostasis, might prove beneficial in this sense.

DISEASE ASSOCIATED GENE EXPRESSION LEVELS

A well described hallmark of ongoing OA is the substantial amount of differentially expressed genes in OA affected cartilage when compared to control cartilage (18,63–72). And while identification of these has traditionally been performed in a gene targeted fashion, advances in transcriptome wide measurements, such as microarray and sequencing technologies, have markedly extended the number of genes that seem to partake in OA pathophysiology. For example, genes involved in extra cellular matrix (ECM) anabolism (e.g. *COL11*, *AGC*, *CILP*, *PRG4*) (17,18), catabolism (e.g. *MMP3*, *MMP9*, *ADAMTS5*) (73), inflammation (e.g. *IL11*, *IL1*, *TREM1*) (73) and development (e.g. *DKK3*, *DIO2*, *GDF5*, *FRZB*) (73,74) have repeatedly been observed to be expressed at different levels in OA affected cartilage. Dynamic expression and subsequent silencing of these genes is arguably required during normal life to cope with everyday physical stresses, consequent microtraumas and to maintain articular cartilage homeostasis in general. Seen with OA pathophysiology, however, it seems as if gene transcription is no longer balanced and tends to favour cartilage catabolism, consequently leading to gradual degradation and calcification of the tissue. In his regard, chondrocytes residing in OA affected articular cartilage appear to be constituently active and no longer able to return to their proposed resting state. Worth mentioning in this context, are the apparent differences between gene expression profiles of distinct joints. Nonetheless, irrespective of the joint of interest, from a pathway point of view OA associated expression profiles reflect the aforementioned shift towards catabolism when compared to unaffected cartilage.

Nevertheless, despite the valuable observations that have resulted from (semi) transcriptome wide profiling of OA affected articular cartilage, it is at the current time not possible to designate the apparent transcriptomic disbalance as either pathologically causal and putatively driving the disease, or consequential and merely constituting a biomarker. While it is eminent that changes in gene expression are compulsory to effectuate degradative remodelling of the ECM, it can be expected that the resulting consequences will in turn affect gene transcription rates in the respective chondrocytes. In human OA research this proposed feedback mechanism opposes a challenge when interpreting results generated from transcriptomic experiments, whereas it is not encountered as such in the context of genetic association analyses. In an attempt to bridge this gap (and while doing so refrain from complex clinical translation of animal and *in-vitro* experiments), studying the underlying mechanisms that might regulate gene expression in human articular cartilage *in-vivo* might just help us in understanding the complexity at hand.

EPIGENETICS IN OSTEOARTHRITIS

Epigenetic mechanisms are traditionally described as heritable modifications of the genome without affecting the actual nucleotide sequence (75,76), although in recent years they have emerged as cellular mechanisms that serve as dynamic regulators of gene expression (77–79). Nonetheless, despite the ambiguous inheritance of epigenetic modifications, its strong correlation with the proximal nucleotide sequence, gene expression levels as well as tissue specificity (80) make it highly applicable for OA research. While a broad spectrum of epigenetic modifiers and levels exist, such as histone modifications (81) and small non-coding RNA expression (82–84), DNA methylation appears studied most intensively (68,85–89). Addition of a methyl group to the 5th carbon atom of cytosines in cytosine-guanine residues (CpGs) within the genomic DNA is known to correlate with gene expression levels, likely due to interference with binding of DNA binding proteins (DBPs) to the genomic DNA (79). Although gene expression is regulated by a vast number of mechanisms, binding of DBPs such as transcription factors is pivotal herein. Additionally, DNA methylation levels at specific positions along the genome are known to be highly tissue specific, partly due to mitotic inheritance of the respective somatic layers during development, but presumably also to maintain cellular differentiation in adult tissues (80). Hence, in light of the strong relation with gene expression as well as constituting distinct tissue profiles, studying DNA methylation seems an appealing endeavour to further dissect the molecular genetic facets of OA (3,90).

Early studies assessing DNA methylation in OA affected articular cartilage have revealed that DNA methylation levels of specific CpGs in or near known OA associated genes often reflect the disease status of the respective tissue. For example, DNA methylation levels of CpGs near multiple matrix metalloproteinases (*MMP3*, *MMP9* and *MMP13*) (91–93), *GDF5* (86,87), *SOX9* (94,95), *IL1* (93,96), *NOS2* (97) and *COL9* (98) were shown to be significantly different between OA affected and unaffected articular cartilage. Notably, these differences in methylation levels coincided with a significant difference in expression levels of the respective genes as well. More recently, multiple studies have reported on (semi-)genome wide DNA methylation profiles of articular cartilage in the context of OA, in part due to the development of affordable genome wide DNA methylation arrays (85,88). By comparing DNA methylation at the methylome level between OA affected and unaffected articular cartilage, numerous differentially methylated CpGs have been reported and confirmed (85,88). However, it is presently still unclear if, and to what extent these CpGs correlate with gene expression levels in articular cartilage. While it was generally accepted that increased methylation or hypermethylation marks decreased transcription of respective genes, an increasing number of CpGs have been reported of which hypermethylation correlates with increased expression. Interestingly, the canonical view of transcriptional downregulation alongside hypermethylation is often observed for CpGs that reside in transcriptional promotors or enhancers, whereas CpGs that do so in opposite

direction are generally located between the transcription start and ending sites of genes (77,79,81). Consequently, given that interpreting cross-sectional DNA methylation differences is challenging on its own, doing so without simultaneous and quantitative knowledge of the transcriptome appears to considerably impede our ability to do so accurately.

SUPERIMPOSING CAUSALITY AND/OR DIRECTIONAL EFFECT

In light of the descriptive nature of the aforementioned transcriptomic and methylomic experiments (i.e. cross-sectional data points), we are inherently refrained from implying causality. This statistical impediment affects *in-vivo* human OA research in multiple ways. Firstly, when we study OA transcriptomic and/or methylomic profiles as described earlier, we cannot differentiate between drivers and markers of the ongoing pathophysiological processes. Secondly, if our interest goes out to understanding the relationship between DNA methylation and gene expression in articular cartilage, we encounter a proposed regulatory feedback loop that prevents us from stating a directional relationship between the two. Within these boundaries, however, identification of CpGs for which their methylation levels quantitatively correlate both with the presence of particular alleles as well as with proximal gene expression, would be indicative of a regulatory relation that propagates from genome to methylome to transcriptome.

AIM AND CONTENT OF THIS THESIS

The here presented thesis aims to address some of the challenges and concerns outlined in the previous sections. As such, the following chapters will elaborate on a number of specific scientific challenges to which the OA research field is currently opposed to.

As candidate gene approaches have resulted in the successful identification of multiple OA predisposing SNPs that exert their disease association through AI in articular cartilage, we have assessed AI on a genome wide scale in articular cartilage. By simultaneously utilizing both transcript sequence as well as abundancy information derived in RNA sequencing data generated from human articular cartilage, **chapter 2** describes a transcriptome wide approach to detect novel SNPs that mark AI in articular cartilage. Subsequently in **chapter 3**, we reveal that the generated data and results in chapter 2 can be used to support ongoing OA GWA studies.

Multiple reports on OA susceptibility SNPs that mark imbalanced expression in articular cartilage of their respective genes have been published, as our group has done so for rs225014 (*DIO2*). Given these observations, we were interested in the regulatory mechanisms that might underlie the reported AI and investigated the role of epigenetic regulation of *DIO2* expression by DNA methylation in knee and hip articular cartilage. **Chapter 4** describes the relation between the alleles of rs225014,

proximal DNA methylation and expression of *DIO2* in articular cartilage. Given the insightful results obtained in chapter 4, we next addressed DNA methylation in articular on a semi-methylome wide scale. In **chapter 5** we integrate methylomic and transcriptomic data derived from preserved and paired OA lesioned articular cartilage and discuss the associations in the context of the genetic background. As part of quality control we observed, to some extent by serendipity, distinct clustering of samples upon dimension reduction of the DNA methylation data described in chapter 5, independent of gender and affection status. **Chapter 6** expands hereon and describes apparent joint specific epigenomic profiles. Consequently, motivated by both the increasing body of literature on DNA methylation research in OA, as well as our own efforts, **chapter 7** summarizes the current status and the putative future perspectives of DNA methylation research in OA.

REFERENCES

1. Ray MB. Osteoarthritis. PostgradMedJ. 1937;13 (143):311–20.
2. Nationaal Kompas Volksgezondheid. 2011.
3. Tsezou A. Osteoarthritis year in review 2014: genetics and genomics. OsteoarthritisCartilage. 2014 Dec;22(1522–9653):2017–24.
4. Holla JF, Steultjens MP, van der Leeden M, Roorda LD, Bierma-Zeinstra SM, den Broeder AA, et al. Determinants of range of joint motion in patients with early symptomatic osteoarthritis of the hip and/or knee: an exploratory study in the CHECK cohort. OsteoarthritisCartilage. 2011 Apr;19(1522–9653):411–9.
5. Elliott KS, Chapman K, Day-Williams A, Panoutsopoulou K, Southam L, Lindgren CM, et al. Evaluation of the genetic overlap between osteoarthritis with body mass index and height using genome-wide association scan data. AnnRheumDis. 2013 Jun;72(1468–2060):935–41.
6. Schenker ML, Mauck RL, Ahn J, Mehta S. Pathogenesis and prevention of posttraumatic osteoarthritis after intra-articular fracture. JAmAcadOrthopSurg. 2014 Jan;22(1067–151X (Print)):20–8.
7. Valdes AM, Spector TD. The genetic epidemiology of osteoarthritis. CurrOpinRheumatol. 2010 Mar;22(1531–6963):139–43.
8. Benichou OD, Hunter DJ, Nelson DR, Guermazi A, Eckstein F, Kwoh K, et al. One-year change in radiographic joint space width in patients with unilateral joint space narrowing: data from the Osteoarthritis Initiative. Arthritis Care Res(Hoboken). 2010 Jul;62(2151–4658):924–31.
9. Fuerst M, Bertrand J, Lammers L, Dreier R, Echtermeyer F, Nitschke Y, et al. Calcification of articular cartilage in human osteoarthritis. Arthritis Rheum. 2009 Sep;60(0004–3591 (Print)):2694–703.
10. Damen J, Schiphof D, Wolde ST, Cats HA, Bierma-Zeinstra SM, Oei EH. Inter-observer reliability for radiographic assessment of early osteoarthritis features: the CHECK (cohort hip and cohort knee) study. OsteoarthritisCartilage. 2014 Jul;22(1522–9653):969–74.
11. Nevitt MC, Zhang Y, Javaid MK, Neogi T, Curtis JR, Niu J, et al. High systemic bone mineral density increases the risk of incident knee OA and joint space narrowing, but not radiographic progression of existing knee OA: the MOST study. AnnRheumDis. 2010 Jan;69(1468–2060):163–8.
12. Man GS, Mologhianu G. Osteoarthritis pathogenesis - a complex process that involves the entire joint. J Med Life . 2014;7 (1):37–41.
13. Beekhuizen M, Bastiaansen-Jenniskens YM, Koevoet W, Saris DB, Dhert WJ, Creemers LB, et al. Osteoarthritic synovial tissue inhibition of proteoglycan production in human osteoarthritic knee cartilage: establishment and characterization of a long-term cartilage-synovium coculture. Arthritis Rheum. 2011 Jul;63(1529–0131):1918–27.
14. Yuan XL, Meng HY, Wang YC, Peng J, Guo QY, Wang AY, et al. Bone-cartilage interface crosstalk in osteoarthritis: potential pathways and future therapeutic strategies. OsteoarthritisCartilage. 2014 Aug;22(1522–9653):1077–89.
15. Sun MM, Beier F. Chondrocyte hypertrophy in skeletal development, growth, and disease. Birth Defects ResCEmbryoToday. 2014 Mar;102(1542–9768):74–82.
16. Sophia Fox AJ, Bedi A, Rodeo SA. The Basic Science of Articular Cartilage: Structure, Composition, and Function. Sports Health . 2009 Nov;1 (6):461–8.
17. Aigner T, Fundel K, Saas J, Gebhard PM, Haag J, Weiss T, et al. Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis. Arthritis Rheum. 2006 Nov;54(0004–3591 (Print)):3533–44.
18. Peffers M, Liu X, Clegg P. Transcriptomic signatures in cartilage ageing. Arthritis ResTher. 2013;15(1478–6362):R98.
19. Goldring MB, Marcu KB. Cartilage homeostasis in health and rheumatic diseases. Arthritis ResTher. 2009;11(1478–6362):224.
20. Ratnayake M, Ploger F, Santibanez-Koref M, Loughlin J. Human chondrocytes respond discordantly to the protein encoded by the osteoarthritis susceptibility gene GDF5. PLoSOne. 2014;9(1932–6203):e86590.
21. Thomas CM, Whittles CE, Fuller CJ, Sharif M. Variations in chondrocyte apoptosis may explain the increased prevalence of osteoarthritis in some joints. RheumatolInt. 2011 Oct;31(1437–160X):1341–8.
22. Goldring MB. Chondrogenesis, chondrocyte differentiation, and articular cartilage metabolism in health and osteoarthritis. TherAdvMusculoskeletDis. 2012 Aug;4(1759–7218):269–85.

23. Saito T, Fukai A, Mabuchi A, Ikeda T, Yano F, Ohba S, et al. Transcriptional regulation of endochondral ossification by HIF-2alpha during skeletal growth and osteoarthritis development. NatMed. 2010 Jun;16(1546–170X):678–86.
24. Meulenbelt I, Bos SD, Chapman K, van der Breggen R, Houwing-Duistermaat JJ, Kremer D, et al. Meta-analyses of genes modulating intracellular T3 bio-availability reveal a possible role for the DIO3 gene in osteoarthritis susceptibility. AnnRheumDis. 2011 Jan;70(1468–2060):164–7.
25. Shao YY, Wang L, Ballock RT. Thyroid hormone and the growth plate. RevEndocrMetab Disord. 2006 Dec;7(1389–9155 (Print)):265–71.
26. Mueller MB, Tuan RS. Functional characterization of hypertrophy in chondrogenesis of human mesenchymal stem cells. Arthritis Rheum. 2008 May;58(0004–3591 (Print)):1377–88.
27. Xiong DH, Liu XG, Guo YF, Tan LJ, Wang L, Sha BY, et al. Genome-wide association and follow-up replication studies identified ADAMTS18 and TGFBR3 as bone mass candidate genes in different ethnic groups. AmJHumGenet. 2009 Mar;84(1537–6605):388–98.
28. Panoutsopoulou K, Southam L, Elliott KS, Wrayner N, Zhai G, Beazley C, et al. Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study. AnnRheumDis. 2011 May;70(1468–2060):864–7.
29. Yerges-Armstrong LM, Yau MS, Liu Y, Krishnan S, Renner JB, Eaton CB, et al. Association analysis of BMD-associated SNPs with knee osteoarthritis. JBone MinerRes. 2014 Jun;29(1523–4681):1373–9.
30. Evans DS, Cailotto F, Parimi N, Valdes AM, Castano-Betancourt MC, Liu Y, et al. Genome-wide association and functional studies identify a role for IGFBP3 in hip osteoarthritis. AnnRheumDis. 2014 Jun 13;(1468–2060).
31. Valdes AM, Evangelou E, Kerkhof HJ, Tamm A, Doherty SA, Kisand K, et al. The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. AnnRheumDis. 2011 May;70(1468–2060):873–5.
32. Meulenbelt I, Min JL, Bos S, Riyazi N, Houwing-Duistermaat JJ, van der Wijk HJ, et al. Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. HumMolGenet. 2008 Jun 15;17(1460–2083):1867–75.
33. Meulenbelt I, Bijkerk C, Breedveld FC, Slagboom PE. Genetic linkage analysis of 14 candidate gene loci in a family with autosomal dominant osteoarthritis without dysplasia. JMedGenet. 1997 Dec;34(0022–2593 (Print)):1024–7.
34. Meulenbelt I, Chapman K, Dieguez-Gonzalez R, Shi D, Tsezou A, Dai J, et al. Large replication study and meta-analyses of DVWA as an osteoarthritis susceptibility locus in European and Asian populations. HumMolGenet. 2009 Apr 15;18(1460–2083):1518–23.
35. Evangelou E, Kerkhof HJ, Styrkarsdottir U, Ntzani EE, Bos SD, Esko T, et al. A meta-analysis of genome-wide association studies identifies novel variants associated with osteoarthritis of the hip. AnnRheumDis. 2013 Sep 4;(1468–2060).
36. Styrkarsdottir U, Thorleifsson G, Helgadottir HT, Bomer N, Metrustry S, Bierma-Zeinstra S, et al. Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene and with rare variants at 1p31. NatGenet. 2014 May;46(1546–1718):498–502.
37. Day-Williams AG, Southam L, Panoutsopoulou K, Rayner NW, Esko T, Estrada K, et al. A Variant in MCF2L Is Associated with Osteoarthritis. AmJHumGenet. 2011 Sep 9;89(1537–6605):446–50.
38. Kerkhof HJ, Lories RJ, Meulenbelt I, Jonsdottir I, Valdes AM, Arp P, et al. A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. Arthritis Rheum. 2010 Feb;62(0004–3591 (Print)):499–510.
39. Evangelou E, Chapman K, Meulenbelt I, Karassa FB, Loughlin J, Carr A, et al. Large-scale analysis of association between GDF5 and FRZB variants and osteoarthritis of the hip, knee, and hand. Arthritis Rheum. 2009 Jun;60(0004–3591 (Print)):1710–21.
40. Evangelou E, Valdes AM, Castano-Betancourt MC, Doherty M, Doherty S, Esko T, et al. The DOT1L rs12982744 polymorphism is associated with osteoarthritis of the hip with genome-wide statistical significance in males. AnnRheumDis. 2013 Jul;72(1468–2060):1264–5.
41. Valdes AM, Spector TD, Doherty S, Wheeler M, Hart DJ, Doherty M. Association of the DVWA and GDF5 polymorphisms with osteoarthritis in UK populations. AnnRheumDis. 2009 Dec;68(1468–2060):1916–20.
42. Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, Saito S, et al. A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. NatGenet. 2007 Apr;39(1061–4036 (Print)):529–33.
43. Chapman K, Takahashi A, Meulenbelt I, Watson C, Rodriguez-Lopez J, Egli R, et al. A meta-analysis of Euro-

pean and Asian cohorts reveals a global role of a functional SNP in the 5' UTR of GDF5 with osteoarthritis susceptibility. HumMolGenet. 2008 May 15;17(1460–2083):1497–504.
44. Valdes AM, Loughlin J, Oene M V, Chapman K, Surdulescu GL, Doherty M, et al. Sex and ethnic differences in the association of ASPN, CALM1, COL2A1, COMP, and FRZB with genetic susceptibility to osteoarthritis of the knee. Arthritis Rheum. 2007 Jan;56(0004–3591 (Print)):137–46.
45. Loughlin J, Dowling B, Chapman K, Marcelline L, Mustafa Z, Southam L, et al. Functional variants within the secreted frizzled-related protein 3 gene are associated with hip osteoarthritis in females. ProcNatlAcad-SciUSA. 2004 Jun 29;101(0027–8424 (Print)):9757–62.
46. Castano Betancourt MC, Cailotto F, Kerkhof HJ, Cornelis FM, Doherty SA, Hart DJ, et al. Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. ProcNatlAcadSciUSA. 2012 May 22;109(1091–6490):8218–23.
47. Zeggini E, Panoutsopoulou K, Southam L, Rayner NW, Day-Williams AG, Lopes MC, et al. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. Lancet. 2012 Sep 1;380(1474–547X):815–23.
48. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature . 2009;461 (7265):747–53.
49. Jordan JM, Kraus VB, Hochberg MC. Genetics of osteoarthritis. CurrRheumatolRep. 2004 Feb;6(1523–3774 (Print)):7–13.
50. Kloppenburg M, Kwok WY. Hand osteoarthritis--a heterogeneous disorder. NatRevRheumatol. 2012 Jan;8(1759–4804):22–31.
51. Ramos YF, Bos SD, van der Breggen R, Kloppenburg M, Ye K, Lameijer EW, et al. A gain of function mutation in TNFRSF11B encoding osteoprotegerin causes osteoarthritis with chondrocalcinosis. AnnRheumDis. 2014 Apr 17;(1468–2060).
52. Peters JE, Lyons PA, Lee JC, Richard AC, Fortune MD, Newcombe PJ, et al. Insight into Genotype-Phenotype Associations through eQTL Mapping in Multiple Cell Types in Health and Immune-Mediated Disease. PLoS Genet . 2016;12 (3):e1005908.
53. Westra HJ, Franke L. From genome to function by studying eQTLs. BiochimBiophysActa. 2014 May 4;(0006–3002 (Print)).
54. Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet . 2013;45 (10):1238–43.
55. Gee F, Clubbs CF, Raine E V, Reynard LN, Loughlin J. Allelic expression analysis of the osteoarthritis susceptibility locus that maps to chromosome 3p21 reveals cis-acting eQTLs at GNL3 and SPCS1. BMCMedGenet. 2014;15(1471–2350):53.
56. Egli RJ, Southam L, Wilkins JM, Lorenzen I, Pombo-Suarez M, Gonzalez A, et al. Functional analysis of the osteoarthritis susceptibility-associated GDF5 regulatory polymorphism. Arthritis Rheum. 2009 Jul;60(0004–3591 (Print)):2055–64.
57. Bos SD, Bovee J V, Duijnisveld BJ, Raine E V, van Dalen WJ, Ramos YF, et al. Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. AnnRheumDis. 2012 Jul;71(1468–2060):1254–8.
58. Raine E V, Dodd AW, Reynard LN, Loughlin J. Allelic expression analysis of the osteoarthritis susceptibility gene COL11A1 in human joint tissues. BMCMusculoskeletDisord. 2013;14(1471–2474):85.
59. Bos SD, Suchiman HE, Kloppenburg M, Houwing-Duistermaat JJ, le Graverand MP, Seymour AB, et al. Allelic variation at the C-reactive protein gene associates to both hand osteoarthritis severity and serum high sensitive C-reactive protein levels in the GARP study. AnnRheumDis. 2008 Jun;67(1468–2060):877–9.
60. Price AL, Helgason A, Thorleifsson G, McCarroll SA, Kong A, Stefansson K. Single-tissue and cross-tissue heritability of gene expression via identity-by-descent in related or unrelated individuals. PLoS Genet. 2011;7 (2).
61. Petretto E, Mangion J, Dickens NJ, Cook SA, Kumaran MK, Lu H, et al. Heritability and tissue specificity of expression quantitative trait loci. PLoS Genet. 2006;2 (10):1625–33.
62. Huang J, Chen J, Esparza J, Ding J, Elder JT, Abecasis GR, et al. eQTL mapping identifies insertion- and deletion-specific eQTLs in multiple tissues. Nat Commun . 2015;6(May):6821.
63. Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. NatGenet. 2012 May;44(1546–1718):491–501.

64. Raine E V, Wreglesworth N, Dodd AW, Reynard LN, Loughlin J. Gene expression analysis reveals HBP1 as a key target for the osteoarthritis susceptibility locus that maps to chromosome 7q22. AnnRheumDis. 2012 Dec;71(1468–2060):2020–7.
65. Upton AR, Holding CA, Dharmapatni AA, Haynes DR. The expression of RANKL and OPG in the various grades of osteoarthritic cartilage. RheumatolInt. 2012 Feb;32(1437–160X):535–40.
66. Nagase H, Nagasawa Y, Tachida Y, Sakakibara S, Okutsu J, Suematsu N, et al. Deiodinase 2 upregulation demonstrated in osteoarthritis patients cartilage causes cartilage destruction in tissue-specific transgenic rats. OsteoarthritisCartilage. 2013 Mar;21(1522–9653):514–23.
67. Kostopoulou F, Gkretsi V, Malizos KN, Iliopoulos D, Oikonomou P, Poultsides L, et al. Central role of SREBP-2 in the pathogenesis of osteoarthritis. PLoSOne. 2012;7(1932–6203):e35753.
68. Takahashi A, de Andrés MC, Hashimoto K, Itoi E, Oreffo ROC. Epigenetic regulation of interleukin-8, an inflammatory chemokine, in osteoarthritis. Osteoarthr Cartil. 2015;23 (11):1946–54.
69. Ijiri K, Zerbini LF, Peng H, Otu HH, Tsuchimochi K, Otero M, et al. Differential expression of GADD45beta in normal and osteoarthritic cartilage: potential role in homeostasis of articular chondrocytes. Arthritis Rheum. 2008 Jul;58(0004–3591 (Print)):2075–87.
70. Duval E, Bigot N, Hervieu M, Kou I, Leclercq S, Galera P, et al. Asporin expression is highly regulated in human chondrocytes. MolMed. 2011;17(1528–3658):816–23.
71. Lisignoli G, Toneguzzi S, Grassi F, Piacentini A, Tschon M, Cristino S, et al. Different chemokines are expressed in human arthritic bone biopsies: IFN-gamma and IL-6 differently modulate IL-8, MCP-1 and rantes production by arthritic osteoblasts. Cytokine. 2002 Dec 7;20(1043–4666 (Print)):231–8.
72. Logar DB, Komadina R, Prezelj J, Ostanek B, Trost Z, Marc J. Expression of bone resorption genes in osteoarthritis and in osteoporosis. JBone MinerMetab. 2007;25(0914–8779 (Print)):219–25.
73. Lambert C, Dubuc JE, Montell E, Verges J, Munaut C, Noel A, et al. Gene expression pattern of cells from inflamed and normal areas of osteoarthritis synovial membrane. Arthritis Rheumatol. 2014 Apr;66(2326–5205):960–8.
74. Leijten JC, Emons J, Sticht C, van GS, Decker E, Uitterlinden A, et al. Gremlin 1, frizzled-related protein, and Dkk-1 are key regulators of human articular cartilage homeostasis. Arthritis Rheum. 2012 Oct;64(1529–0131):3302–12.
75. Eccleston A, DeWitt N, Gunter C, Marte B, Nath D. Epigenetics. Nature . 2007;447 (7143):395–395.
76. Bird A. Perceptions of epigenetics. Nature . 2007;447 (7143):396–8.
77. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol . 2011;12 (1):R10.
78. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet . 2012;13 (7):484–92.
79. Muers M. Gene expression: Disentangling DNA methylation. Nat Rev Genet . 2013;14 (8):519.
80. Slieker RC, Bos SD, Goeman JJ, Bovee J V, Talens RP, van der Breggen R, et al. Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. EpigeneticsChromatin. 2013;6(1756–8935):26.
81. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. NatGenet. 2003 Mar;33 Suppl(1061–4036 (Print)):245–54.
82. Mendell JT, Olson EN. MicroRNAs in stress signaling and human disease. Cell. 2012 Mar 16;148(1097–4172):1172–87.
83. Zhang L, Yang M, Marks P, White LM, Hurtig M, Mi QS, et al. Serum non-coding RNAs as biomarkers for osteoarthritis progression after ACL injury. OsteoarthritisCartilage. 2012 Dec;20(1522–9653):1631–7.
84. Barter MJ, Young DA. Epigenetic mechanisms and non-coding RNAs in osteoarthritis. CurrRheumatolRep. 2013 Sep;15(1534–6307):353.
85. Fernandez-Tajes J, Soto-Hermida A, Vazquez-Mosquera ME, Cortes-Pereira E, Mosquera A, Fernandez-Moreno M, et al. Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster of osteoarthritic patients. AnnRheumDis. 2014 Apr;73(1468–2060):668–77.
86. 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. HumGenet. 2014 May 27;(1432–1203).
87. Reynard LN, Bui C, Canty-Laird EG, Young DA, Loughlin J. Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. HumMolGenet. 2011 Sep 1;20(1460–2083):3450–

60.
88. Rushton MD, Reynard LN, Barter MJ, Refaie R, Rankin KS, Young DA, et al. Characterization of the cartilage DNA methylome in knee and hip osteoarthritis. Arthritis Rheumatol. 2014 May 16;(2326–5205).
89. Delgado-Calle J, Fernandez AF, Sainz J, Zarrabeitia MT, Sanudo C, Garcia-Renedo R, et al. Genome-wide profiling of bone reveals differentially methylated regions in osteoporosis and osteoarthritis. Arthritis Rheum. 2013 Jan;65(1529–0131):197–205.
90. Loughlin J, Reynard LN. Osteoarthritis: Epigenetics of articular cartilage in knee and hip OA. Nat Rev Rheumatol . 2015;11 (1):6–7.
91. Roach HI, Yamada N, Cheung KSC, Tilley S, Clarke NMP, Oreffo ROC, et al. Association between the abnormal expression of matrix-degrading enzymes by human osteoarthritic chondrocytes and demethylation of specific CpG sites in the promoter regions. Arthritis Rheum. 2005;52 (10):3110–24.
92. Bui C, Barter MJ, Scott JL, Xu Y, Galler M, Reynard LN, et al. cAMP response element-binding (CREB) recruitment following a specific CpG demethylation leads to the elevated expression of the matrix metalloproteinase 13 in human articular chondrocytes and osteoarthritis. FASEB J . 2012;26 (7):3000–11.
93. Hashimoto K, Otero M, Imagawa K, De Andrés MC, Coico JM, Roach HI, et al. Regulated transcription of human matrix metalloproteinase 13 (MMP13) and interleukin-1β (IL1B) genes in chondrocytes depends on methylation of specific proximal promoter CpG sites. J Biol Chem. 2013;288 (14):10061–72.
94. Im G-I, Kim K-I, Park S-Y. Increased methylation status of SOX9 gene promoter in human osteoarthritic cartilage. Ann Rheum Dis . 2013;71.
95. Kim K Il, Park YS, Im G Il. Changes in the epigenetic status of the SOX-9 promoter in human osteoarthritic cartilage. J Bone Miner Res. 2013;28 (5):1050–60.
96. Hashimoto K, Oreffo ROC, Gibson MB, Goldring MB, Roach HI. DNA demethylation at specific CpG sites in the IL1B promoter in response to inflammatory cytokines in human articular chondrocytes. Arthritis Rheum. 2009;60 (11):3303–13.
97. Andrés MC, Imagawa K, Hashimoto K, Gonzalez A, Roach HI, Goldring MB and Oreffo ROC, et al. Loss of methylation in CpG sites in the proximal coding region and NF-B enhancer elements of iNOS are responsible for gene induction in human articular chondrocytes. Bone . 2011;48:S155.
98. Imagawa K, De Andrés MC, Hashimoto K, Itoi E, Otero M, Roach HI, et al. Association of reduced type IX collagen gene expression in human osteoarthritic chondrocytes with epigenetic silencing by DNA hypermethylation. Arthritis Rheumatol. 2014;66 (11):3040–51.

2

EXPLORATION OF THE ARTICULAR CARTILAGE TRANSCRIPTOME REVEALS *CRLF1* AS NOVEL OSTEOARTHRITIS SUSCEPTIBILITY GENE

Wouter den Hollander[1*], Cindy Boer[2*], Irina Pulyakhina[3,4], Nils Bomer[1], Ruud van der Breggen[1], Wibowo Arindrarto[5], Nico Lakenberg[1], Joris Deelen[1], Marian Beekman[1], Jeroen F.J. Laros[5], Peter A.C. 't Hoen[3], Margreet Kloppenburg[6], Eline P.E. Slagboom[1], Rob G.H.H. Nelissen[7], Yolande F.M. Ramos[1] Joyce van Meurs[2‡], Ingrid Meulenbelt[1‡].

[1]Dept. of Molecular Epidemiology, LUMC, Leiden, The Netherlands. [2]Dept. of Internal Medicine, Erasmus Medical Center, Rotterdam, the Netherlands. [3]Dept. of Human Genetics, LUMC, Leiden, The Netherlands. [4]Wellcome Trust Centre for Human Genetics Oxford, Oxfordshire, United Kingdom. [5]Medical Statistics and Bioinformatics, LUMC, Leiden, The Netherlands. [6]Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands. [7]Dept. of Orthopaedics, LUMC, Leiden, The Netherlands.

* *These authors contributed equally to this work*
‡ *These authors jointly supervised this work*

ABSTRACT

Multiple osteoarthritis (OA) susceptibility single nucleotide polymorphisms (SNPs) mark imbalanced expression of positional genes in articular cartilage, reflected by unequally expressed alleles among heterozygotes (allelic imbalance, AI). Here, we explored the articular cartilage transcriptome from OA patients for AI events to identify putative disease driving genetic variation by RNA sequencing. We observed 2070 SNPs that consistently marked AI of 1031 unique genes in articular cartilage, of which 32 were additionally significantly differentially expressed (0.5>FC>2, FDR<0.05) between preserved and paired lesioned cartilage. Among those was *CRLF1* (FC=3.17, FDR=$7.86*10^{-5}$), which appeared subject to AI, marked by lower expression of the rs7256319 alternative allele T compared to the reference allele C. Additionally, the T allele harbored a protective signal in a combined genetic association meta-analysis (OR=0.881, CI_{95}=0.781-0.994, P=0.0393). Finally, we show that increased *CRLF1* signaling markedly affects expression of anabolic genes *in-vitro*, while catabolic markers remain unaffected.

INTRODUCTION

Due to the increased proportion of elderly in the human population, osteoarthritis (OA) has become one of the major musculoskeletal diseases (1). While all joint tissues have been implicated in OA pathology, the disease is characterized primarily by progressive degradation and calcification of articular cartilage (2). Both gene targeted (3-5), as well as genome-wide research (6-9) showed that a multitude of genes are involved in the currently irreversible destruction of articular cartilage that precedes total joint replacement surgery; presently the only effective treatment for end-stage OA. In this regard, numerous studies have reported on altered regulation of gene expression that reflects, attenuates and/or stimulates OA-mediated cartilage degradation (10-13). Moreover, multiple OA risk alleles of single nucleotide polymorphisms (SNPs) were shown to consistently modulate OA pathology due to altered transcription of the respective genes in articular cartilage (14-18). Hence, it is clear that *in-cis* genetic regulation of transcription plays a substantial role in cartilage homeostasis and, therefore, in OA pathophysiology.

SNPs that confer risk for OA frequently act via allele-specific gene regulation as reflected by unequally expressed alleles among heterozygous carriers, commonly referred to as an allelic imbalance (AI) (15, 19-22). Notable examples are *DIO2* (19) and *GDF5* (15), genes of which the OA-associated risk allele affects transcription in articular cartilage. Additionally, AI among OA risk SNPs that are not situated within gene bodies have been addressed by measuring AI among SNPs in strong linkage disequilibrium (LD), as shown for *ALDH1A2* (22). Despite the evidence for *in-cis* genetic regulation of transcription in OA susceptibility, genome-wide association studies (GWAS) have thus far failed to explain the larger part of the hereditary component of OA (23). In this regard, a large number of the tested SNPs in GWAS likely bear no biological function in relation to the addressed phenotype or disease relevant tissues (24), resulting in massive inflation of, possibly biologically irrelevant, statistical tests and thus the multiple testing correction penalty. Consequently, large numbers of SNPs that do bear biological functionality in the context of OA are missed. Furthermore, SNPs that reside within LD blocks are hard to interpret, as association analysis is inherently unable to distinguish disease-relevant alleles from mere statistically associated alleles.

In previous studies, we and others have used targeted approaches to address AI events of putative, as well as established OA susceptibility genes (19-22, 25, 26). Given the successful identification of the transcriptional consequences of multiple OA-associated SNPs, we have here aimed to characterize, on a transcriptome-wide scale, novel SNPs that tag AI of genes expressed in articular cartilage. By means of RNA sequencing we addressed AI events and combined them with differential expression between preserved and lesioned articular cartilage, from the RAAK study (12), of the respective genes. Finally, we applied association analyses and functional experiments to reveal a novel OA susceptibility gene.

RESULTS

TRANSCRIPTOME-WIDE DISCOVERY OF ARTICULAR CARTILAGE ALLELIC IMBALANCE EVENTS.

To understand how genetic variation contributes *in-cis* to transcriptional regulation in articular cartilage on a transcriptome-wide scale, we first called heterozygous SNPs (dbSNP144) using RNA sequencing data from articular cartilage derived from patients who underwent total joint replacement surgery of either the hip (N=22) or knee (N=25) joint due to primary OA (**Supplementary Table S1**). After filtering by the number of read counts per position ($\boldsymbol{R}_{reference} \geq 10$, $\boldsymbol{R}_{alternative} \geq 10$ and $\boldsymbol{R}_{total} \geq 25$), selecting for heterozygous SNPs present in at least 2 individuals, removing SNPs present in multiple distinct transcripts and discarding the HLA locus, we defined $\boldsymbol{\varphi}$ for 13853 SNPs as the measure of imbalance (**Figure 1A**), which denotes the fraction of $\boldsymbol{R}_{alternative}$ among $\boldsymbol{R}_{total}$. Possibly due to reference bias, a considerable number of SNPs marked AI by $\boldsymbol{\varphi} < 0.1$ or $\boldsymbol{\varphi} > 0.9$ (N=418) and were subsequently removed prior to further analyses. A meta-analysis per SNP (null hypothesis: median $\boldsymbol{\varphi}=0.49$) and subsequent multiple testing correction (FDR) revealed 2070 SNPs that significantly marked AI among respectively 1031 genes (**Figure 1B**, **Supplementary Table S2**). We assessed five SNPs that have been reported to mark AI in articular cartilage (**Supplementary Figure S1**). As such, we were able to replicate AI of *ALDH1A2*, marked by the alleles of rs3204689 ($\boldsymbol{\varphi}=0.42$, FDR=0.003) (22), and, additionally, report on multiple other SNPs at the *ALDH1A2* locus that do so. These SNPs were, expectedly, in high LD and thus suggest that the observed SNP marking AI might not be the actual regulatory SNP. In that regard, we were unable to detect significant AI reports for rs225014 ($\boldsymbol{\varphi}=0.49$, FDR=0.98, *DIO2*) (19), rs1676486 ($\boldsymbol{\varphi}=0.51$, FDR=0.49, *COL11A1*) (21), rs11177 ($\boldsymbol{\varphi}=0.47$, FDR=0.59, *GNL3*) (20) or rs6617 ($\boldsymbol{\varphi}=0.51$, FDR=0.63, *SPCS1*) (20). This observation may be due to lack of power in our dataset (insufficient heterozygous individuals), insufficient positional read depth, or may imply that the actual regulatory SNPs for these genes are in incomplete LD with the SNPs that mark AI reported by us and others. Finally, it could be that targeted assays are simply more sensitive.

DIFFERENTIAL GENE EXPRESSION BETWEEN PRESERVED AND PAIRED OA LESIONED CARTILAGE.

While SNPs marking AI in articular cartilage could contribute to OA pathophysiology in various ways, e.g. in cartilage development or homeostasis, it can be expected that those located in genes which additionally mark the articular cartilage's disease state are more likely to contribute to or attenuate disease progression. Therefore, we went back to the original expression data and determined differential expression, using the edgeR R package, in patients for which paired RNA sequencing data of both preserved and OA-lesioned articular cartilage was generated (6 hip joints, 15 knee joints). Of the 10468 Ensembl gene identifiers with at least 5 counts per million, 118 and 48 were observed to be respectively significantly (FDR<0.05) down- (FC<0.5) and upregulated (FC>2) in lesioned compared to

preserved cartilage (**Figure 2A**, **Supplementary Table S3**). As has been reported by microarray studies that have utilized a similar design, Gene Ontology (GO) term enrichment analysis (**Supplementary Table S4**) revealed significant enrichment for inflammatory (e.g. *SCUBE1*, *CFH* and *CXCL14,* **Figure 2B-D**), response to wound healing (e.g. *NOTCH3*, *BMP5* and *SERPINE1*, **Figure 2E-G**) and joint development associated pathways (e.g. *SPP1*, *MMP3* and *COL9A1*, **Figure 2H-J**). Of the 166 differentially expressed genes, 32 were additionally subject to AI, marked by 74 SNPs (**Supplementary Table S5**).

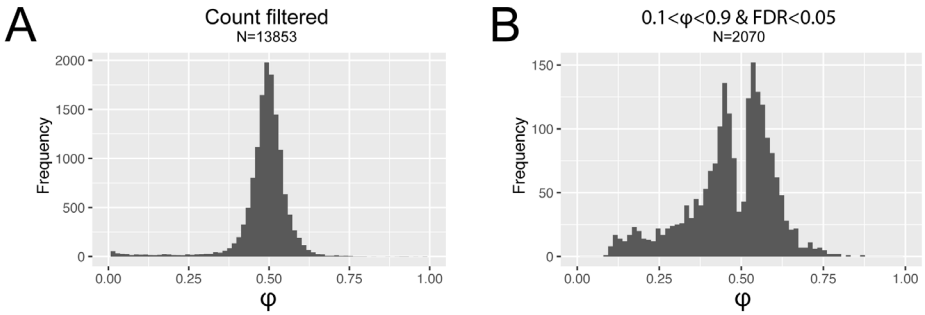


Figure 1. Distribution of AI events in articular cartilage. (**A**) All 13435 called variants after selecting for at least 2 heterozygotes, selecting SNPs present in only single genes and removing low counts. (**B**) After filtering by allelic fraction (0.1<ϕ<0.9) and FDR<0.05 for all heterozygotes in the same direction per SNP, 352 SNPs remained that marked consistent AI of 219 unique genes.

GENETIC ASSOCIATION ANALYSIS OF AI SNPS IN DISEASE MODULATED GENES.

As a proof of principle, we performed genetic association analysis among participants of the Genetics osteoARthritis and Progression study (GARP, N=380) and healthy controls from the Leiden Longevity Study (LLS, N=2315) with the top 10 AI SNPs (single most significant SNP per gene) among differentially expressed genes (**Supplementary Table S6**). After adjustment for age, sex and BMI, we observed a significant protective association for the alternative allele T of rs7256319 (C>T, MAF=0.39, OR=0.803, CI_{95}=0.681-0.946, P=0.009), located in the cytokine receptor-like factor 1 gene (*CRLF1*) encoding a member of the cytokine type I receptor family (CLF1). The CLF1 protein is reported to signal via the ciliary neurotrophic factor receptor (CNTFR) gene (*CNTFR*) after heterodimerization with cardiotrophin-like cytokine factor 1 (CLCF1) gene (*CLCF1*) (27). We were able to replicate the protective signal by meta-analysis, having additionally included three independent cohorts from the Rotterdam Study (RS1: 188 cases, 1903 controls, RS2: 57 cases, 742 controls RS3: 36 cases, 1185 controls), among which cases were defined as having underwent a total joint replacement surgery. This analyses confirmed the direction of effect and provided a significant association in the combined meta-analysis (OR=0.881, CI_{95}=0.781-0.994, P=0.0393, I (2)= 0%, $P_{heterogeneity}$=0.417, **Supplemental Figure S2**).

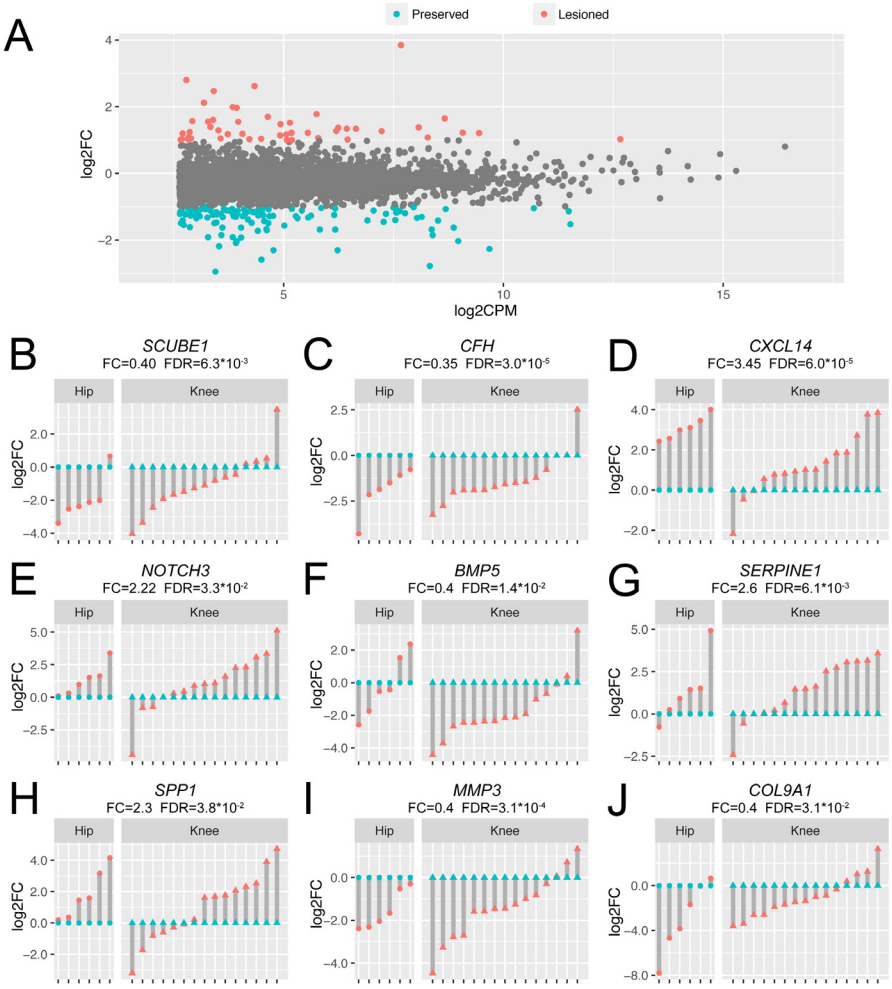


Figure 2. Paired differential gene expression between preserved and lesioned articular cartilage. (**A**) Volcano plot, showing FDR cutoff (0.05) and foldchange cutoff (0.5>FC>2.0) indicated in red. (**B-D**) Differential expression of selected genes from inflammatory pathways. (**E-G**) Differential expression of selected genes from developmental pathways. (**H-J**) Differential expression of selected genes from developmental related pathways. Hip and knee joints are depicted by triangles and circles, respectively. Preserved and lesioned samples are depicted by blue and red, respectively.

CRLF1 SIGNALING PLAYS AN ACTIVE ROLE IN END-STAGE OA ASSOCIATED CARTILAGE DEGRADATION.

The protective T allele marked consistent lower *CRLF1* expression (***φ***=0.29, FDR=4.02*10^{-21}, **Figure 3A, Supplementary Table 2**) compared to the reference allele C, which was confirmed by custom TaqMan assay performed in respectively 5 preserved and 5 lesioned articular cartilage samples, originating from 6 independent patients that underwent total knee replacement surgery (**Figure 3B**). In parallel, and similar to distinct publically available microarray datasets, expression of the three genes (*CRLF1, CNTFR* and

CLCF1) in the current dataset was significantly different between preserved and OA-lesioned cartilage (**Supplementary Table S3**) with a significant upregulation in OA affected cartilage for *CRLF1* (FC=3.17, FDR=7.86*10^{-5}, **Figure 3C**) as well as for *CLCF1* (FC=1.56, FDR=0.01, **Figure 3D**), while the receptor for the complex (CNTFR) revealed significant downregulation (FC=0.31, FDR=3.43*10^{-6}, **Figure 3E**). Together, these data demonstrated that the protective effect of the alternative allele T of rs7256319 likely acts via mitigation of *CRLF1/CLCF1* signaling with ongoing OA pathophysiology.

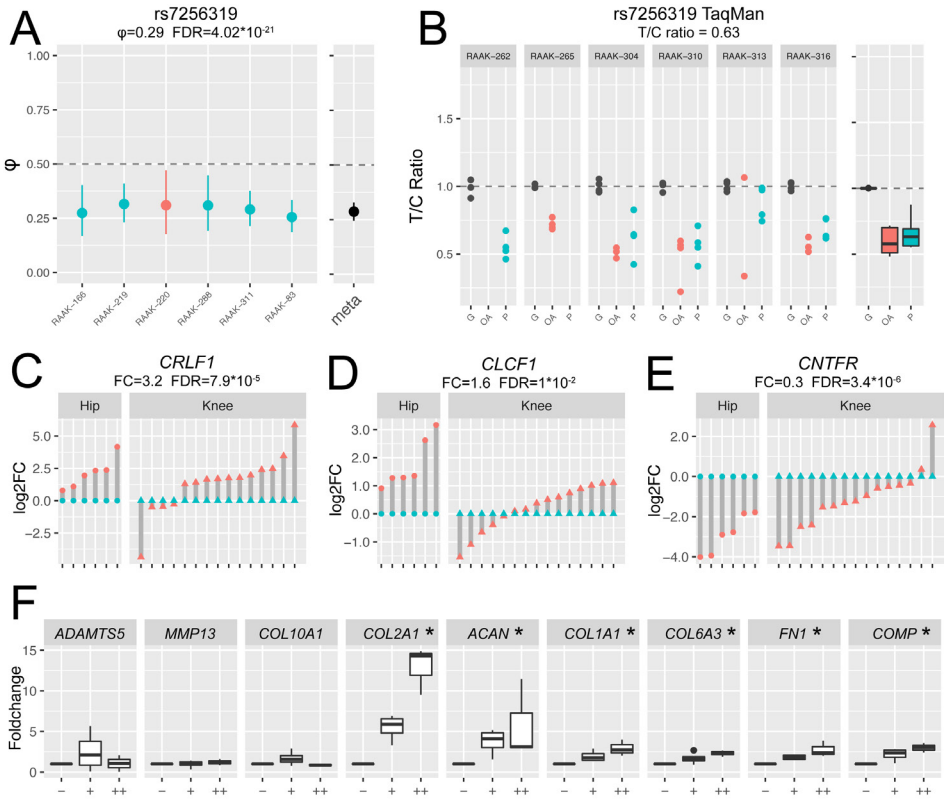


Figure 3. *CRLF1* signaling is modulated in articular cartilage during OA progression. (**A**) *CRLF1* expression is subject to AI, marked by lower expression of the alternative allele T of rs7256319 (ϕ=0.29, FDR=1.26*10^{-14}). (**B**) Replication of the observed AI tagged by rs7256319 by TaqMan genotyping in 6 additional knee samples ratified the observed lower expression of the alternative allele T. Genomic DNA (grey) was used as reference ratio. (**C**) *CRLF1* expression was significantly upregulated in OA lesioned articular cartilage compared with paired preserved articular cartilage (FC=3.17, P_{FDR}=4.06*10^{-4}). (**D**) *CLCF1* expression was significantly upregulated in OA lesioned articular cartilage compared with paired preserved articular cartilage (FC=1.57, P_{FDR}=0.02). (**E**) The CLF1/CLC complex receptor gene *CNTFR* was significantly downregulated in OA lesioned articular cartilage compared with paired preserved articular cartilage (FC=0.32, P_{FDR}=5.26*10^{-5}). (**F**) Expression of catabolic markers *ADAMTS5*, *MMP13* and *COL10A1* were not affected by addition of either 10 (+) or 50 (++) ng mL^{-1} CLF1/CLC protein complex to the culture medium of primary chondrocytes isolated from 4 individual knee cartilage samples. Expression of anabolic markers *COL2A1*, *AGC*, *COL1A1*, *COL6A3*, *FN1* and *COMP* were significantly affected as reflected by a dose response curve upon addition of the protein complex. Preserved samples are indicated in blue, lesioned samples in red, knees as circles and hips as triangles.

CLF1/CLCF1 MODULATES ARTICULAR CARTILAGE ANABOLISM *IN-VITRO*.

Finally, we studied downstream effects of increased *CRLF1* signaling in cartilage micro-masses derived from primary human chondrocytes isolated from 4 patients who underwent a total joint replacement of the knee due to primary OA. After proliferation and formation of the extracellular matrix (ECM) on day 3, micro-masses were treated with 10 or 50 ng mL^{-1} of human recombinant CLF1/CLCF1 protein complex for 4 consecutive days and RNA was isolated to assess transcriptional differences. We did not observe a difference in expression of the catabolic genes *ADAMTS5*, *MMP13* and *COL10A1* with either concentration of added CLF1/CLCF1 protein complex compared to control cultures. However, a marked dose response upregulation was observed for multiple anabolic genes, such as *COL2A1*, *AGC*, *COL1A1*, *COL6A3*, *FN1* and *COMP* (**Figure 3F**).

DISCUSSION

Our approach in the current paper comprises a concept framework for complex traits to identify disease-relevant genetic variation, as reflected by allele associated transcription levels in the disease's pivotal tissue. We aim to present the reported observations as a legacy dataset for researchers from the field to probe for their gene or SNP of interest and have described by proof of concept how the *CRLF1* locus seems causally involved in the progression of OA-associated cartilage degradation.

CRLF1, the gene encoding for the CLF1 protein, harbors the rs7256319 C>T SNP, which marked imbalanced expression of its respective alleles in articular cartilage, reflected by consistent lower expression of the alternative allele T in comparison with the reference allele C among heterozygotes. As has been reported previously (12, 27), *CRLF1* was significantly upregulated in OA-affected compared to preserved articular cartilage, as was its signaling partner *CLCF1*, while the protein complex signaling receptor gene *CNTFR* was downregulated. The decreased expression of the alternative allele of rs7256319 in heterozygotes, significant upregulation of *CRLF1* in lesioned cartilage and the respective protective genetic association seem to indicate that mitigating *CRLF1*/*CLCF1* signaling will be beneficial towards ongoing cartilage degradation due to primary OA.

Less straightforward, however, may be the subsequent observation that increasing concentrations of CLF1/CLC protein complex, applied to primary chondrocytes derived cartilage constructs, resulted in a significant dose responsive upregulation of multiple anabolic cartilage genes, while catabolic genes remained unaffected. This, as well as upregulation of catabolic genes and downregulation of anabolic genes are generally considered relevant determinants in OA. We propose that increased expression of anabolic markers reflect the incapability of chondrocytes to return to a steady-state upon disease

associated CLF1/CLC-CNTFR signalling, likely required for adapting to mechanical stress and subsequent articular cartilage micro trauma. Alternatively, increased expression of these anabolic genes might influence the stiffness of the ECM and thereby affect mechanical resilience of articular cartilage. In any way, our data suggest that consistent innate increased expression of *CRLF1*, either during development of the cartilage anlagen and/or with ongoing pathology, affects the propensity of articular cartilage to engage an OA phenotype. Paradoxically, treatment of 2D cultured mouse ATDC5 cells with human CLF1/CLC complex was shown to induce downregulation of *Acan* and *Col2a1*, while catabolic genes remained unaffected (28). Aside from this apparent species-specific and/or difference in culture protocol (3D vs 2D) response, the role of the CLF1/CLC complex in OA seems depicted by the inability to accordingly regulate articular cartilage anabolism, as opposed to affect degradation.

Up for speculation remains how rs7256319 marks allelic *CRLF1* expression mechanistically, especially due to its location within an intron, suggesting intron retention and/or alternatively spliced transcripts, which we were, however, unable to detect consistently (data not shown). A number of transcription factors is predicted to bind at the rs7256319 locus (e.g. *JUN*, *REST* and *SP1*, (based on HaploReg 4.1)) and regulation might thus be affected by altered binding motifs. Possibly, regulation by SNPs in high LD with rs7256319 that are situated in other regulatory elements, such as rs3170474, which disrupts another set of transcription factor binding sets (e.g. *HEY1*, *INSM1* and *PAX5*), is located in the 3' UTR of *C19orf60* and was shown to be an expression quantitative trait locus (eQTL) in brain tissues, are responsible for the marked allelic imbalance.

Although outside the scope of our current efforts, we believe that alongside rs7256319, the 2070 SNPs that mark AI in articular cartilage are likely to contain multiple compelling association signals. While in canonical GWAS a strict genome-wide significance level of 5×10^{-8} is imposed due to the vast amount of SNPs that is tested for, we here postulate that selecting for SNPs that are more likely to contribute to the disease *a priori*, by means of marking AI in articular cartilage, could aid the search for OA susceptibility loci substantially. Further downstream selection criteria, such as, but not limited to, significant differential expression between preserved and OA lesioned cartilage, will help tailor genetic association analyses even more and might attribute SNPs to specific disease facets, such as progression of cartilage degradation, as we have shown here. Of note, in this regard, it deserves mention that inherent to our study design, we have potentially missed genes that affect joint morphology, cartilage integrity during development and/or change expression during early stage OA.

A number of SNPs is known to mark AI in articular cartilage, as has been shown by gene-targeted approaches. We were able to replicate the earlier observed AI of *ALDH1A2* and to lesser extent of *DIO2*, marked by rs3204689 (22) and rs225014 (19) respectively, while rs1676486 (*COL11A1*) (21),

rs11177 (*GNL3*) (20) and rs6617 (*SPCS1*) (20) did not mark consistent AI in our dataset. Additionally, we did not observe heterozygotes for rs143383 (*GDF5*) (25) or rs3815148 (*HBP1*) (26). These findings indicate first and foremost that additional replication, preferably using a different technique, is required to increase confidence in the observed AI, as we have done for rs7256319. Secondly, it stresses the fact that observed AI reflects regulatory properties of the respective LD block and does not *per se* identify genetic variation that affects respective gene expression levels mechanistically. Furthermore, despite the applied filtering steps and statistics, the list of significant AI SNPs potentially contains a number of false positives, of which some could have originated from alignment bias. While future novel alignment and other bioinformatic approaches (29) might address these issues from a more fundamental perspective, we have here aimed to reduce them by inclusion of multiple filtering steps ($0.1<\boldsymbol{\varphi}<0.9$, present in at least two heterozygotes and null hypothesis adjustment).

In summary, we present a framework and resulting dataset for researchers in the OA research field to probe for disease relevant genetic variation that affects respective gene expression in the disease's pivotal affected tissue. Furthermore, conceptual downstream analyses and experiments revealed *CRLF1* to be actively involved in OA associated articular cartilage degradation, as reflected by lower expression of the rs7256319 protective allele T among heterozygous carriers and aberrant tissue homeostasis upon increased CLF1/CLCF1 signalling *in-vitro*.

MATERIALS AND METHODS

COHORTS

Ethical approval for the RAAK study was obtained from the medical ethics committee of the LUMC (P08.239) and informed consent was obtained from 68 participants. From 21 patients preserved as well as lesioned cartilage was sampled (6 hip patients, 15 knee patients), complemented by an additional 21 preserved (14 hip patients, 7 knee patients) and 5 lesioned (2 hip, 3 knee patients) samples (**Supplementary Table S1**). For cartilage sampling details see (12, 14). The GARP study consists of 380 siblings with symptomatic OA at multiple joint sites. BMI, age and sex was available in 378 subjects with mean BMI = 27.03, age range 50-75 years, 82.0% female) (9, 30). The LLS study consists of 2415 individuals of whom genetic data is available for 1583 offspring of long-lived individuals and 732 of their partners. BMI, age and sex was available in 1976 subjects with mean BMI = 25.41, age range 30-80 years, 54.9% female. (31) Genotyping and quality control of the data from GARP and LLS participants was done as described previously (11). In the current paper, to select for symptomatic OA patients, we performed in silico association of rs7256319 to those patients of the Rotterdam studies that underwent a joint replacement surgery as result of their OA.

RNA-SEQUENCING DATA.

Post RNA isolation (Qiagen RNeasy Mini Kit, RIN>7), paired-end 100 bp RNA library sequencing (Illumina TruSeq RNA Library Prep Kit, Illumina HiSeq 2000) resulted in an average of 10 million clusters. Reads were aligned using GSNAP against the hg19 reference genome, while known Dutch SNPs (GoNL) were masked to aid with potential reference alignment bias. AI events were assessed on SNPs called using SNVMix2 with default settings (32) with minimum coverage of 25 and at least 10 reads (***R***) per allele. AI is reported by the average fraction (***φ***) of the alternative allele reads ($\boldsymbol{R}_{alternative}$) among the total number of reads ($\boldsymbol{R}_{total}$=$\boldsymbol{R}_{alternative}$+$\boldsymbol{R}_{reference}$) at the position of the respective genetic variation per sample (***i***):

$$\boldsymbol{\varphi} = \frac{1}{n}\sum_{i=1}^{n}\frac{R_{i,alternative}}{R_{i,reference}}$$

To detect SNPs that robustly mark imbalance two binomial tests were performed per heterozygote and per SNP under the null hypothesis that the amount of imbalance is either greater or smaller than 0.49. Subsequently, P-values per SNP were corrected for multiple testing (FDR) by the number of heterozygotes of the respective SNP and considered significant if all FDR corrected P-values were <0.05 and in the same direction among all heterozygotes. Using the edgeR package, fragments per gene were used to assess the dispersion by quantile-adjusted conditional maximum likelihood (qCML) (33). Subsequently, differential gene expression analysis was performed pairwise between preserved and lesioned samples for which we had RNA of both (N=21, **Supplementary Table S1**) followed by FDR correction. GO term enrichment analysis was performed using the online available tool DAVID (34).

TAQMAN ASSAY

Conventional TaqMan genotyping was performed on both genomic DNA and articular cartilage cDNA (35) from 6 (2 female, 4 male) patients who underwent total joint replacement surgery of the knee due to primary OA. An allele-specific custom TaqMan assay for rs7256319 (Thermo Fisher Scientific) was used to quantify the allele ratio in cDNA samples and were normalized against the gDNA ratio, which was used as a 1:1 allele ratio reference.

CELL CULTURE AND RT-QPCR

Within the ongoing RAAK study human primary chondrocytes were isolated from macroscopically intact cartilage of OA patients (4 females) undergoing total knee arthroplasty, as described previously (19). In short, cartilage was rinsed with phosphate buffered saline (PBS), cut into small pieces, and incubated overnight with 3 mg mL^{-1} collagenase I (Worthington Biochemical Corporation, USA) in

Dulbecco's Modified Eagle's Medium (DMEM; Gibco) containing antibiotics, 10% fetal bovine serum (Gibco) and 0.5 ng mL^{-1} recombinant human basic FGF (PeproTech). Next day, cells were filtered through a 100 μm mesh to remove undigested cartilage fragments and debris. Chondrocytes were passaged twice before harvest and incubation in 3D-pellets (2.5 x 10^5 cells/pellet) in serum-free chondrogenic medium in the absence or presence of 10 or 50 ng mL^{-1} recombinant human CRLF1/CLC complex (R&D Systems). Chondrogenic medium with or without recombinant proteins was refreshed at day 2 and day 6, and pellets were harvested for RNA isolation at day 7 (Qiagen RNeasy Mini Kit). Gene expression was determined with RT-qPCR using FastStart SYBR Green Master reaction mix (Roche Applied Science) according to the manufacturer's protocol and corrected for ADP-ribosylation factor related protein 1 (*ARFRP1*; primer sequences are shown in **Supplementary Table S7**).

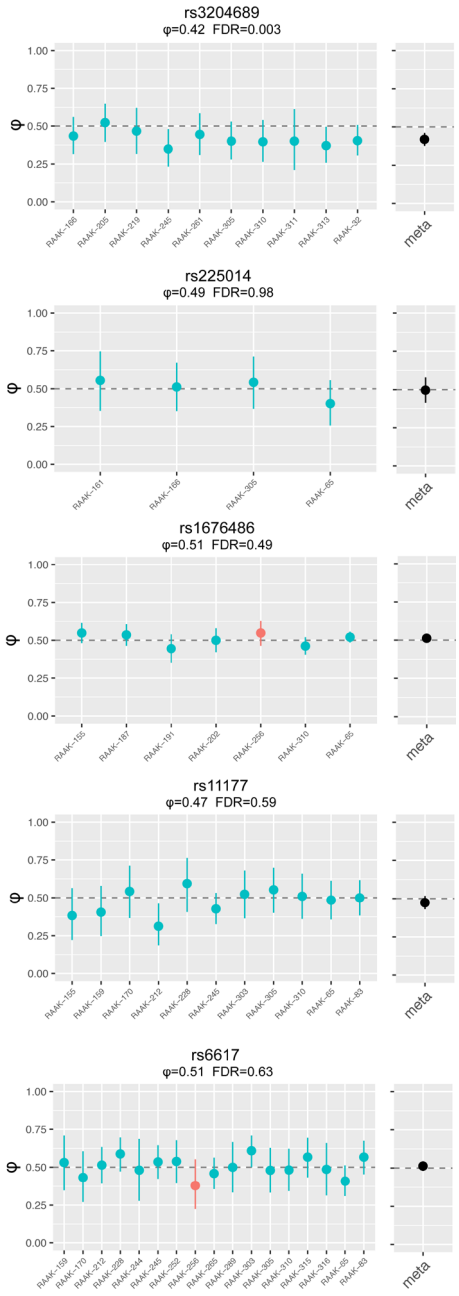
ACKNOWLEDGEMENTS

The Leiden University Medical Centre, the Dutch Arthritis Association and Pfizer Inc., Groton, CT, USA support the GARP study, whilst the LLS was supported by the Netherlands Organization of Scientific Research (MW 904-61-095, 911-03-016, 917-66-344 and 911-03-012), Leiden University Medical Centre, and by the "Centre of Medical System Biology" and the "Netherlands Consortium of Healthy Aging" in the framework of the Netherlands Genomics Initiative (NGI). Furthermore, the research leading to the RAAK biobank and the current results has received funding from the Dutch Arthritis Association (DAA 2010_017) and the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement no. 259679. We thank Nico Lakenberg, Ruud van der Breggen and Eka Suchiman, for their help in preparing DNA and RNA samples.

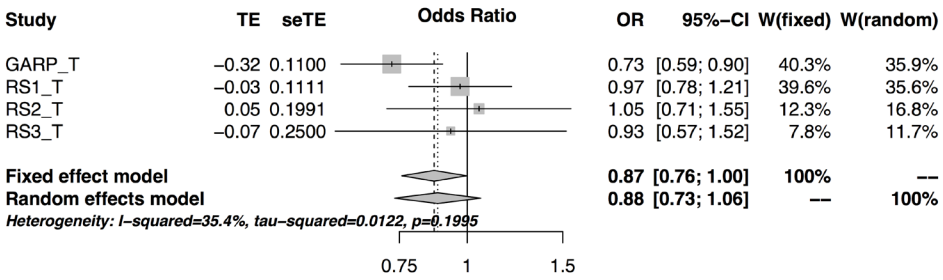
REFERENCES

1. Bijlsma,J.W., Berenbaum,F., & Lafeber,F.P. Osteoarthritis: an update with relevance for clinical practice. *The Lancet* 377, 2115-2126 (2011).
2. Man,G.S. & Mologhianu,G. Osteoarthritis pathogenesis- a complex process that involves the entire joint. *J. Med. Life* 7, 37-41 (2014).
3. Saito,T. *et al.* Transcriptional regulation of endochondral ossification by HIF-2[alpha] during skeletal growth and osteoarthritis development. *Nat Med* 16, 678-686 (2010).
4. Sherwood,J. *et al.* A homeostatic function of CXCR2 signalling in articular cartilage. *Ann. Rheum. Dis.* (2014).
5. Blaney Davidson,E.N. *et al.* TGF-beta is a potent inducer of Nerve Growth Factor in articular cartilage via the ALK5-Smad2/3 pathway. Potential role in OA related pain? *Osteoarthritis Cartilage* 23, 478-486 (2015).
6. Jeffries,M.A. *et al.* Genome-wide DNA methylation study identifies significant epigenomic changes in osteoarthritic cartilage. *Arthritis Rheumatol.* (2014).
7. Rushton,M.D. *et al.* Characterization of the cartilage DNA methylome in knee and hip osteoarthritis. *Arthritis Rheumatol.* (2014).
8. Zeggini,E. *et al.* Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet* 380, 815-823 (2012).
9. Meulenbelt,I. *et al.* Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. *Human Molecular Genetics* 17, 1867-1875 (2008).
10. Xu,Y. *et al.* Identification of the pathogenic pathways in osteoarthritic hip cartilage: commonality and discord between hip and knee OA. *Osteoarthritis. Cartilage.* 20, 1029-1038 (2012).
11. den Hollander,W. *et al.* Transcriptional associations of osteoarthritis mediated loss of epigenetic control in articular cartilage. *Arthritis & Rheumatology*n/a (2015).
12. Ramos,Y.F. *et al.* Genes Involved in the Osteoarthritis Process Identified through Genome Wide Expression Analysis in Articular Cartilage; the RAAK Study. *PLoS. ONE.* 9, e103056 (2014).
13. Zhang,R. *et al.* Gene expression analyses of subchondral bone in early experimental osteoarthritis by microarray. *PLoS. ONE.* 7, e32356 (2012).
14. Bomer,N. *et al.* Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis. *Ann. Rheum. Dis.* 74, 1571-1579 (2015).
15. Reynard,L.N., Bui,C., Canty-Laird,E.G., Young,D.A., & Loughlin,J. Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. *Hum. Mol. Genet.* 20, 3450-3460 (2011).
16. Styrkarsdottir,U. *et al.* Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene and with rare variants at 1p31. *Nat Genet* advance online publication, (2014).
17. Gee,F., Clubbs,C.F., Raine,E.V., Reynard,L.N., & Loughlin,J. Allelic expression analysis of the osteoarthritis susceptibility locus that maps to chromosome 3p21 reveals cis-acting eQTLs at GNL3 and SPCS1. *BMC. Med. Genet.* 15, 53 (2014).
18. Raine,E.V., Dodd,A.W., Reynard,L.N., & Loughlin,J. Allelic expression analysis of the osteoarthritis susceptibility gene COL11A1 in human joint tissues. *BMC. Musculoskelet. Disord.* 14, 85 (2013).
19. Bos,S.D. *et al.* Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. *Annals of the Rheumatic Diseases* 71, 1254-1258 (2012).
20. Gee,F., Clubbs,C.F., Raine,E.V., Reynard,L.N., & Loughlin,J. Allelic expression analysis of the osteoarthritis susceptibility locus that maps to chromosome 3p21 reveals cis-acting eQTLs at GNL3 and SPCS1. *BMC. Med. Genet.* 15, 53 (2014).
21. Raine,E.V., Dodd,A.W., Reynard,L.N., & Loughlin,J. Allelic expression analysis of the osteoarthritis susceptibility gene COL11A1 in human joint tissues. *BMC. Musculoskelet. Disord.* 14, 85 (2013).
22. Styrkarsdottir,U. *et al.* Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene and with rare variants at 1p31. *Nat Genet* 46, 498-502 (2014).
23. Loughlin,J. Osteoarthritis year 2010 in review: genetics. *Osteoarthritis Cartilage* 19, 342-345 (2011).
24. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 348, 648-660 (2015).
25. Southam,L. *et al.* An SNP in the 5'-UTR of GDF5 is associated with osteoarthritis susceptibility in Europeans and with in vivo differences in allelic expression in articular cartilage. *Hum. Mol. Genet.* 16, 2226-2232 (2007).

26. Raine,E.V.A., Wreglesworth,N., Dodd,A.W., Reynard,L.N., & Loughlin,J. Gene expression analysis reveals HBP1 as a key target for the osteoarthritis susceptibility locus that maps to chromosome 7q22. *Annals of the Rheumatic Diseases* 71, 2020-2027 (2012).
27. Elson,G.C. *et al.* CLF associates with CLC to form a functional heteromeric ligand for the CNTF receptor complex. *Nat. Neurosci.* 3, 867-872 (2000).
28. Tsuritani,K. *et al.* Cytokine receptor-like factor 1 is highly expressed in damaged human knee osteoarthritic cartilage and involved in osteoarthritis downstream of TGF-beta. *Calcif. Tissue Int.* 86, 47-57 (2010).
29. Castel,S.E., Levy-Moonshine,A., Mohammadi,P., Banks,E., & Lappalainen,T. Tools and best practices for data processing in allelic expression analysis. *Genome Biol.* 16, 195 (2015).
30. Riyazi,N. *et al.* Evidence for familial aggregation of hand, hip, and spine but not knee osteoarthritis in siblings with multiple joint involvement: the GARP study. *Ann. Rheum. Dis.* 64, 438-443 (2005).
31. Westendorp,R.G. *et al.* Nonagenarian siblings and their offspring display lower risk of mortality and morbidity than sporadic nonagenarians: The Leiden Longevity Study. *J. Am. Geriatr. Soc.* 57, 1634-1637 (2009).
32. Goya,R. *et al.* SNVMix: predicting single nucleotide variants from next-generation sequencing of tumors. *Bioinformatics.* 26, 730-736 (2010).
33. Robinson,M.D., McCarthy,D.J., & Smyth,G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 26, 139-140 (2010).
34. Huang,d.W., Sherman,B.T., & Lempicki,R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44-57 (2009).
35. Wilkins,J.M. *et al.* Extreme context-specificity in differential allelic expression. *Hum. Mol. Genet.* (2007).



Supplementary Figure S1. The extent of AI plotted for SNPs that have been reported to mark AI in known OA susceptibility genes.



Supplementary Figure S2. Forest plot of meta-analysis performed in GARP and RS for rs7256319.

Supplementary tables upon request.

Supplementary Table S1.

Sample characteristics.

Supplementary Table S2.

All observed SNPs in at least two heterozygotes.

Supplementary Table S3.

Differential gene expression between preserved and paired lesioned samples.

Supplementary Table S4.

Significantly enriched GO terms among the differentially expressed genes (N=166, 0.5>FC>2, FDR<0.05).

Supplementary Table S5.

Primer sequences used for RT-qPCR analysis.

Supplementary Table S6.

Overview of differentially expressed genes (0.5>FC>2) between preserved and paired OA lesioned articular cartilage that are additionally influenced by *in-cis* genetic regulation as reflected by AI. FC=Foldchange. FDR=Benjamini-Hochberg corrected P-value.

Supplementary Table S7.

qPCR primer sequences.

3

GENOME-WIDE ASSOCIATION AND FUNCTIONAL STUDIES IDENTIFY A ROLE FOR MATRIX-GLA PROTEIN IN OSTEOARTHRITIS OF THE HAND

Wouter den Hollander[1*], Cindy G. Boer[2*], Debbie Hart[3], Michelle S. Yau[4,5], Yolande F.M. Ramos[1], Sarah Metrustry[3], Linda Broer[2], Joris Deelen[1,6], L. Adrienne Cupples[7], Fernando Rivadeneira[2], Margreet Kloppenburg[8], Marjolein Peters[2], Tim D. Spector[3], Albert Hofman[9,10], P. Eline Slagboom[1], Rob G.H.H. Nelissen[11], André G. Uitterlinden[2,9], David T. Felson[5,12], Anna M. Valdes[13], Ingrid Meulenbelt[1‡], Joyce J.B. van Meurs[2‡]

[1]Department of Medical Statistics and Bioinformatics, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. [2]Department of Internal Medicine, Genetic Laboratory, Erasmus Medical Center, Rotterdam, the Netherlands. [3]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [4]Institute for Aging Research, Hebrew SeniorLife, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA. [5]Clinical Epidemiology Research and Training Unit, Boston University School of Medicine, Boston, MA, USA. [6]Max Planck Institute for Biology of Ageing, Cologne, Germany. [7]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; The Framingham Heart Study, Framingham, MA, USA. [8]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands. [9]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. [10]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [11]Department of Orthopedics, Leiden University Medical Center, Leiden, The Netherlands. [12]Arthritis Research UK Epidemiology Unit, University of Manchester, Manchester, UK. [13] School of Medicine, University of Nottingham, Nottingham, UK.

* *These authors contributed equally to this work*

‡ *These authors jointly supervised this work*

http://dx.doi.org/10.1136/annrheumdis-2017-211214

ABSTRACT

Objective Osteoarthritis (OA) is the most common form of arthritis and the leading cause of disability in the elderly. Of all the joints, genetic predisposition is strongest for OA of the hand, however only few genetic risk loci for hand OA have been identified. Our aim was to identify novel genes associated with hand OA and examine the underlying mechanism.

Methods We performed a genome-wide association study of a quantitative measure of hand OA in 12,784 individuals (discovery: 8,743, replication: 4,011). Genome-wide significant signals were followed up by analysing gene and allele specific expression in a RNA-sequencing dataset (n=96) of human articular cartilage

Results We found two significantly associated loci in the discovery set: at chr12 ($P=3.5 \times 10^{-10}$) near the MGP gene and at chr12 ($P=6.1 \times 10^{-9}$)near the CCDC91 gene. The DNA variant near the MGP-gene was validated in three additional studies, which resulted in a highly significant association between the MGP-variant and hand OA (rs4764133, $Beta_{meta}=0.83$, $P_{meta}=1.8*10^{-15}$). This variant is high linkage disequilibrium with a coding variant in *Matrix Gla-Protein* (*MGP*), a vitamin K-dependent inhibitor of cartilage calcification. Using RNA-sequencing data from human primary cartilage tissue (n=96), we observed that the hand OA MGP-risk allele was significantly lower expressed compared to the reference allele (40.7%, $P<5*10^{-16}$).

Conclusions Our results indicate that the association between the MGP-variant and increased risk for hand OA is caused by a lower expression of *MGP*, which may increase the burden of hand OA by decreased inhibition of cartilage calcification.

INTRODUCTION

Osteoarthritis (OA) is the most frequent joint disorder worldwide. An estimated 22% of the adult population has a joint affected by OA and this incidence increases to 49% in individuals over 65 years of age (1). All synovial joints can be affected by OA, with hand OA as one of the most common forms of OA. Hand OA is characterized by osteophyte formation, bony enlargements of finger joints and cartilage degradation in the joints. One of the factors contributing to cartilage degradation is the increase of calcified cartilage in the joint (2,3). In addition, hand OA is related to the occurrence of OA at other sites, most notably with knee OA (4,5). Patients affected by hand OA suffer from pain and disability, impacting their quality of life. OA is a leading cause of chronic disability (6), yet currently no effective therapeutic treatments against osteoarthritis are known. It is therefore imperative to dissect the underlying mechanism of disease aetiology as this may enhance effective and targeted drug development.

OA has a strong genetic component. Depending on the joint affected, the heritability of OA is estimated in the range of 40-60% (7,8), with hand OA having the largest heritability, i.e. ~60% (9, 10). Therefore, in recent years, several large-scale genetic studies have been performed to identify the underlying genes and pathways leading to OA. Multiple significant associated loci for OA of the hip and knee have been identified through genome-wide association studies (GWAS) (11–18). However, thus far, only one report has described a robust association with OA of the hand (19). In this previous report, common variants in the *ALDH1A2* and rare variants in chromosome 1p31 were genome-wide significantly associated with hand OA using a discovery cohort of 837 cases and 77,325 controls.

In this study, we aimed to identify novel genes and pathways involved in the aetiology of OA of the hand by performing a large-scale GWAS. We used a semi- quantitative measure for OA of the hand in order to increase statistical power. We gathered a large sample size of 12,754 individuals for analysis, by combining data from three studies in the discovery phase and an additional three cohorts for replication. Next, we conducted functional follow-up of our top finding to investigate the underlying mechanism.

METHODS

DISCOVERY GWAS, REPLICATION AND META-ANALYSIS

For a detailed description on the GWAS methods, participating studies, quality control procedures for genotyping and imputation, see **Supplementary Text S1** and **Supplementary Table S1**.

DETAILED PHENOTYPE DESCRIPTION OF KL SUM-SCORE

We have used a semi-quantitative bilateral measure of osteoarthritis of the hand based on the radiographic Kellgren and Lawrence score (KL-score) (20). Using radiographs of both hands, the KL-score was determined for each joint in the hand. Using these KL-scores we defined the KLsum-score: the total KL-score, the sum, of the following hand joints for both hands (left and right): all Distal Interphalangeal (DIP) joints, all Proximial Interphalangeal (PIP) joints, all Metacarpophalangeal (MCP) joints, the Interphalangeal (IP) joint and the first Carpometacarpal (CMC1) joint. Which gives the sum of 15 joints on each hand, and in total 30 joints for both hands together, resulting in a minimum score of 0 and a maximum score of 120. The Leiden Studies cohort no Kellgren-Lawrence scoring was done of the MCP joints, resulting in a KLsum-score of maximum 88. Individuals lacking KL-grading for both or one hands and individuals with missing age or gender information were excluded from all analyses in all cohorts. As the KL-sum score has a skewed distribution the top finding of the meta-analysis was repeated in the discovery cohorts using a Poisson regression.

VISUALIZATION OF THE ASSOCIATED LOCI AND THE REGULATORY LANDSCAPE

For the top GWAS associated SNP, the LD region (r^2>0.8) was determined using the 1000G Phase-1 population using the HaploReg V3 tool (21). Using the ROADMAP generated reference epigenomes we determined if any of the variants in high LD were located in potential gene regulatory regions in primary osteoblasts (generated by ENCODE) and bone marrow derived chondrocytes (ROADMAP) (22,23). The 18-state chromatin reference epigenomes were downloaded from the ROADMAP epigenomes data portal (23). SNPs and regulatory annotations were visualized busing the UCSC genome-browser on GRCh37/hg19 (24). For each variant it was also determined if the alternative allele would disrupt a protein binding motif, this was done using the HaploReg V3 tool (21).

RNA-SEQUENCING DATA

Post RNA isolation (Qiagen RNeasy Mini Kit, RIN>7) of 40 knee (15 paired preserved (P) and OA lesioned (OAL), 7 P only, 3 OAL only) and 28 hip (6 paired P and OAL, 14 P only, 2 OAL only) cartilage samples (**Supplementary Table S2**), paired-end 2x100bp RNA library sequencing (Illumina TruSeq RNA-Library Prep Kit, Illumina HiSeq2000) resulted in an average of 10 million fragments per sample. Reads were aligned using GSNAP against GRCh37/hg19, in which SNPs from the Genome

of the Netherlands consortium with MAF>1% were masked to prevent alignment bias. Number of fragments per gene were used to assess quantile-adjusted conditional maximum likelihood (qCML) (edgeR, R-package).Subsequently, differential gene expression analysis was performed pairwise between P and OAL samples for which we had RNA of both (n=21). ASE was assessed using SNVMix2 (25) with default settings (min coverage=25, 10 reads per allele). The extent of allele specific expression (ASE) was defined as the fraction of risk allele among all counts at the respective location. Meta-analysis was done only across P samples or OAL when no P counterpart sample was present. P-values were calculated using canonical binominal test (metagen R-package).

TAQMAN ASSAY

Conventional TaqMan genotyping was performed on both genomic-DNA (gDNA), articular cartilage and Subchondral bone cDNA. An allele-specific custom TaqMan assay for rs1800801 (Thermo Fisher Scientific) was used to quantify the allele ratio in cDNA samples and were normalized against the gDNA ratio, which was used as an 1:1 allele ratio reference. Each sample has been measured in four (cartilage) or eight (subchondral bone) times, while calculations and statistics were performed as described previously (19,26). Cartilage samples which yielded fewer than four measurements (N=2) were discarded prior to further analyses. All subchondral bone samples were assessed by eight technical replicates.

RESULTS

GWAS OF KLSUM-SCORE

We conducted a genome-wide association study (GWAS) of a semi-quantitative measure of hand OA, a bilateral summed score of Kellgren- and Lawrence scores (20), that grades radiographic OA severity, across all hand joints (KLsum-score, range of 0 to 120). The discovery set consisted of three Rotterdam Study cohorts (RSI, RSII and, RSIII) and included 8,743 participants with KLsum-scores. Replication was done in another 4,011 individuals from three different cohorts; Leiden studies(LS), Framingham heart study(FHS), and Twins-UK (TUK). General characteristics of the discovery cohorts and replication cohorts can be found in **Supplementary Table S3** and in **Supplementary Text S1**.

The discovery analysis yielded two novel independent genome-wide significant loci ($P \leq 5*10^{-8}$) on chromosome 12, an intergenic region between *MGP* and *ERP27* and an intronic region in *CCDC91*. We also identified seven other novel loci with suggestive significance ($P<5*10^{-6}$) (**Figure 1**). In total, nine loci were selected for replication in 4,011 individuals from three different cohorts (LS, FHS, and TUK). Using a Bonferroni corrected P-value$<5.56*10^{-3}$, we significantly replicated one of nine loci,

rs4764133 ($Beta_{meta}$=0.83, SE_{meta}=0.10, $P\text{-value}_{replication}=3.4*10^{-7}$, $P\text{-value}_{meta}=1.8*10^{-15}$) with the same direction of effect as identified in the discovery analysis (**Table 1**, and **Supplementary Figure S1**). This locus maintained genome-wide significance and another locus near *ENPP3* reached near genome-wide significance (chr6:132063842:D, $Beta_{meta}$=0.58, SE_{meta}=0.11, $P\text{-value}_{meta}=3.8*10^{-7}$) in the combined discovery and replication joint meta-analysis (**Table 1**). Since the KLsum-score has a skewed distribution the top hit was also re-analysed in the discovery set using a Poisson regression (rs4764133, $Beta_{poisson}$=0.12, $SE_{poisson}$=0.02, $P\text{-value}_{poisson}=1.98*10^{-11}$).

Our top replicated and genome-wide significant finding, rs4764133 [T] ($P_{meta}=1.80*10^{-15}$, Beta=0.83, MAF=0.39) is located in a non-coding intergenic region between *MGP* (Matrix Gla-protein) and *ERP27* (Endoplasmic Reticulum Protein 27). However, variants in high linkage disequilibrium (LD) with rs4764133 ($r^2 \geq 0.8$) span a ~80Kb region encompassing multiple genes, including *MGP* and an open-reading frame *C12orf60* (**Figure 2A**). Moreover, several of these variants are located in an mRNA transcript, including a nonsynonymous variant in *MGP*, and variants in 3' and 5'UTR of *MGP* and *C12orf60* (**Table 2**, **Figure 2B**). The nonsynonymous variant in *MGP,* rs4236, is predicted to be non-damaging (STIFT=1, tolerated; PolyPhen=0, benign) causing a threonine to alanine amino acid substitution. Two variants are located in predicted active promoter region of *MGP* (rs1800801) and *C12orf60* (rs9668569) in chondrogenic cells and primary osteoblasts (**Table 2**).

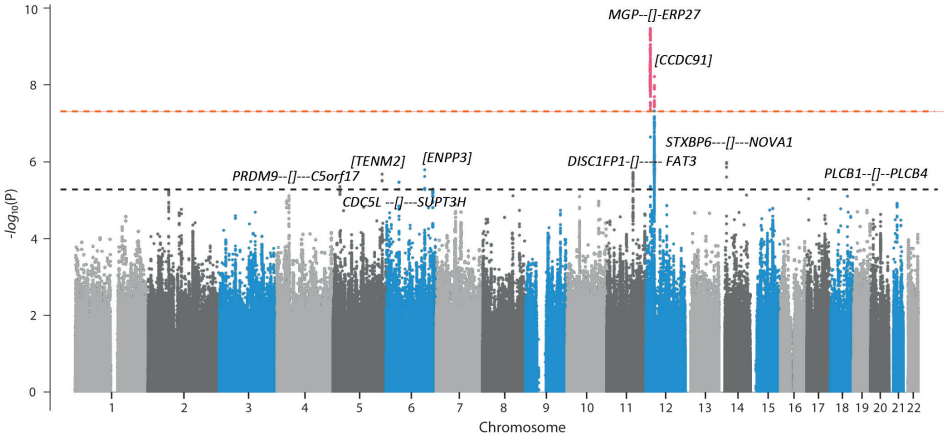


Figure 1. GWAS results for association with the KLsum score in the discovery phase. Manhattan plot for association with the KLsum-score, adjusted for age and sex, in the discovery cohorts of RSI, RSII and RSIII. The -log10 P-values, for each of the ~11 million SNPs analyzed (remaining after EASYQC quality control) as part of the genome wide association with the KLsum-score, plotted against their position per chromosome. The red dotted horizontal line corresponds to the genome-wide significant threshold ($P = 5x10^{-8}$). The dotted grey line corresponds to the selection for replication threshold ($P = 5x10^{-6}$). SNP location represented by [], if the SNP is localized intergenic the dashes denotes the distance,-≤10 kb,--≤100kb,---≤1000kb,----≤1Mkb,-----≥1Mkb.

Table 1. Results GWAS quantitative bilateral phenotype of osteoarthritis of the hand (KLsum-score), discovery, replication and meta-analysis

						Discovery*			Replication**			Combined			
SNP	**Chr**	**Position (hg19)**	**Effect Allele**	**Other Allele**	**EAF†**	**Beta**	**SE**	**P-value**	**Beta**	**SE**	**P-value**	**Beta**	**SE**	**Pval**	**Genomic location††**
rs1494593	5	23574856	T	C	0.88	-0.83	0.18	4.46E-06	-0.15	0.32	6.31E-01	-0.67	0.16	2.28E-05	*PRDM9--[]---C5orf17*
rs114370021	5	167398535	A	G	0.27	-0.84	0.18	2.12E-06	0.48	0.32	1.28E-01	-0.53	0.16	6.96E-04	*[TENM2]*
rs7770034	6	44447004	A	G	0.48	-0.54	0.12	3.42E-06	-0.36	0.21	8.71E-02	-0.50	0.10	1.01E-06	*CDC5L --[]---SUPT3H*
6:132063842:D‡	*6*	*132063842*	*D*	*I*	*0.27*	*0.64*	*0.13*	*1.28E-06*	*0.41*	*0.23*	*7.82E-02*	*0.58*	*0.11*	*3.79E-07*	*[ENPP3]*
11:90657297:D	11	90657297	D	I	0.11	0.93	0.19	1.35E-06	0.33	0.33	3.11E-01	0.78	0.17	2.91E-06	*DISC1FP1-[]----FAT3*
rs4764133	**12**	**15064363**	**T**	**C**	**0.39**	**0.75**	**0.12**	**3.45E-10**	**1.11**	**0.22**	**3.34E-07**	**0.83**	**0.10**	**1.80E-15**	***MGP--[]- ERP27***
rs7139060	12	28693144	A	G	0.67	-0.73	0.12	6.12E-09	0.11	0.22	6.16E-01	-0.52	0.11	1.47E-06	*[CCDC91]*
rs1950427	14	25955502	T	C	0.12	0.86	0.18	1.09E-06	-0.28	0.30	3.56E-01	0.57	0.15	1.80E-04	*STXBP6---[]---NOVA1*
rs6108226	20	8960884	T	C	0.77	0.70	0.15	3.94E-06	0.20	0.27	4.43E-01	0.58	0.13	1.14E-05	*PLCB1--[]--PLCB4*

* Discovery: RS-I, RS-II, RS-III, *n*= 8,743

** Replication : GARP, LSS, TwinsUK & FHS, *n*= 4,011

† EAF: Effect Allele Frequency

††SNP location represented by [], if the SNP is localized intergenic the dashes denotes the distance,-≤10 kb,--≤100kb,---≤1000kb,----≤1Mkb,-----≥1Mkb

‡For TwinsUK a proxy SNP was used: rs3850251 r^2=1 D'=1 (as calculated in the RSI, RSII and RSIII cohorts)

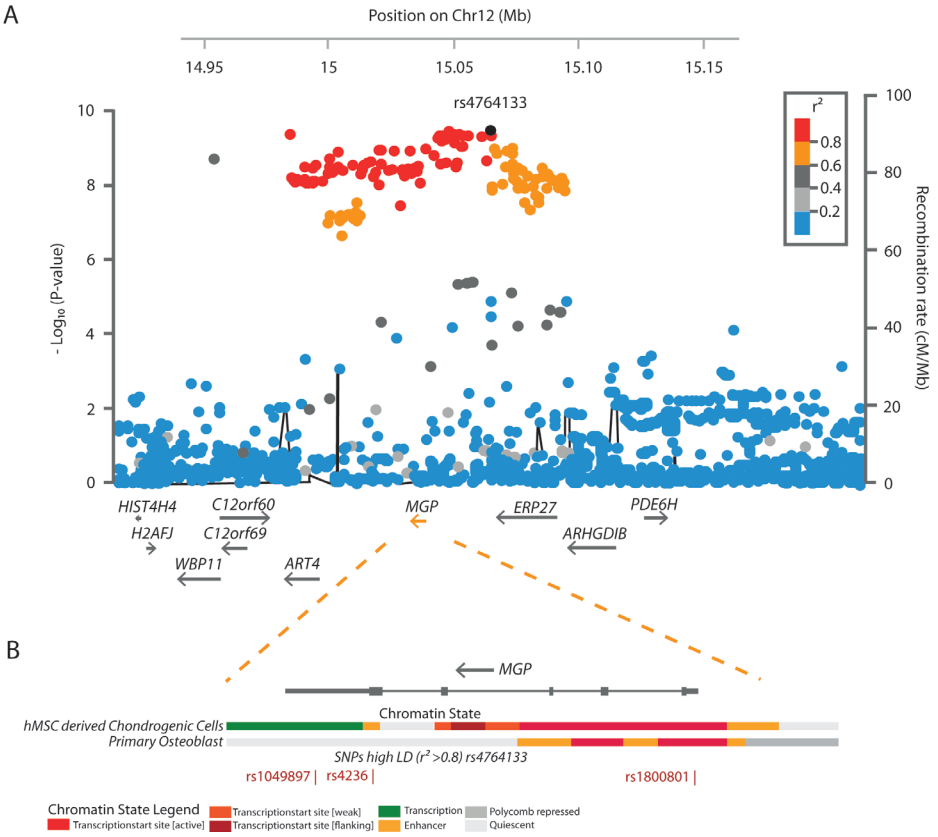


Figure 2. Locus zoom plot for rs4764133. locus zoom plot for rs4764133, 150 kb upstream and downstream of rs4764133 has been taken as plotted region (**A**). Zoom in on *MGP* and three SNPs in high LD with top SNP that are located in the *MGP* mRNA transcript (**B**). Also represented is ROADMAP chromatin 18-state data of two tissue types: human Mesenchymal Stem Cell (hMSC) derived cultured chondrogenic cells and primary osteoblasts. In both these cell types the chromatin contains active marks surrounding the *MGP* promoter.

Table 2. rs4764133 LD block (r^2>0.8) annotation of potential functional elements in osteoblasts and chondrogenic cells, X marks no potential functional annotation i.e. enhancer region, promoter region or altered protein binding motifs.

				Regulatory Chromatin Marks**		
SNP	**P-value Discovery**	**r^2**	**Annotation***	**Chondrogenic cells**	**Osteoblasts**	**Altered Protein Binding Motifs (Haploreg V3)**
rs1049897	3.48E-09	0.88	*MGP* 3'-UTR	Transcription	X	X
rs4236	4.16E-09	0.86	*MGP* non-synonymous	Enhancer region	X	HNF4, PLAG1
rs1800801	1.12E-09	0.95	*MGP* 5'UTR	Promoter region	Promoter region	Zfp410
rs7310951	4.04E-09	0.86	*C12orf60*	Enhancer region	X	DMRT7, Gfi1, Pax-5
rs12320004	4.04E-09	0.86	*C12orf60*	Enhancer region	X	BHLHE40, P300, HEN1, LBP-1, RAD21,TATA, Zfx
rs10772814	3.76E-09	0.88	*C12orf60*	Enhancer region	X	HNF4
rs10492151	1.21E-09	0.95	*C12orf60*	Enhancer region	X	AIRE, Hoxa13
rs725445	3.58E-08	0.82	*C12orf60*	Enhancer region	X	Hand1
rs725444	3.92E-09	0.87	*C12orf60*	Enhancer region	X	Foxf1, Foxi1, Foxo, Foxq1, Mef2
rs4764131	6.31E-10	0.97	*C12orf60*	Enhancer region	Enhancer region	Myc
rs9668569	5.91E-10	0.97	*C12orf60*	Promoter region	Promoter region	X
rs2430687	2.44E-09	0.89	*C12orf60*	Enhancer region	Enhancer region	BHLHE40
rs12311463	6.91E-10	0.97	*C12orf60*	Enhancer region	Enhancer region	Pou1f1, Pou2f2, TATA
rs67482087	4.61E-10	0.95	*C12orf60*	Enhancer region	Enhancer region	Foxp1, Irx, Pou1f1, Pou2f2, Pou3f3, TATA
rs67436073	6.76E-10	0.97	*C12orf60*	Enhancer region	Enhancer region	Foxj2, Foxk1, Foxo, GATA, Mef2, Pou2f2, Pou3f2, Pou6f1, TATA, Zfp
rs11276	8.05E-09	0.96	*C12orf60* non-synonymous	X	X	SPIB,NF-AT
rs3088189	9.46E-09	0.96	*C12orf60* synonymous	X	X	SPIB
rs1861698	3.56E-09	0.96	*C12orf60* synonymous	X	X	Bbx, Pou1f1, TATA

*Gene annotation based on the hg19 release of the UCSC Genome Browser

** Regulatory chromatin marks taken from the ROADMAP Epigenomes project chromatin state learning core 18-state model

Next, we investigated the association of rs4764133 with bilateral severe hand OA and bilateral finger OA using the discovery set (RS-I, RS-II and RS-III). We found a strong association with finger OA (P-value=3.09*10^{-8}, OR=1.25) and nominal significant association with severe hand OA (p-value=2.80*10^{-2}, OR=1.36), which has a low frequency in the population (**Supplementary Table S4**), we also see a nominal significant association with cartilage thickness in the hip joint (minimal joint space width (mJSW)). To see if rs4764133 also confers risk for other forms of osteoarthritis, i.e. osteoarthritis of the hip and knee, we used the GWAS summary data of the TreatOA consortium (27) and the recently published mJSW meta-analysis (18). No association was found between rs4764133 and hip or knee OA (**Supplementary Table S4**). However, we did find a nominal significant association for rs1049897 (r^2= 0.98 with rs4764133)(P-value=1.28*10^{-2}, Beta=-0.398).

GENE EXPRESSION ANALYSES

In order to identify potential causal genes located in the LD block surrounding rs4764133, we assessed gene expression of *MGP, ERP27, ART4, SMOC3* (*C12orf69*) and *C12orf60* in articular cartilage, the primary OA affected tissue. RNA sequencing was obtained on articular cartilage from primary OA patients who had total joint replacement surgeries of either the knee (n=25) or hip (n=22) joint. Expression levels of *ERP27, C12orf60, ART4* and *SMOC3* were substantially lower than *MGP* expression levels in articular cartilage (**Supplementary Figure S2A**). Nonetheless, neither *MGP*, *ERP27, ART4* nor *SMOC3* and *C12orf60* showed significant difference in gene expression between paired preserved(P) and OA lesioned(OAL) articular cartilage. However, while these genes are not differentially expressed in OA affected cartilage, it is possible that the identified GWAS SNPs affect gene transcription. When we analysed the relationship between the top SNP and expression analysis in a classical eQTL (expression Quantitative Trait Loci) analysis, we did not to detect significant correlations between rs1049897, rs4236 or rs1800801 and absolute *MGP, ERP27, ART4, SMOC3* or *C12orf60* expression levels (**Supplementary Figure S2B**). However, we did observe several variants in high LD located in the mRNA transcript of *MGP* and *C12orf60*, allowing us to assess allele specific expression (ASE) for these genes. We were unable to study ASE for *ART4*, *SMOC3* and *ERP27,* since no SNP in high LD with rs4764133 is present in the coding region. In ASE the influence of exonic alleles on gene expression *in-cis* is measured within heterozygote subjects, circumventing strong effects from environmental or trans-acting influences. This property results in ASE analysis to be a more statistically powerful approach, when compared to classical eQTL analysis (28). Subsequently, we found that the OA risk alleles for three coding variants in high LD with the lead variant, rs4236 (**Supplementary Figure 3SA**, 39.6% C allele, P<5*10^{-16}), rs1049897 (**Supplementary Figure S3B**, 44.4% A allele, P<5*10^{-10}), and rs1800801 (**Figure 3A**, 40.7% T allele, P<5*10^{-16}), were significantly correlated with lower expression of *MGP*, marking imbalanced expression among heterozygotes, independent of the disease status of the articular cartilage. No allele specific expression was observed between SNPs

rs11276, rs3088189 and rs1861698 (residing in *C12orf60* and in high LD with the lead SNP, r^2>0.8, **Table 2**). Technical and biological replication was performed using a custom allele specific TaqMan assay for rs1800801 in eight additional heterozygous individuals for which we isolated RNA from P cartilage (n=2), OAL (n=2) or both (n=4) from primary knee OA patients and confirmed the observed imbalance in preserved articular cartilage (**Figure 3B**, relative allelic difference=0.92, P<1*10^{-6}), as well as in 8 knee subchondral bone samples (**Figure 3C**, relative allelic difference=0.78, P<1*10^{-4}).

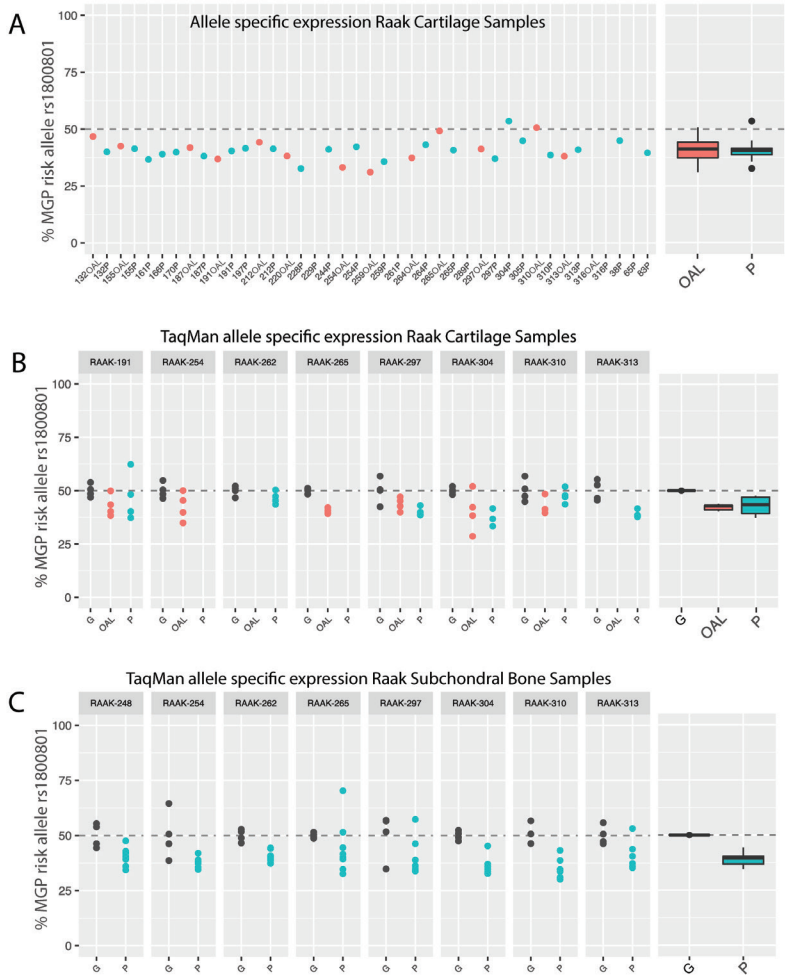


Figure 3. Allelic imbalanced expression of *MGP* marked by the alleles among heterozygotes of rs1800801 (**A**), in the assessed cartilage RNA sequencing dataset. Validation of selected rs1800801 using a custom TaqMan assay confirmed the imbalance (**B**). Allelic imbalance was also assed in with a custom TagMan assay in subchondral bone samples (**C**). Preserved (P) and OA lesioned (OAL) samples are shown respectively in blue and red, and genomic DNA (TaqMan control) in black (G). For ASE results for rs4236 and rs1049879, see **Supplementary Figure S3** and for information on the samples see **Supplementary Table S2**.

DISCUSSION

Here, we show for the first time, that there is a robust genome-wide significant association between rs4764133, located near *MGP,* and hand OA. Furthermore, we performed functional validation showing that *MGP* coding variants in LD with rs4764133 are associated with allele-specific expression of *MGP* which may increase risk of hand OA by lowering inhibition of articular cartilage calcification, since MGP is an essential inhibitor of cartilage calcification (29,30). These findings suggest that MGP could be considered a prioritized drug target for hand OA, since genetically supported drug targets double the success rate of therapeutics in clinical development (31).

MGP is an essential inhibitor of cartilage calcification, and genetic deficiencies of *MGP* in humans and mice have been linked to abnormal mineralization of soft tissues, including cartilaginous tissue (29,32). Furthermore *MGP* has been previously implicated in relation to OA. A small candidate study reported marginally significant association between hand OA and genetic variants in *MGP* (rs1800802 and rs4236) (33). This is consistent with our findings that the minor allele for rs4764133 and related coding variants in high LD (r^2>0.8), rs1800802 and rs4236, increase the risk of hand OA and that we found high expression of *MGP* in both preserved and OA lesioned articular cartilage. In contrast, another study showed that an MGP protein complex is excreted by healthy articular chondrocytes, but not by OA affected chondrocytes (34), although we only assed *MGP* expression and not MGP protein complex excretion.

Although the loci with allele specific expression (ASE) are known to be enriched for eQTLs (35), we were unable to detect an association between the *MGP*-genotype and *MGP* RNA-expression levels in cartilage. This could have been due to our modest sample size (knee joint, n=25 and or hip joint, n=22) in combination with large heterogeneity of the tissue. Notably, the available cartilage samples originated from different joint sites (knee, hip) and different disease stage (preserved versus affected), and had large age range of the individuals. Also, it is known that ASE is a more powerful technique than classical eQTL analysis to identify functional SNPs influencing expression of genes (28). While the extent of imbalance could be considered relatively modest, an increasing number of OA associated SNP alleles appear to mark ASE by comparable amount (19,36–38). From a more biological perspective, one could consider a prolonged, albeit slight, deviation from homeostasis due to modest ASE of cartilage relevant genes to be of substantial influence over time. This latter hypothesis could contain the molecular basis for increased risk towards developing OA among the ageing population. Additionally, we observed that the rs1800801 alleles also affected expression of *MGP* in subchondral bone samples. This could imply that, in parallel to an effect in cartilage, the presumed disturbed cartilage homeostasis is further affected by the underlying bone. Further enabling the view that OA is a pathology of the entire joint.

Our findings may give an explanation for the known vitamin K association with OA: MGP mediated calcification inhibition is dependent on γ-carboxylation by vitamin K (39). It has been shown that low vitamin K intake is correlated with OA (40). Thus vitamin K intake may be a potential therapeutic treatment in OA. Recently, a first randomized control trial testing the effects of vitamin K on OA was published, which reported no overall effect of vitamin K on hand OA (41). Despite the low power of the trial, there was a significant beneficial effect on joint space narrowing (cartilage degradation) among those individuals that were vitamin K deficient at the start of the trial (41). Thus, an adequately powered study of vitamin K may be justified based on the found *MGP* association. Furthermore, genetic predisposition for hand OA, was not taken into account in the trial. Perhaps, genetic predisposition for hand OA (*MGP*-risk variants) in combination with insufficient vitamin K intake might potentiate cartilage calcification and subsequent risk for developing hand OA. Therefore, future OA trails, therapeutic and preventive treatments might benefit from taking a personalized medicine approach since genetically supported drug targets double the success rate of therapeutics in clinical development (31).

Styrkarsdottir *et al.* (2014) reported on common genetic variants that associate with severe hand OA, among the replication cohorts were the Leiden and Rotterdam cohorts (19). Although we observe suggestive signals at the reported locus (*ALDH1A2* gene, 1p31) the respective variants did not meet the genome-wide significance threshold in our analyses (**Supplementary Table S5**). This difference is likely caused by the markedly different phenotypes that were used for either analyses. Where Styrkarsdottir *et al.* studied a dichotomous severe hand OA phenotype, our phenotype was semi-quantitatively phenotype.

To conclude, we here present coding variants in *MGP*, that are associated with radiographic hand OA, and the hand OA risk allele marks lower expression of *MGP* in articular cartilage. Our findings suggest that *MGP* might play an important role in hand OA pathogenesis through pathways related to articular cartilage calcification and vitamin K. Better understanding of MGP gene and protein regulation and its relation to vitamin K intake and OA, may reveal novel therapeutic drug targets for hand OA.

ACKNOWLEDGMENTS

This study was funded by The Netherlands Society for Scientific Research (NWO) VIDI Grant 917103521. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the

Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

The generation and management of GWAS genotype data for the Rotterdam Study (RS I, RS II, RS III) was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. The GWAS datasets are supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/ Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters, MSc, and Carolina Medina-Gomez, MSc, for their help in creating the GWAS database, and Karol Estrada, PhD, Yurii Aulchenko, PhD, and Carolina Medina-Gomez, MSc, for the creation and analysis of imputed data

The Leiden University Medical Centre, the Dutch Arthritis Association and Pfizer Inc., Groton, CT, USA support the GARP study, whilst the LLS was supported by the Netherlands Organization of Scientific Research (MW 904-61-095, 911-03-016, 917-66-344 and 911-03-012), Leiden University Medical Centre, and by the “Centre of Medical System Biology” and the “Netherlands Consortium of Healthy Aging” in the framework of the Netherlands Genomics Initiative (NGI). Furthermore, the research leading to the RAAK biobank and the current results has received funding from the Dutch Arthritis Association (DAA 2010_017) and the European Union’s Seventh Framework Programme (FP7/2007-2011) under grant agreement no. 259679. We thank Nico Lakenberg, Ruud van der Breggen and Eka Suchiman, for their help in preparing DNA and RNA samples.

TwinsUK is funded by the Wellcome Trust, Medical Research Council, European Union, the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy’s and St Thomas’ NHS Foundation Trust in partnership with King’s College London.

The Framingham Heart Study of the National Heart, Lung, and Blood Institute of the National Institutes of Health and Boston University School of Medicine was supported by the National Institutes of Health (Contract No. HHSN268201500001I, N01-HC-25195, AG18393, AR47785) and its contract with Affymetrix, Inc. for genotyping services (N02-HL-6-4278). Analyses reflect intellectual input

and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. M.S.Y. is supported by the National Institutes of Aging (T32AG023480).

AUTHOR CONTRIBUTIONS

W.H. and C.G.B. contributed equally to this work. D.H., M.S.Y., Y.F.M.R. and S.M. performed replication analysis for this work, L.B. provided analysis help. L.S.C. and F.R. provided data. M.K. provided phenotypic contribution to the GARP study. M.P. provided data and analyses. T.D.S. contributed data for replication. A.H. contributed data of the RS cohorts. J.D., M.B. and P.E.S contributed to genotyping data and analyses of LLS cohort. R.G.H.H.N. provided contribution to the RAAK study. A.G.U. contributed genotype data of RS cohorts , D.T.F. and A.V. contributed replication data for this work. I.M., and J.B.J.M. jointly supervised this work.

REFERENCES

1 CDC. Prevalence of Doctor-Diagnosed Arthritis and Arthritis-Attributable Activity Limitation — United States, 2010–2012. *MMWR Morb Mortal Wkly Rep* 2013;62:869–73. doi:10.1001/jama.296.22.2671

2 Patel N, Buckland-Wright C. Advancement in the zone of calcified cartilage in osteoarthritic hands of patients detected by high definition macroradiography. *Osteoarthr Cartil* 1999;7:520–5.http://www.idealibrary.com (accessed 2 Nov 2016).

3 Fuerst M, Bertrand J, Lammers L, *et al.* Calcification of articular cartilage in human osteoarthritis. *Arthritis Rheum* 2009;60:2694–703. doi:10.1002/art.24774

4 Nelson AE, DeVellis RF, Renner JB, *et al.* Quantification of the whole-body burden of radiographic osteoarthritis using factor analysis. *Arthritis Res Ther* 2011;13:R176. doi:10.1186/ar3501

5 Jonsson H, Helgadottir GP, Aspelund T, *et al.* Hand Osteoarthritis Severity is Associated with Total Knee Joint Replacements Independently of BMI. The Ages-Reykjavik Study. *Open Rheumatol J* 2011;5:7–12. doi:10.2174/1874312901105010007

6 CDC. Prevalence of Disabilities and Associated Health Conditions Among Adults --- United States, 1999. *MMWR Morb Mortal Wkly Rep* 2009;58:421–6.

7 Spector TD, Cicuttini F, Baker J, *et al.* Genetic influences on osteoarthritis in women: a twin study. *BMJ* 1996;312:940–3. doi:10.1136/bmj.312.7036.940

8 Felson DT. Risk factors for osteoarthritis: understanding joint vulnerability. *Clin Orthop Relat Res* 2004;:S16–21. doi:10.1097/01.blo.0000144971.12731.a2

9 Bijkerk C, Houwing-Duistermaat JJ, Valkenburg H a., *et al.* Heritabilities of radiologic osteoarthritis in peripheral joints and of disc degeneration of the spine. *Arthritis Rheum* 1999;42:1729–35. doi:10.1002/1529-0131(199908)42:8<1729::AID-ANR23>3.0.CO;2-H

10 Ishimori ML, Altman RD, Cohen MJ, *et al.* Heritability patterns in hand osteoarthritis: the role of osteophytes. *Arthritis Res Ther* 2010;12:R180. doi:10.1186/ar3144

11 Kerkhof HJM, Lories RJ, Meulenbelt I, *et al.* A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. *Arthritis Rheum* 2010;62:499–510. doi:10.1002/art.27184

12 Nakajima M, Takahashi A, Kou I, *et al.* New sequence variants in HLA class II/III region associated with susceptibility to knee osteoarthritis identified by genome-wide association study. *PLoS One* 2010;5:e9723. doi:10.1371/journal.pone.0009723

13 Castaño Betancourt MC, Cailotto F, Kerkhof HJ, *et al.* Genome-wide association and functional 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. doi:10.1073/pnas.1119899109

14 Consortium A, Collaborators A. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet* 2012;380:815–23. doi:10.1016/S0140-6736 (12)60681-3

15 Day-Williams AG, Southam L, Panoutsopoulou K, *et al.* A variant in MCF2L is associated with osteoarthritis. *Am J Hum Genet* 2011;89:446–50. doi:10.1016/j.ajhg.2011.08.001

16 Evangelou E, Kerkhof HJ, Styrkarsdottir U, *et al.* A meta-analysis of genome-wide association studies identifies novel variants associated with osteoarthritis of the hip. *Ann Rheum Dis* 2013;:1–7. doi:10.1136/annrheumdis-2012-203114

17 Miyamoto Y, Shi D, Nakajima M, *et al.* Common variants in DVWA on chromosome 3p24.3 are associated with susceptibility to knee osteoarthritis. *Nat Genet* 2008;40:994–8. doi:10.1038/ng.176

18 Castaño-Betancourt MC, Evans DS, Ramos YFM, *et al.* Novel Genetic Variants for Cartilage Thickness and Hip Osteoarthritis. *PLOS Genet* 2016;12:e1006260. doi:10.1371/journal.pgen.1006260

19 Styrkarsdottir U, Thorleifsson G, Helgadottir HT, *et al.* Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene and with rare variants at 1p31. *Nat Genet* 2014;46:498–502. doi:10.1038/ng.2957

20 Kellgren JH, Lawrence JS. Radiological Assessment of Osteo-Arthrosis. *Ann Rheum Dis* 1957;16:494–502. doi:10.1136/ard.16.4.494

21 Ward LD, Kellis M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012;40:930–4. doi:10.1093/nar/gkr917

22 Chadwick LH. The NIH Roadmap Epigenomics Program data resource. *Epigenomics* 2012;4:317–24. doi:10.2217/epi.12.18

23 Kundaje A, Meuleman W, Ernst J, *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* 2015;518:317–30. doi:10.1038/nature14248

24 Kent WJ, Sugnet CW, Furey TS, *et al.* The Human Genome Browser at UCSC. *Genome Res* 2002;12:996–1006. doi:10.1101/gr.229102

25 Goya R, Sun MGF, Morin RD, *et al.* SNVMix: predicting single nucleotide variants from next-generation sequencing of tumors. *Bioinformatics* 2010;26:730–6. doi:10.1093/bioinformatics/btq040

26 Bos SD, Bovée JVMG, Duijnisveld BJ, *et al.* Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. *Ann Rheum Dis* 2012;71:1254–8. doi:10.1136/annrheumdis-2011-200981

27 Kerkhof HJM, Doherty M, Arden NK, *et al.* Large-scale meta-analysis of interleukin-1 beta and interleukin-1 receptor antagonist polymorphisms on risk of radiographic hip and knee osteoarthritis and severity of knee osteoarthritis. *Osteoarthr Cartil* 2011;19:265–71. doi:10.1016/j.joca.2010.12.003

28 Pastinen T. Genome-wide allele-specific analysis: insights into regulatory variation. *Nat Rev Genet* 2010;11:533–8. doi:10.1038/nrg2815

29 Luo G, Ducy P, McKee MD, *et al.* Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein. *Nature* 1997;386:78–81. doi:10.1038/386078a0

30 Yagami K, Suh JY, Enomoto-Iwamoto M, *et al.* Matrix GLA protein is a developmental regulator of chondrocyte mineralization and, when constitutively expressed, blocks endochondral and intramembranous ossification in the limb. *J Cell Biol* 1999;147:1097–108.http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2169349&tool=pmcentrez&rendertype=abstract

31 Nelson MR, Tipney H, Painter JL, *et al.* The support of human genetic evidence for approved drug indications. *Nat Genet* 2015;47:856–60. doi:10.1038/ng.3314

32 Munroe PB, Olgunturk RO, Fryns JP, *et al.* Mutations in the gene encoding the human matrix Gla protein cause Keutel syndrome. *Nat Genet* 1999;21:142–4. doi:10.1038/5102

33 Misra D, Booth SL, Crosier MD, *et al.* Matrix Gla protein polymorphism, but not concentrations, is associated with radiographic hand osteoarthritis. *J Rheumatol* 2011;38:1960–5. doi:10.3899/jrheum.100985

34 Wallin R, Schurgers LJ, Loeser RF. Biosynthesis of the vitamin K-dependent matrix Gla protein (MGP) in chondrocytes: a fetuin-MGP protein complex is assembled in vesicles shed from normal but not from osteoarthritic chondrocytes. *Osteoarthritis Cartilage* 2010;18:1096–103. doi:10.1016/j.joca.2010.05.013

35 Hangoc G, Williams DE, Falkenburg JH, *et al.* Influence of IL-1 alpha and-1 beta on the survival of human bone marrow cells responding to hematopoietic colony-stimulating factors. *J Immunol* 1989;142:4329–34. http://www.ncbi.nlm.nih.gov/pubmed/2656864 (accessed 22 Aug 2016).

36 Gee F, Clubbs CF, Raine EVA, *et al.* Allelic expression analysis of the osteoarthritis susceptibility locus that maps to chromosome 3p21 reveals cis-acting eQTLs at GNL3 and SPCS1. *BMC Med Genet* 2014;15:53. doi:10.1186/1471-2350-15-53

37 Raine EVA, Dodd AW, Reynard LN, *et al.* Allelic expression analysis of the osteoarthritis susceptibility gene COL11A1 in human joint tissues. *BMC Musculoskelet Disord* 2013;14:85. doi:10.1186/1471-2474-14-85

38 Bos SD, Bovée JVMG, Duijnisveld BJ, *et al.* Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. *Ann Rheum Dis* 2012;71:1254–8. doi:10.1136/annrheumdis-2011-200981

39 Schurgers LJ, Uitto J, Reutelingsperger CP. Vitamin K-dependent carboxylation of matrix Gla-protein: a crucial switch to control ectopic mineralization. *Trends Mol Med* 2013;19:217–26. doi:10.1016/j.molmed.2012.12.008

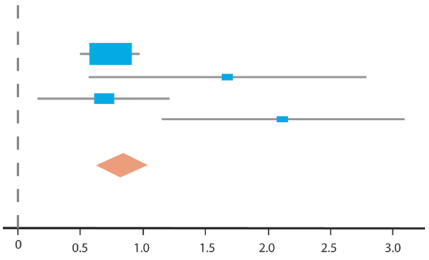
40 Shea MK, Kritchevsky SB, Hsu F-C, *et al.* The association between vitamin K status and knee osteoarthritis features in older adults: the Health, Aging and Body Composition Study. *Osteoarthritis Cartilage* 2015;23:370–8. doi:10.1016/j.joca.2014.12.008

41 Neogi T, Felson DT, Sarno R, *et al.* Vitamin K in hand osteoarthritis: results from a randomised clinical trial. *Ann Rheum Dis* 2008;67:1570–3. doi:10.1136/ard.2008.094771

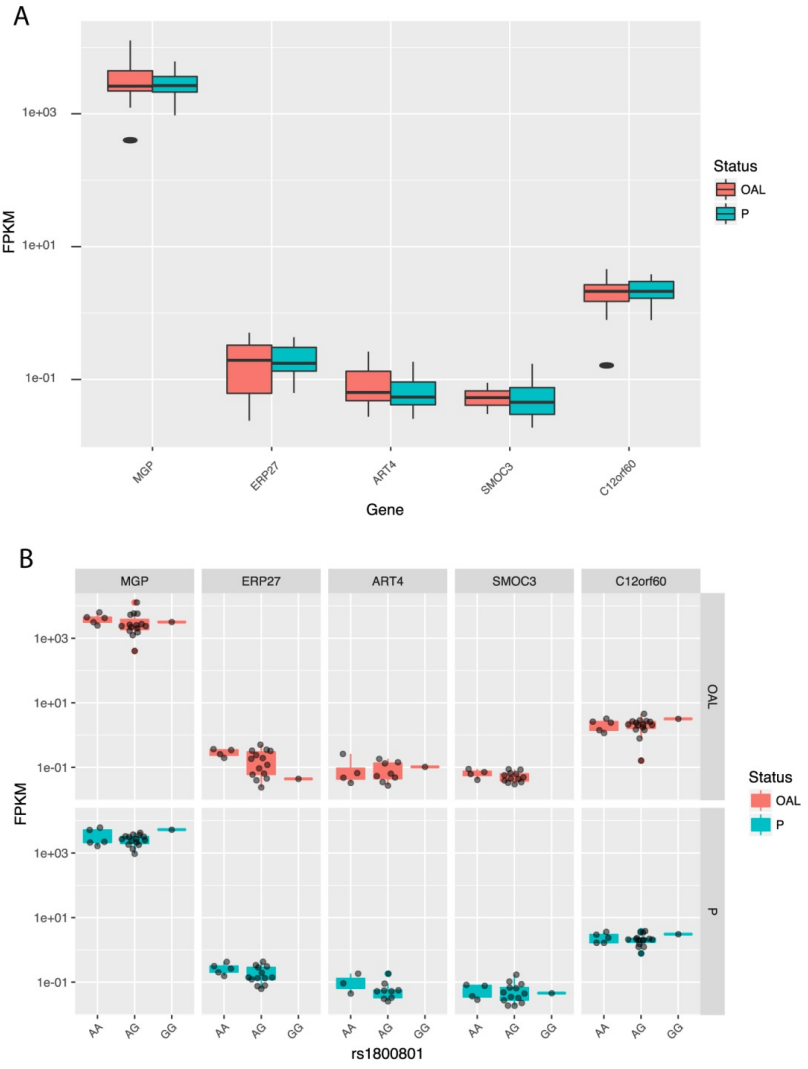
Forrest Plot *rs4764133*

Study name		***n***	**Beta (interval)**
Rotterdam Study (RS1-3)	*(the Netherlands)*	8743	0.75 *(0.63-0.87)*
Leiden Hand OA Studies	*(the Netherlands)*	836	1.68*(1.11 -2.24)*
Twins UK	*(United Kingdom)*	2000	0.69*(0.42-0.96)*
Framingham Heart Study	*(United States of America)*	1175	2.12*(1.62-2.61)*
Total		12754	0.83 *(0.73-0.93)*

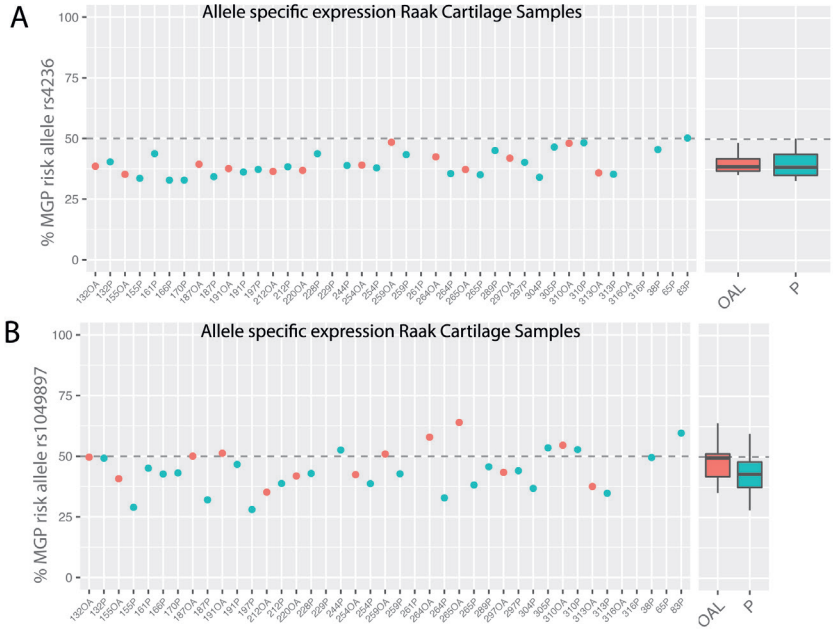
Heterogeneity : Chi²= 63.27, df = 3 (P = 0.02), I²=69.07



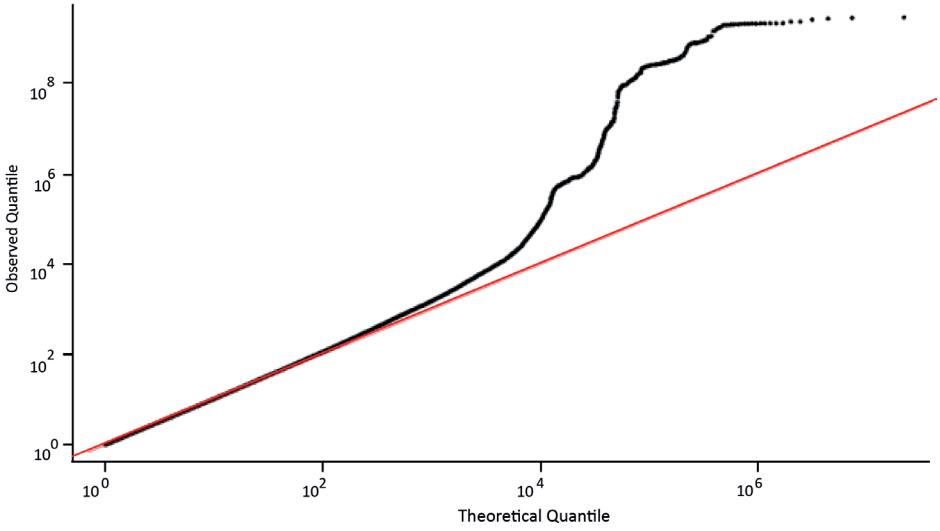
Supplementary Figure S1.



Supplementary Figure S2.



Supplementary Figure S3.



Supplementary Figure S4.

4

OSTEOARTHRITIS SUSCEPTIBILITY POLYMORPHISM RS225014 MODULATES EPIGENETIC REGULATION OF *DIO2* IN ARTICULAR CARTILAGE

W. den Hollander[1,2,‡], N. Bomer[1,3], Y.F.M. Ramos[1,3], S.D. Bos[1,4], R. van der Breggen[1], N. Lakenberg[1], B.A. Pepers[5], A.E. van Eeden[1], E.W. Tobi[1,3], B.J. Duijnisveld[6], B.T. Heijmans[1,3], W.M.C. van Roon-Mom[5], P.E. Slagboom[1,3,4], R.G.H.H. Nelissen[6], I. Meulenbelt[1,2,3,4].

[1]Dept. of Molecular Epidemiology, LUMC, Leiden, The Netherlands; [2]Reumafonds, Dutch Arthritis Foundation; [3]IDEAL; [4]The Netherlands Genomics Initiative, sponsored by NCHA, Leiden-Rotterdam, The Netherlands;[5] Dept. of Human Genetics, LUMC, Leiden, The Netherlands;[6] Dept. of Orthopedics, LUMC, Leiden, The Netherlands.

http://dx.doi.org/10.1136/annrheumdis-2013-204739

ABSTRACT

OBJECTIVES: An earlier observed allelic imbalance and up-regulation of the osteoarthritis (OA) susceptibility gene Type II Deiodinase (*DIO2*) in OA affected articular cartilage indicate that *DIO2* associated OA susceptibility is mediated by aberrant expression of *DIO2*. We set out to characterize possible regulatory properties of putative CCCTC-binding factor (CTCF) sites, proximal and distal CpG dinucleotides in the *DIO2* locus.

METHODS: Electrophoretic Mobility Shift Assays were performed to assess binding of CTCF to the rs225014 locus. OA affected and macroscopically preserved articular cartilage were sampled from end stage OA patients undergoing total joint arthroplasty (RAAK study) of the knee (N = 29) or hip (N = 23) to assess CpG dinucleotide methylation and *DIO2* expression. Chromatin Immunoprecipitation assays were performed to investigate the relation between methylation dependent CTCF binding and *DIO2* expression in three primary chondrocyte cell lines.

RESULTS: OA susceptibility polymorphism rs225014 alleles did not influence local CTCF binding, however a functional CTCF binding site was present 90 base pairs upstream of rs225014. We observed that *DIO2* expression is under epigenetic control of a CpG dinucleotide located in a CTCF binding site 2031 base pairs upstream of the *DIO2* transcription start site. Furthermore, *DIO2* expression in carriers of the rs225014 risk allele showed an enhanced up-regulation upon methylation changes at this CpG dinucleotide, compared to homozygous wildtype carriers.

CONCLUSIONS: *DIO2* associated OA susceptibility is likely to be brought about by aberrant up-regulation of *DIO2* in articular cartilage, due to methylation at a CpG dinucleotide sensitive to the OA process.

INTRODUCTION

Osteoarthritis (OA) is a common disease in the elderly causing pain and disability in articular joints. (1, 2) Although OA is characterized by diseased joint tissues due to wear and tear, it has a significant genetic component. As with many complex genetic diseases, multiple minor genetic defects are considered to contribute to the onset and development of OA. (3) Research aiming to identify determinants conferring osteoarthritis susceptibility, by applying molecular epidemiological approaches, has resulted in a number of successes. Early genetic studies on OA have provided several compelling genes harboring OA susceptibility alleles including the growth differentiation factor 5 gene *GDF5* (4, 5), *SMAD3* (6) and the deiodinase iodothyronine type II and III genes (*DIO2* and *DIO3*), of which the latter two are identified by our own group. (7, 8) In addition, large European consortia (such as arcOGEN and TREAT~OA) have compiled genome wide association (GWA) data and performed meta-analyses on OA and OA-related quantitative traits such as joint space width, which resulted genome wide significant signals. (3, 9-12)

A considerable number of OA susceptibility genes have been found to be part of a common pathways involved in the developmental process of endochondral ossification. (13) Furthermore, it has previously been shown that activation of such genes in articular cartilage occurs during the OA disease process. (14, 15) To ensure cartilage integrity, articular chondrocytes should remain in a differentiated and maturational arrested state; however a hypothesized loss of epigenetic control in articular cartilage could result in reactivation of genes involved in endochondral ossification, leading to loss and mineralization of articular cartilage, a process known to contribute to OA. (7, 16-18)

The linkage signal that identified *DIO2* as OA susceptibility gene was significantly explained by the minor allele (C) of the T>C single nucleotide polymorphism (SNP) rs225014. Rs225014 is a non-synonymous SNP located in exon 2 of *DIO2* causing a Thr92Ala transition. (7) Although this transition does not influence enzyme velocity (19), we have recently reported a highly consistent allelic imbalance (AI) in articular cartilage of the Research and Articular Osteoarthritis Cartilage (RAAK) study, where the OA-associated rs225014 C allele was being 1.3 times more abundantly expressed relative to the T allele. (20) No SNP was found in high linkage disequilibrium (LD) with rs225014 (7) that could explain the AI, therefore it was hypothesized that the observed expression of *DIO2* is directly influenced by rs225014 alleles. A putative CCCTC-binding factor (CTCF) site is overlapping with rs225014, indicative of potential regulatory properties of the locus. (21) CTCF is considered to facilitate long-range chromatin interactions in order to regulate gene expression; distal transcriptional elements on the genome are brought in close proximity to transcriptional start sites (TSSs) of genes to alter expression. (22) We hypothesize that *DIO2* associated OA susceptibility is brought about by aber-

rant up-regulation of *DIO2* expression in articular cartilage. Here we set out to identify genetic and epigenetic elements that regulate *DIO2* expression in articular cartilage and how these are related to rs225014 alleles and the ongoing OA disease process.

MATERIALS AND METHODS

SUBJECTS

Macroscopically, we identified and sampled weight bearing preserved (away from the affected area) and OA affected cartilage (within the affected area) from 52 Caucasian end stage OA patients who underwent a total knee replacement (N = 29) or a total hip replacement (N = 23) in the Leiden University Medical Centre. Ethical approval was obtained from the medical ethics committee of the LUMC (P08.239) and informed consent was obtained from all study participants.

NUCLEIC ACID ISOLATION AND GENOTYPING

Snap frozen cartilage was powderized using a Retsch Mixer Mill 200 with continuous liquid nitrogen cooling. DNA was isolated using the Promega Wizard Genomic DNA Purification kit according to the manufacturer's protocol. RNA was isolated using the Qiagen RNAeasy Mini kits, followed by cDNA synthesis using 1 µg of RNA and random hexamer primers (First Strand cDNA Synthesis Kit, Thermo Scientific). Samples were genotyped for rs225014 using restriction fragment length polymorphism analysis with RsaI (Forward primer (F): 5'-AGTGGCAATGTGTTTAATGTGA-3', Reverse primer (R): 5'-CACACACGTTCAAAGGCTACC-3'). DNA fragment length of wildtype alleles were called after gel electrophoresis and were 121, 30 and 389 base pairs. The risk allele C affects the first cut site resulting in two fragments consisting of 151 and 389 base pairs. All primers were ordered at Invitrogen.

MEASUREMENTS

For Electrophoretic mobility shift assays (EMSAs) synthetic oligonucleotides containing the putative CTCF binding site were 5'-end labeled by γ- (32)P-ATP and subsequently purified by gel filtration on Sephadex G-25 Medium columns. The CTCF-11 zinc finger (11ZF) DNA binding domain, full-length CTCF using pIVEX1.4 WG CTCF-11ZF and CTCF-FL constructs were synthesized with the RTS 100 Wheat Germ CECF kit (5 PRIME). For binding reactions, we used buffer containing standard PBS with 5 mM $MgCl_2$, 0.1 mM $ZnSO_4$, 1 mM DTT, 0.05% NP40, 50 ng/µl poly(dI-dC) and 10% glycerol. The reaction mixtures were incubated for 30 min at RT and analyzed by 5% native PAGE in 0.5x Tris-borate-EDTA buffer. **Supplementary Table 1** contains the sequences of the used probes.

Expression of *DIO2* was assessed using TaqMan probe Hs00988260_m1 (Applied Biosystems), normalized for *GAPDH* expression (real-time PCR, F: 5'-TGCCATGTAGACCCCTTGAAG-3', R: 5'-ATGGTACATGACAAGGTGCGG-3') and subsequently log transformed for downstream analysis.

Using the ZymoResearch EZ DNA Methylation kit isolated genomic DNA was treated with sodium bisulphite (BS), thereby reducing unmethylated cytosine residues to uracil, while methylated cytosines remain unchanged. The methylated fraction of CpG dinucleotides was assessed with MALDI-TOF mass spectrometry (Epityper, Sequenom), a commonly applied to quantify CpG methylation. (23-25) Samples were randomly distributed on PCR plates prior to BS treatment and PCR amplification. PCR amplification and MALDI-TOF measurements were performed in triplicate as technical replicates.

Using MethPrimer 9 amplicons (**Supplementary Table 1**) were designed, covering a total of 23 measureable CpG dinucleotides upstream, downstream and intragenic of *DIO2*. Amplicons were designed to cover conserved transcription factor binding sites (TFBS) according to Human Genome Assembly, build 19. Methylation of several CpG dinucleotides was measured redundantly by separate amplicons, measurements of CpG dinucleotides with the most successful observations per amplicon were used for downstream analyses. Finally, the 9 amplicons constituted 4 independent regions.

CELL CULTURE AND CHROMATIN IMMUNOPRECIPITATION

Primary carticular chondrocytes were isolated from cartilage derived from three OA patients who underwent total joint arthroplasty of the hip (RAAK study). Cartilage tissue was incubated overnight in DMEM (high glucose; Gibco, Bleiswijk, The Netherlands) supplemented with 10% fetal bovine serum (FBS; Gibco), antibiotics (100 U/ml penicillin, 100 µg/ml streptomycin; Gibco) and 2 mg/ml collagenase Type I at 37 ºC in a humidified 5% CO_2 / 95% atmosphere. Subsequently, primary chondrocytes were resuspended and filtered through a 100 µm mesh to remove undigested cartilage fragments and extracellular matrix debris. Cells were expanded at 37 °C in a humidified 5% CO_2 / 95% atmosphere in DMEM supplemented with 10% FBS, penicillin (100 units/mL), streptomycin (100 units/mL) and 0.5 ng/ml FGF-2 (PeproTech, Heerhugowaard, The Netherlands) for 2 passages. 24 Hours into the second passage 1.5 µM of the demethylating agent 5-aza-2 –deoxycytidine (AZA) (Sigma Aldrich; Zwijndrecht, The Netherlands) was added. Cells were harvested for DNA and RNA isolations after being grown to confluence, obtained after three more days.

For chromatin immunoprecipitation (ChIP) chromatin was prepared and immunoprecipitation was performed as previously described. (26) Sonification was optimized to result in ~500 base pair long fragments. For immunoprecipitation rabbit anti-CTCF antibody was used (Anti-CTCF; 07-729 Millipore). Protein A and G sepharose beads were ordered from Sigma (P3391) and GE Healthcare

(17-0618-01) respectively. Finally the immunoprecipitated chromatin was used as input for real-time-PCR to assess the relative binding of CTCF, used primers are listed in **Supplementary Table 1**.

DATA ANALYSIS

Methylation of CpG dinucleotides with fewer than two out of three triplicate measurements or with a SD > 0.1 were discarded prior to analysis. CpG site-containing fragments that had equal or overlapping mass, making them irresolvable by mass spectrometry, and CpG sites containing fragments whose measurement was confounded by SNPs were removed prior to analysis. (27) Samples with bisulphite conversion rates < 98% were discarded.

Principal Component Analysis (PCA) was performed to reduce redundant variation of methylation at CpG dinucleotides which had a high in-between correlation. (27) For each of the 4 regions, Principal Components (PCs) were constructed if the KMO sampling adequacy was < 0.5 and Bartlett's test of sphericity was < 0.01. PCs with Eigen values > 1 were considered relevant. The individual loadings of each of the PCA were checked to assess whether all CpG dinucleotides were covered by the constructed PCs, CpG dinucleotides for which all component loadings were < 0.4 were considered not significant and these CpG dinucleotides were discarded from the PCA and analyzed separately in subsequent analyses.

All statistical analyses were performed by fitting Generalized Linear Mixed Models (GLMMs). To account for inter-individual differences a random effect for sample donor was added to each model. Homozygous carriers (N = 3) of the rs225014 risk allele were pooled with heterozygous carriers (N = 27). Analyses were carried out using the R programming language with the *lme4* (GLMMs) (28), *psych* (29) and *GPArotation* (PCA) (30) packages. To assess the relation between the cartilage phenotype and methylation of separate features, we fitted the following model: $Methylation_i \sim (1|Donor) + Phenotype$. To identify functional CpG dinucleotides, we fitted the following model: $DIO2\ Expression \sim (1|)\ Donor + Methylation_i$. Where, in both models, $Methylation_i$ represents the methylated fraction of the *i*-th CpG feature. To explore other possible significant covariates, we fitted the following model: $DIO2\ Expression \sim (1|Donor) + Joint\ site + rs225014\ alleles + Methylation_{CpG\text{-}2031}$. Finally, p-values were adjusted for multiple testing using Bonfferoni correction.

RESULTS

THE PUTATIVE CTCF BINDING SITE OVERLAPPING RS225014 IS NOT FUNCTIONAL IN ARTICULAR CHONDROCYTES

To assess the regulatory properties of the rs225014 overlapping CTCF binding site, we investigated binding of CTCF at this locus and tested whether rs225014 alleles directly influence local binding of CTCF. We performed EMSAs using full-length CTCF (CTCF-FL) and a truncated protein, containing just the 11 Zinc Finger binding domain (CTCF-11ZF). We have used three different probes containing the rs225014 common allele (T, **Figure 1**, lane 1-2), the minor allele (C, **Figure 1**, lane 3-4) and a non-existent allele (G, **Figure 1**, lane 5-6). The latter is in highest agreement with the consensus CTCF-binding sequence. (31) No band shift was observed for either protein, irrespective of the three rs225014 alleles, indicating that CTCF does not bind to the putative CTCF sequence at SNP rs225014. Nevertheless, the EMSA confirmed a CTCF site 90 base pairs upstream of rs225014 (DIO2-CTCF2, **Figure 1**, lane 7-8), as a clear band shift was observed for this predicted CTCF binding site.

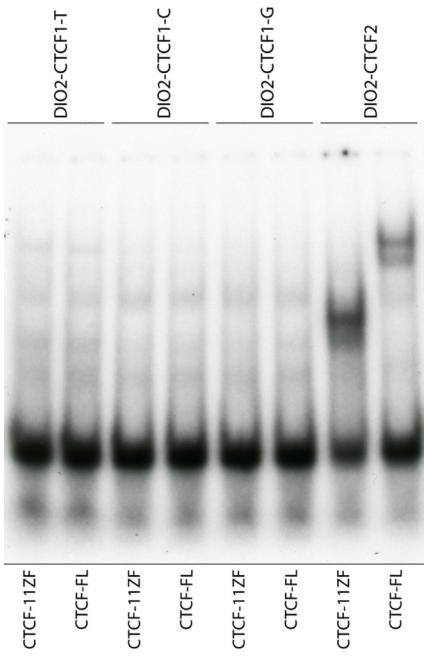


Figure 1. EMSAs for two CTCF binding sites located on (DIO2-CTCF1) and next to (DIO2-CTCF2) rs225014. Full length CTCF (CTCF-FL) and truncated CTCF (CTCF-11ZF) containing just the 11 Zinc Finger binding domain were used. No band shift was observed for DIO2-CTCF1, irrespective of rs225014 alleles (lane 1-6). A band shift for DIO2-CTCF2 was observed for both protein constructs (lane 7-8).

CPG DINUCLEOTIDES LOCATED ACROSS *DIO2* ARE METHYLATED IN ARTICULAR CARTILAGE

Being unable to confirm binding of CTCF at the putative CTCF binding site overlapping the rs225014 locus, we set out to quantify CpG dinucleotide methylation across the *DIO2* locus. We therefore selected macroscopically preserved and OA affected cartilage from patients undergoing total joint arthroplasty of the knee (N = 29) or hip (N = 23) of the RAAK study (**Supplementary Table 2**). Amplicons containing CpG dinucleotides were designed across TFBSs located intragenic, upstream and downstream of the *DIO2* open reading frame. **Figure 2** shows the 4 various regions (A-D) across the *DIO2* locus, methylation measurements of respective CpG dinucleotides per region and the overall mean across pooled samples. Two, four and two CpG dinucleotides were discarded from region A, B and C respectively prior to analysis, due to the inability to discriminate them by mass spectrometry. Region A, B and D showed relatively high levels of methylation, whereas region C, overlapping the *DIO2* promotor, showed overall hypomethylation. Where applicable, we summarized CpG dinucleotides per region in Principal Components (PCs), which resulted in 15 independent variables (5 PCs and 10 individual CpGs), constructed from the original 23 (**Supplementary Table 3** and **Supplementary Table 4**).

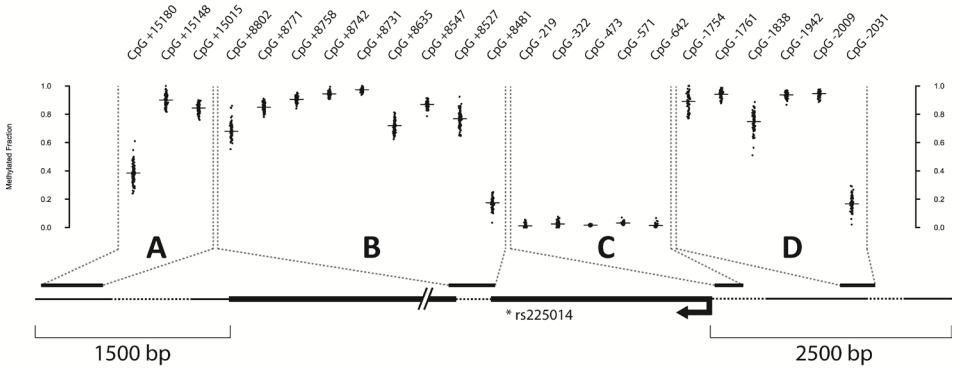


Figure 2. Schematic overview of the *DIO2* open reading frame (ORF) and flanking regions. Lower panel shows the coding strand of the genome (chr14:80,661,868-80,680,970, hg19) with the *DIO2* ORF located on the complementary strand, depicted as a solid bold line. Conserved TFBSs are depicted as dashed lines. CpG dinucleotides are indicated by their relative position to the *DIO2* TSS. (**A-D**) Observed methylation values of CpG dinucleotides in 4 interrogated regions of pooled preserved and OA cartilage samples.

A CPG DINUCLEOTIDE 2031 BASE PAIRS UPSTREAM OF THE *DIO2* TSS (CPG -2031) MODULATES EXPRESSION SIGNIFICANTLY

Upon investigating differential methylation and expression between OA and preserved cartilage, we observed that methylation at multiple CpG features across *DIO2* were responsive to the OA disease process as reflected by significant differential methylation between matched preserved and OA

cartilage samples (**Table 1**). For the single CpG-2031 in region D, we observed the most significant difference in methylation between OA and preserved cartilage (GLMM, N = 103, Beta = 0.028, *P* = 0.0007, Bonferroni adjusted, **Figure 3A**). Suggestive evidence was observed for a difference in *DIO2* expression between preserved and OA affected cartilage, albeit not significant (**Figure 3B**, GLMM, N = 87, Beta = 0.22, *P* = 0.063). To prioritize on functional CpG dinucleotides, we subsequently focused on those features that additionally associated with *DIO2* expression (**Supplementary Table 5**) in cartilage. As seen in **Figure 3C**, we observed a significant positive association between methylation and expression of *DIO2* only for CpG-2031 (GLMM, N = 87, Beta = 4.959, *P* = 0.0016, Bonferroni adjusted).

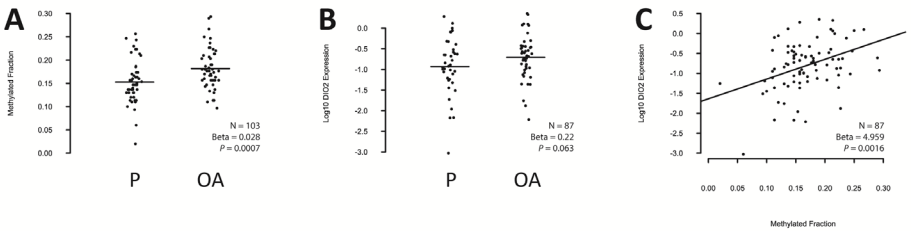


Figure 3. Summary statistics and functional evaluation of methylation at CpG-2031 in articular cartilage. (**A**) A significant difference in methylation was found at CpG-2031 between preserved and OA affected cartilage. (**B**) Suggestive evidence for a difference in *DIO2* expression between preserved and OA affected cartilage. (**C**) Methylation at CpG-2031 correlated significantly with *DIO2* expression.

Table 1. Independent features tested for differential methylation between preserved and OA affected cartilage. Asterisks indicate significant *P*-values after Bonferroni adjustment.

Region	Feature	N	Beta	p-value	Adjusted p-value
A	PC 1	102	0.079	0.56018	1
B	PC 1	95	0.223	0.03132	0.4699
	PC 2	95	-0.152	0.26965	1
	CpG +8802	76	0.001	0.92605	1
	CpG +8742	101	-0.003	0.36266	1
	CpG +8635	103	0.001	0.58860	1
C	CpG-219	94	0.000	0.92106	1
	CpG-322	98	0.002	0.60637	1
	CpG-473	102	0.001	0.15828	1
	CpG-571	97	0.002	0.28463	1
	CpG-642	98	0.003	0.08636	1
D	PC 1	96	-0.520	0.00015	0.0023 *
	PC 2	96	-0.301	0.04685	0.7028
	CpG-1754	63	0.032	0.00449	0.0673
	CpG-2031	103	0.028	0.00005	0.0007 *

To confirm the regulatory properties of CpG-2031 on *DIO2* expression, we performed ChIP assays for a predicted CTCF binding site overlapping the location of CpG-2031. To test whether CpG methylation influences local CTCF binding we treated three primary chondrocyte cell lines, derived from OA patients undergoing total joint arthroplasty of the hip, with 1.5 µM AZA, a widely used demethylating agent. As shown in **Figure 4A**, all cell lines show a decrease in methylation at CpG-2031 that corresponded to an increase in bound CTCF at the respective position (**Figure 4B**) and subsequent down regulation of *DIO2* expression (**Figure 4C**), confirming that methylation at CpG-2031 affects expression of *DIO2*, which may be attributable to local methylation dependent binding of CTCF.

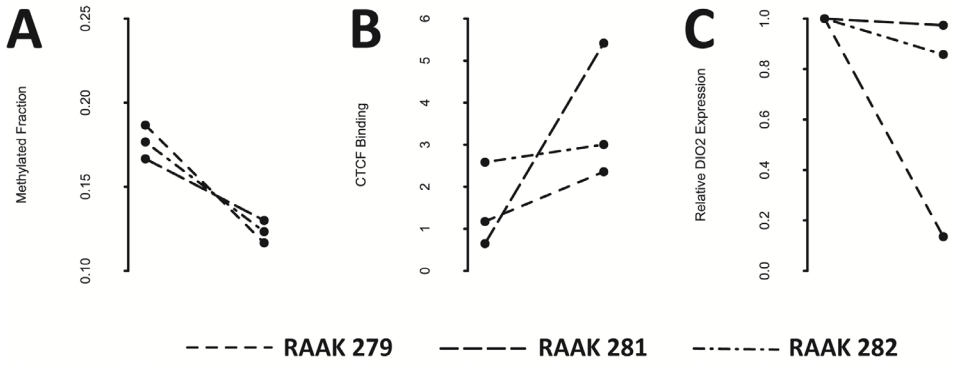


Figure 4. Response of primary chondrocyte cell lines upon treatment with 1.5 µM AZA. (**A**) Decrease of CpG -2031 methylation upon AZA treatment. (**B**) Increase in CTCF binding at CpG-2031 upon AZA treatment. (**C**) Decrease in *DIO2* expression upon AZA treatment.

RS225014 RISK ALLELE CARRIERS SHOW AN ENHANCED CORRELATION BETWEEN CPG -2031 METHYLATION AND *DIO2* EXPRESSION

Having observed the association between *DIO2* expression and OA reflective methylation at CpG -2031, we explored in a multivariate analysis the individual effects of CpG-2031 methylation, joint site and rs225014 alleles as independent variables on *DIO2* expression in articular cartilage as dependent variable. We could ratify the significant association between CpG-2031 methylation and *DIO2* expression (GLMM, N = 87, $Beta_{CpG\ -2031}$ = 4.526, $P_{CpG-2031}$ = 0.0006) and observed a slight effect for joint site (GLMM, N = 87, $Beta_{joint}$ =-0.264, P_{joint} = 0.048). However, the most compelling observation was the significant effect of rs225014 alleles on *DIO2* expression (GLMM, N = 87, $Beta_{rs225014}$ = 0.448, $P_{rs225014}$ = 0.0003), independently of methylation at CpG-2031.

Finally, to further elucidate these effects, we explored whether the earlier observed results were affected by stratification of the rs225014 risk allele. We observed an increase in methylation (**Figure 5A**, GLMM, N = 59, Beta = 0.034, *P* = 0.00002) and expression (**Figure 5B**, GLMM, N = 51, Beta = 0.35, *P* = 0.012) differences between preserved and OA affected cartilage in carriers of the rs225014

risk allele. Most strikingly, expression of *DIO2* in cartilage derived from rs225014 risk allele carriers was more responsive to methylation changes at CpG-2031 (**Figure 5C**, GLMM, N = 51, Beta = 5.58, *P* = 0.0006). Together, these data indicate that *DIO2* expression among carriers of the rs225014 risk allele, compared to homozygous common allele carriers, is more sensitive to subtle methylation changes at CpG-2031 in articular cartilage.

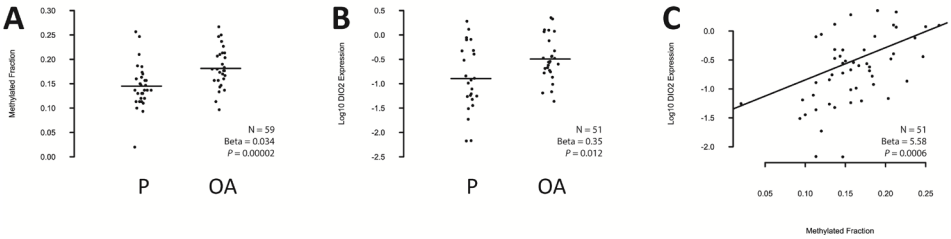


Figure 5. (**A**) A more pronounced difference between preserved and OA affected cartilage was observed in carriers of the rs225014 risk allele. (**B**) *DIO2* expression in carriers of the rs225014 risk allele is significantly higher in OA affected cartilage compared to preserved cartilage. (**C**) The functional relation between CpG-2031 and *DIO2* expression is more pronounced in carriers of the rs225014 risk allele.

DISCUSSION

In the current study, we have identified a CpG dinucleotide at 2031 base pairs upstream of the *DIO2* TSS that modulated *DIO2* expression significantly (**Figure 3C**) and appeared highly sensitive to the ongoing OA process (**Table 1** and **Figure 3A**). Furthermore, by means of ChIP analyses it was shown that the regulatory properties of methylation at CpG-2031 are most likely to be exerted via methylation dependent binding of CTCF at the respective position (**Figure 4**). Moreover, regarding epigenetically regulated *DIO2* expression, we showed that subtle differential methylation at this particular CpG dinucleotide associated with up regulation of *DIO2* expression, most vividly among rs225014 risk allele carriers (**Figure 5C**). Given the role of *DIO2* in growth plate chondrocytes, i.e. the initiation of endochondral ossification (32, 33), such an up-regulation should be considered detrimental to articular cartilage integrity and increasing the propensity to develop OA. Together these data provide compelling insights into how the *DIO2* OA susceptibility allele of rs225014 confers risk to OA; upon repeated challenges throughout life, chondrocytes residing in articular cartilage of rs225014 risk allele carriers are less able to maintain cartilage homeostasis due to the fact that subtle changes in methylation at CpG-2031 result in detrimental up-regulation of *DIO2*.

In previous studies we have demonstrated an early mode of action of *DIO2* susceptibility alleles on hip morphology (34), whereas in articular cartilage we have reported on up regulation of DIO2

protein in OA as compared to preserved cartilage and a highly consistent AI of *DIO2* in heterozygous carriers of the rs225014 risk allele. (20) In the current study, we have prioritized on CpG dinucleotides that significantly modulated expression of *DIO2* and have provided insight into the mechanisms how the rs225014 risk allele confers risk to OA at later ages. Nevertheless, we were unable to unravel the molecular mechanism of the AI marked by the rs225014 alleles. This implies that *DIO2* expression in articular chondrocytes is subject to multiple regulatory mechanisms, possibly intertwined, as is suggested by the vivid *DIO2* up-regulation upon CpG-2031 methylation changes in rs225014 risk allele carriers.

Demethylation of CpG dinucleotides commonly up-regulates gene expression, however, here a positive correlation between methylation at CpG-2031 and *DIO2* expression was observed. This positive correlation is likely to be mediated via the confirmed methylation dependent binding of CTCF, a general insulator protein known to facilitate regulation of gene expression, at CpG-2031. We hypothesize that a putative three dimensional structure, brought on about by bound CTCF in hypomethylated samples, exerts inhibitory effects on *DIO2* expression, possibly via recruitment of inhibitory transcription factors.

A significant up-regulation of *DIO2* mRNA and protein in OA affected cartilage compared to healthy controls is reported in literature (14, 15, 20), however, when comparing preserved with OA affected cartilage we were unable to detect a significant difference in *DIO2* expression. Hence, it is likely that preserved cartilage from OA affected joints has already to some extent been seized by the ongoing OA disease process, but has not yet manifested in macroscopically observable differences. However, we did observe a significant difference in expression between preserved and OA affected cartilage in carriers of the rs225014 risk allele. Even so, our approach does minimize environmental influences, to which CpG methylation is highly susceptible (35-37), as the preserved and matched OA affected cartilage originate from a single joint, thereby allowing robust pair wise comparisons. To study the epigenetic mechanisms underlying the reported up-regulation of *DIO2* mRNA in OA affected cartilage compared to unaffected cartilage, CpG dinucleotide methylation in true healthy controls should be quantified. However, appropriate sample sizes are required, as large variation in methylation of CpG dinucleotides is observed among subjects.

It has been shown that regulatory elements on the genome, such as CpG dinucleotides, quite commonly do not influence gene expression of the nearest gene (38); i.e. differential methylation of CpG dinucleotides does not imply an effect on gene expression of the nearest gene whatsoever (**Table 1** and **Supplementary Table 5**). Therefore, in the view of the authors, functional follow up by means of epigenetic regulation of OA susceptibility genes, should have a focus on CpG dinucleotides correlating significantly with expression of the gene of interest.

In conclusion, we have shown that a CpG dinucleotide 2031 base pairs upstream of the *DIO2* promoter is differentially methylated between preserved and OA affected cartilage. The functionality of this CpG dinucleotide is illustrated by local methylation dependent CTCF binding and, moreover, by a significant correlation with *DIO2* expression. Furthermore, compelling evidence was observed for an enhanced response of *DIO2* expression upon methylation changes at CpG-2031 in rs225014 risk allele carriers in articular cartilage.

ACKNOWLEDGEMENTS:

The research leading to these results has received funding from the Dutch Arthritis Association under Project NR-10-1-402 and the European Union's Seventh Framework Programme (FP7/2007-2011) under Grant Agreement No. 259679. The genotypic work was supported by the Netherlands Organization of Scientific Research (MW 904-61-095, 911-03-016, 917 66344 and 911-03-012), Leiden University Medical Centre and the Centre of Medical System Biology and Netherlands Consortium for Healthy Aging in the framework of the Netherlands Genomics Initiative (NGI). The RAAK study was supported by the Leiden University Medical Center, the Dutch Arthritis Association, the Centre of Medical System Biology and the Netherlands Consortium for Healthy Ageing in the framework of the NGI.

REFERENCES

1 Nuesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Juni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. *BMJ* 2011;342:d1165.

2 Valdes AM, Spector TD. The clinical relevance of genetic susceptibility to osteoarthritis. *Best Pract Res Clin Rheumatol* 2010;24:3-14.

3 Panoutsopoulou K, Southam L, Elliott KS, Wrayner N, Zhai G, Beazley C, *et al.* Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study. *Ann Rheum Dis* 2011;70:864-7.

4 Chapman K, Takahashi A, Meulenbelt I, Watson C, Rodriguez-Lopez J, Egli R, *et al.* A meta-analysis of European and Asian cohorts reveals a global role of a functional SNP in the 5' UTR of GDF5 with osteoarthritis susceptibility. *Hum Mol Genet* 2008;17:1497-504.

5 Evangelou E, Chapman K, Meulenbelt I, Karassa FB, Loughlin J, Carr A, *et al.* Large-scale analysis of association between GDF5 and FRZB variants and osteoarthritis of the hip, knee, and hand. *Arthritis Rheum* 2009;60:1710-21.

6 Valdes AM, Spector TD, Tamm A, Kisand K, Doherty SA, Dennison EM, *et al.* Genetic variation in the SMAD3 gene is associated with hip and knee osteoarthritis. *Arthritis Rheum* 2010;62:2347-52.

7 Meulenbelt I, Min JL, Bos S, Riyazi N, Houwing-Duistermaat JJ, van der Wijk HJ, *et al.* Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. *Hum Mol Genet* 2008;17:1867-75.

8 Meulenbelt I, Bos SD, Chapman K, van der BR, Houwing-Duistermaat JJ, Kremer D, *et al.* Meta-analyses of genes modulating intracellular T3 bio-availability reveal a possible role for the DIO3 gene in osteoarthritis susceptibility. *Ann Rheum Dis* 2011;70:164-7.

9 Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet* 2012;380:815-23.

10 Day-Williams AG, Southam L, Panoutsopoulou K, Rayner NW, Esko T, Estrada K, *et al.* A variant in MCF2L is associated with osteoarthritis. *Am J Hum Genet* 2011;89:446-50.

11 Kerkhof HJ, Lories RJ, Meulenbelt I, Jonsdottir I, Valdes AM, Arp P, *et al.* A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. *Arthritis Rheum* 2010;62:499-510.

12 Castano Betancourt MC, Cailotto F, Kerkhof HJ, Cornelis FM, Doherty SA, Hart DJ, *et al.* Genome-wide association and functional 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.

13 Bos SD, Slagboom PE, Meulenbelt I. New insights into osteoarthritis: early developmental features of an ageing-related disease. *Curr Opin Rheumatol* 2008;20:553-9.

14 Ijiri K, Zerbini LF, Peng H, Otu HH, Tsuchimochi K, Otero M, *et al.* Differential expression of GADD45beta in normal and osteoarthritic cartilage: potential role in homeostasis of articular chondrocytes. *Arthritis Rheum* 2008;58:2075-87.

15 Karlsson C, Dehne T, Lindahl A, Brittberg M, Pruss A, Sittinger M, *et al.* Genome-wide expression profiling reveals new candidate genes associated with osteoarthritis. *Osteoarthritis Cartilage* 2010.

16 Reynard LN, Bui C, Canty-Laird EG, Young DA, Loughlin J. Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. *Hum Mol Genet* 2011;20:3450-60.

17 Young DA, Bui C, Barter MJ. Understanding CpG methylation in the context of osteoarthritis. *Epigenomics* 2012;4:593-5.

18 Young DA. More evidence for a role of CpG methylation in the pathogenesis of osteoarthritis. *Arthritis Rheum* 2012.

19 Canani LH, Capp C, Dora JM, Meyer EL, Wagner MS, Harney JW, *et al.* The type 2 deiodinase A/G (Thr92Ala) polymorphism is associated with decreased enzyme velocity and increased insulin resistance in patients with type 2 diabetes mellitus. *J Clin Endocrinol Metab* 2005;90:3472-8.

20 Bos SD, Bovee JV, Duijnisveld BJ, Raine EV, van Dalen WJ, Ramos YF, *et al.* Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. *Ann Rheum Dis* 2012.

21 Rosenbloom KR, Dreszer TR, Long JC, Malladi VS, Sloan CA, Raney BJ, *et al.* ENCODE whole-genome data in the UCSC Genome Browser: update 2012. *Nucleic Acids Res* 2012;40:D912-D917.

22 Holwerda S, de LW. Chromatin loops, gene positioning, and gene expression. *Front Genet* 2012;3:217.

23 Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, *et al.* Quantitative high-throughput

analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci U S A* 2005;102:15785-90.
24 Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, *et al.* Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *FASEB J* 2010.
25 Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, *et al.* DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum Mol Genet* 2009;18:4046-53.
26 Ramos YF, Hestand MS, Verlaan M, Krabbendam E, Ariyurek Y, van GM, *et al.* Genome-wide assessment of differential roles for p300 and CBP in transcription regulation. *Nucleic Acids Res* 2010;38:5396-408.
27 Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE. Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. *Hum Mol Genet* 2007;16:547-54.
28 Douglas Bates, Martin Maechler, Ben Bolker. lme4: Linear mixed-effects models using S4 classes. 2012.
29 Williams R. psych: Procedures for Psychological, Psychometric and Personality Reasearch. 2012.
30 Bernaards CA, Jennrich RI. Gradient Projection Algorithms and Software for Arbitrary Rotation Criteria in Factor Analysis. 65[Educational and Psychological Measurement], 676-696. 2005.
31 Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, *et al.* Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 2007;128:1231-45.
32 Robson H, Siebler T, Stevens DA, Shalet SM, Williams GR. Thyroid hormone acts directly on growth plate chondrocytes to promote hypertrophic differentiation and inhibit clonal expansion and cell proliferation. *Endocrinology* 2000;141:3887-97.
33 Miura M, Tanaka K, Komatsu Y, Suda M, Yasoda A, Sakuma Y, *et al.* Thyroid hormones promote chondrocyte differentiation in mouse ATDC5 cells and stimulate endochondral ossification in fetal mouse tibias through iodothyronine deiodinases in the growth plate. *J Bone Miner Res* 2002;17:443-54.
34 Waarsing JH, Kloppenburg M, Slagboom PE, Kroon HM, Houwing-Duistermaat JJ, Weinans H, *et al.* Osteoarthritis susceptibility genes influence the association between hip morphology and osteoarthritis. *Arthritis Rheum* 2011;63:1349-54.
35 Milagro FI, Campion J, Cordero P, Goyenechea E, Gomez-Uriz AM, Abete I, *et al.* A dual epigenomic approach for the search of obesity biomarkers: DNA methylation in relation to diet-induced weight loss. *FASEB J* 2011;25:1378-89.
36 Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T, *et al.* Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci Transl Med* 2010;2:49ra67.
37 Byun HM, Nordio F, Coull BA, Tarantini L, Hou L, Bonzini M, *et al.* Temporal stability of epigenetic markers: sequence characteristics and predictors of short-term DNA methylation variations. *PLoS ONE* 2012;7:e39220.
38 Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. *Nature* 2012;489:109-13.

Supplementary tables can be found at doi: 10.1136/annrheumdis-2013-204739

Supplementary Table 1.
Used EMSA probes and Epityper primers.

Supplementary Table 2.
Characteristics of samples used in methylation and expression assessment.

Supplementary Table 3.
PCA eigenvalues per region. PCs with eigenvalues < 1 were discarded.

Supplementary Table 4.
PCA loadings for each region. A dash indicates inappropriate KMO sampling adequacy or failed Bartlett's test for sphericity, these CpG dinucleotides were analyzed separately.

Supplementary Table 5.
Independent features tested for their association with *DIO2* expression in articular cartilage. Asterisks indicate significant *P*-values after Bonferroni adjustment.

5

TRANSCRIPTIONAL ASSOCIATIONS OF OSTEOARTHRITIS MEDIATED LOSS OF EPIGENETIC CONTROL IN ARTICULAR CARTILAGE

Wouter den Hollander[1], Yolande F.M. Ramos[1], Nils Bomer[1,2], Stefan Elzinga[1], Ruud van der Breggen[1], Nico Lakenberg[1], Wesley J. de Dijcker[1], Eka D. Suchiman[1], Bouke J. Duijnisveld[4], Jeanine J. Houwing-Duistermaat[5], Eline P. Slagboom[1,2,3], Steffan D. Bos[1,3], Rob G.H.H. Nelissen[4], Ingrid Meulenbelt[1,3].

[1]Dept. of Molecular Epidemiology, LUMC, Leiden, The Netherlands. [2]IDEAL The Netherlands. [3]Genomics Initiative, sponsored by the NCHA, Leiden, The Netherlands. [4]Dept. of Orthopaedics, LUMC, Leiden, The Netherlands.[5]Dept. of Biostatistics and Bioinformatics, LUMC, Leiden, The Netherlands.

https://doi.org/10.1002/art.39162

ABSTRACT

Objective To identify osteoarthritis (OA) progression modulating pathways in articular cartilage and their respective regulatory epigenetic and genetic determinants in end stage disease.

Methods Transcriptional activity of CpGs (t-CpGs) was assessed using gene expression and DNA methylation data of respectively 33 and 31 pairs of preserved and lesioned articular cartilage. Disease responsive t-CpGs were identified by means of differential methylation between preserved and lesioned cartilage. Transcriptionally relevant genetic determinants were addressed by means of proximal SNPs near the OA responsive t-CpGs. Statistical analyses were corrected for age, sex, joint and technical covariates. A random effect was included to correct for possible correlations between paired samples.

Results Of the 9838 transcribed genes in articular cartilage, 2324 correlated with the methylation status of 3748 t-CpGs, both negative (N=1741) and positive (N=2007) correlations were observed. Hypomethylation and hypermethylation (FDR<0.05, $|\Delta\beta|>0.05$) were observed for 62 and 25 t-CpGs, respectively, covering 70 unique genes. Enrichment for developmental and ECM maintenance pathways indicated possible reactivation of endochondral ossification. Finally, we observed 31 and 26 genes of which, respectively, methylation and expression was additionally affected by genetic variation.

Conclusion We identified tissue specific genes involved in OA disease progression, reflected by genetic and pathological epigenetic regulation of transcription, primarily at genes involved in development. Therefore, transcriptionally active SNPs near these genes may serve as putative susceptibility alleles. Our results comprise an important step in understanding the reported widespread epigenetic changes occurring in OA affected articular cartilage and subsequent development of future treatments targeting disease driving pathways.

INTRODUCTION

Osteoarthritis (OA) is the most prevalent arthritic disease among the elderly (1) and is currently recognized as a disease of the whole joint. (2) A well described hallmark of OA is articular cartilage degradation. (3) The single cell type present in articular cartilage is the articular chondrocyte, which is a highly specialized, maturational arrested, non-proliferating cell. To ensure articular cartilage integrity throughout life it needs to adapt its behaviour in response to external signals, such as mechanical stress, ageing or micro-traumas. (4) To facilitate these adaptations, the chondrocyte requires phenotypic plasticity with proper dynamic control of gene expression to shift between active metabolic and maturational arrested states. In this respect, chondrocytes in OA affected cartilage were shown to have lost their maturational arrested state, regain growth-plate morphology and start to proliferate, while degrading and calcifying the articular cartilage matrix. (5;6)

A likely candidate for maintaining the chondrocyte phenotype is through epigenetic control of gene expression, such as DNA methylation; a biochemical process which is utilized by cells to adapt to environmental challenges such as age or disease by dynamic control of gene expression. (7;8) In this respect, methylome wide studies of articular cartilage in OA have revealed numerous differentially methylated loci between healthy and diseased tissue, while only a small minority of these loci were subsequently studied in terms of gene expression differences. (9-12) Therefore, up to now it remains unclear to what extent the large number of differentially methylated CpGs in OA confer relevant gene expression changes in articular cartilage. Moreover, a growing body of literature describes how aberrant gene expression is influenced, in addition to DNA methylation, by genetic risk alleles in complex genetic diseases, a mechanism outlines previously in OA. (13-17) These reports imply the need for combining multiple levels of genome wide data to gain a more robust understanding of the transcriptional processes that occur with complex genetic diseases, such as OA.

In a previous study we have described functional DNA methylation differences between knee and hip articular cartilage, independent of OA pathophysiology. (10) Although the entire epigenomic profile of knee and hip articular cartilage is primarily defined by differentially methylated regions (DMRs) between the two joints, literature suggests highly gene specific DNA methylation changes in association with OA onset and progression. (9;11;12;14-17) Therefore, in the current study we set out to identify gene specific DNA methylation differences, independent of the joint, between pre-served and lesioned cartilage in patients undergoing total joint replacement surgery due to primary end-stage OA. Moreover, we combined DNA methylation changes with a previously assessed gene expression dataset of overlapping samples (18) to assess OA related changes in the epigenetically regulated transcriptome. Finally, by integrating the results with genome wide single nucleotide poly-

morphism (SNP) data, we aim to identify OA relevant, tissue specific genetic variants that influence gene expression in articular cartilage. The applied consecutive stepwise approach will provide novel OA susceptibility genes, as well as the respective transcriptional determinants. To our knowledge this is the first study comprehensively combining genetic, epigenomic and transcriptomic data to gain a functional understanding of joint independent DNA methylation changes in relation to OA pathophysiology.

MATERIALS AND METHODS

THE RAAK COHORT.

Ethical approval was obtained from the medical ethics committee of the LUMC (P08.239) and informed consent was obtained from all participants. Participant details are listed in **Supplementary Table S1**. For sampling details see (10;14;18).

METHYLATION DATA.

Methylation data were obtained and processed as previously. (10) In short, DNA was isolated using the Promega Wizard Genomic DNA Purification kit according to the manufacturer's protocol. Next, the DNA was bisulphite treated using the ZymoResearch EZ DNA Methylation kit. DNA methylation was assessed using Illumina Infinium HumanMethylation450 BeadChips. Samples were randomly dispersed, while sample pairs were assured to be on the same chip. Using the *minfi* and *lumi* R-packages the methylation dataset was filtered for probes that contained SNPs or mapped ambiguously to the genome and colour channels were separately quantile normalized. Validation and replication using the EpiTYPER platform were done so as previously reported, (14) primer sequences are listed in **Supplementary Table S2**.

EXPRESSION DATA.

Normalized expression data from the RAAK study were processed and normalized as described previously (GSE57218). (18) RT-qPCR validation primers are listed in **Supplementary Table S2**

GENOTYPE DATA.

Using Illumina HumanOmniExpressExome chips genome wide genotyping data was constructed for 216 samples from the RAAK study. SNPs with <95% call rate, Hardy-Weinberg equilibrium $<10^{-4}$, minor allele frequency <0.01 or located on the sex chromosomes were removed prior to imputation together with Leiden Longevity Study data against the 1000 Genomes V3 March 2012 reference panel. (19) Next, SNPs that were homozygous in all the overlapping samples with the methylation

(N=23) and expression (N=24) datasets were removed prior to analyses, as were SNPs of which the imputation quality of 0.4 was not met. (20;21)

STATISTICAL ANALYSES.

All statistical procedures were carried out in R-3.0.2. Analyses were corrected for technical covariates as well as sex, joint and age. To correct for putative correlations between preserved and OA affected AC from the same joint, a random effect for patient ID was included using the *lme4* package. (22) Correction for multiple testing was performed using the bonferroni procedure per gene in the functional methylation analysis as well as the genetic analysis, all other multiple testing corrections were performed using the Benjamini-Holm method. Methylation measurements are reported as β-values. (10;23) CpGs were considered differentially methylated when the mean paired difference was at least 0.05β, as smaller differences would be hard to address statistically and/or interpret biologically. Pathway enrichment was performed using the online annotation tools DAVID and STRING-DB. A full analysis summary scheme is shown in **Figure 1.**

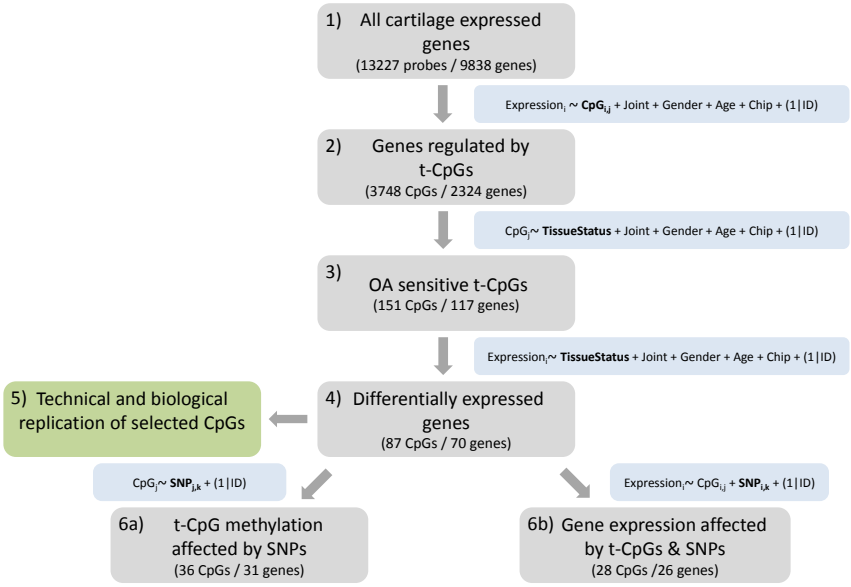


Figure 1. Overview of the applied analysis strategy.

RESULTS

TRANSCRIPTIONALLY ACTIVE CPG DINUCLEOTIDES (T-CPGS) IN ARTICULAR CARTILAGE.

Recently, we have assessed the late stage transcriptomic profile of articular cartilage of patients who underwent total joint replacement surgery, due to primary OA (GSE57218, **Supplementary Table S1**). (18) For the 13277 probes (covering 9838 unique genes, **Figure 1, step 1**) that were expressed to detectable extent in articular cartilage, we set out to explore whether they associated with DNA methylation of proximal CpGs. To identify articular cartilage relevant CpGs in terms of transcriptional association, DNA methylation data of CpGs within 10kb of annotated genes was correlated to respective gene expression data of 13 sample pairs (4 knees and 9 hips) of preserved and lesioned articular cartilage. After multiple testing correction for the number of CpGs for each individual gene, we observed 3748 CpGs that significantly correlated with proximal gene expression, covering a total of 2324 unique genes (24%, **Supplementary Table S3, Figure 1, step 2**), hereafter indicated as transcriptionally active CpGs (t-CpGs). Notably, both negative (e.g. *SPINT2*, *CILP*, *BFSP1*, *TMEM140*, **Figure 2A-D**) as well as positive correlations (e.g. *COL1A2*, *THBS2*, *MSX1*, *RUNX3*, **Figure 2E-H**) were observed for the, respectively, 1741 and 2007 t-CpGs.

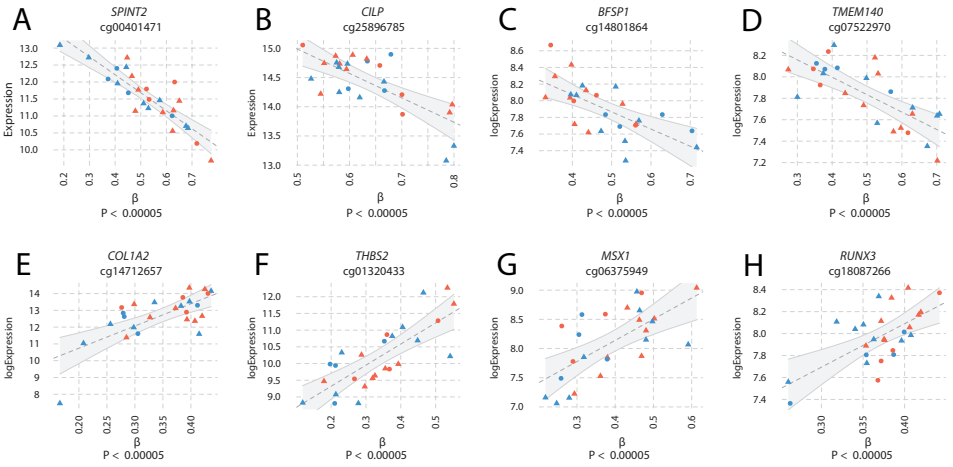


Figure 2. Examples of epigenetically regulated cartilage expressed genes by DNA methylation. Preserved and lesioned samples are respectively coloured blue and red, while knee and hip samples are respectively depicted as circles and triangles. In grey the 95% confidence intervals are plotted. (**A-D**) Examples of down regulated genes upon increased methylation. (**E-H**) Examples of genes of which expression is positively correlated to increased DNA methylation.

OA ASSOCIATED DNA METHYLATION CHANGES AT T-CPGS.

Next, we determined in a total of 31 sample pairs (17 knees and 14 hips) which of the detected 3748 t-CpGs were sensitive to the ongoing OA disease process, as reflected by differential methylation between paired preserved and lesioned cartilage. In total we observed 5282 differentially methylated CpGs (FDR<0.05, |Δβ|>0.05, **Figure 3A, Supplementary Table S4**), of which 2188 and 3094 were respectively hyper- and hypomethylated. Among these OA associated CpGs 151 overlapped with the observed t-CpGs, covering a total of 117 unique genes (**Figure 1, step 3**). Hypermethylation was observed in 59 OA responsive t-CpGs, while hypomethylation was seen in 92 OA responsive t-CpGs covering respectively 46 and 75 genes. Among those are genes known to be involved in OA pathophysiology (e.g. *FOXA2*, *RUNX1*, *COL6A3* and *CD44*, **Figure 3B-E**), as well as multiple genes not earlier reported (e.g. *UACA*, *DLX5*, *DYSF* and *IGFBP7*, **Figure 3F-I**). Next, to focus solely on genes of which expression is involved in OA progression, we selected t-CpG regulated genes of which expression was additionally significantly different between preserved and lesioned tissue. As of such, we continued with 25 and 62 respectively hyper- and hypomethylated t-CpGs, covering 70 unique genes (**Figure 1, step 4**). Subsequent gene enrichment analysis revealed significant enrichment among the 70 genes for pathways earlier reported to be implicated in OA pathophysiology, such as ECM maintenance and developmental processes (**Supplementary Table S5** and **Supplementary Figure S1**).

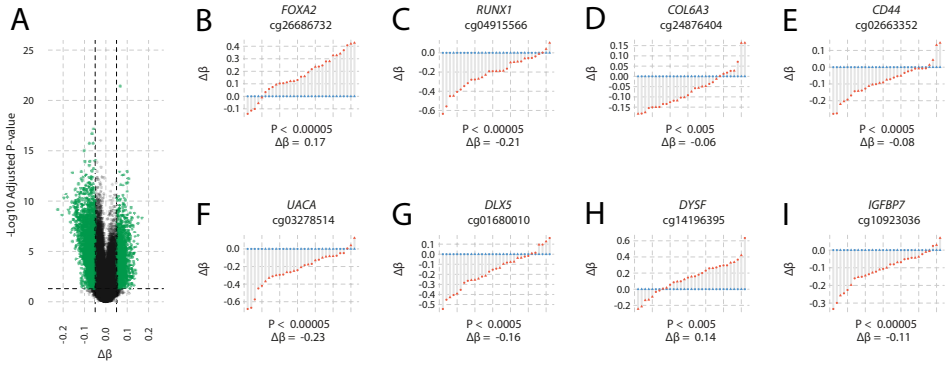


Figure 3. (**A**) Volcano plot showing the cut-offs taken to identify all differentially methylated CpGs between preserved and lesioned cartilage, significant (FDR < 0.05) differentially methylated (|Δβ| > 0.05) CpGs are depicted as green dots. (**B-E**) Significant differential methylation between preserved and lesioned cartilage in known OA associated genes. (**F-I**) Significant differential methylation between preserved and lesioned cartilage in genes not earlier implicated in OA. Preserved samples are set to 0 and depicted in blue, while its paired lesioned sample is depicted in red. Knee and hip joints are respectively shown as circles and triangles.

TECHNICAL VALIDATION AND BIOLOGICAL REPLICATION OF OA RESPONSIVE T-CPG DINUCLEOTIDES.

Using the EpiTYPER® platform, a commonly used technique for measuring DNA methylation, (14) we set out to technically validate 8 CpGs in 17 pairs of preserved and lesioned samples. We found a high degree of similarity between the two techniques, reflected by large Pearson correlation coefficients (mean r>0.85, **Supplementary Figure S2, Figure 1, step 5**). Next, we addressed the previously observed relation between the replicated CpGs and respective gene expression. Except for *IGFBP7*, we were able to validate the transcriptional involvement of all selected t-CpGs and/or disease associated dysregulation of the respective gene (**Supplementary Figure S3**). For biological validation, DNA methylation of the selected CpGs was measured in an additional 31 pairs of preserved and lesioned cartilage. All CpGs showed highly similar, significant DNA methylation changes as were seen in both the discovery and validation samples (**Supplementary Figure S3**).

THE INFLUENCE OF GENETIC FACTORS ON T-CPG METHYLATION AND EXPRESSION IN ARTICULAR CARTILAGE.

Finally, we investigated the stable regulatory genetic environment, as reflected by transcriptionally active SNPs in proximity of the 70 genes. Presence of such SNPs may causally affect cartilage homeostasis of epigenetically controlled genes and confer potential OA susceptibility. The genotypes of all SNPs (dbSNP build 138) 10kb up- and downstream of the 70 genes were assessed in 23 sample pairs of the methylation dataset. Using multivariate analysis with methylation as dependant variable, we identified 36 OA responsive t-CpGs that are significantly affected by at least one SNP (**Supplementary Table S6, Figure 1, step 6a**), covering 31 unique genes. In parallel, using multivariate analysis with expression as dependent variable, we explored whether epigenetic regulation of the 70 genes was additionally affected by the alleles of proximal SNPs. As of such, we observed 26 genes of which expression was modulated by the local genetic background in conjunction with 28 t-CpGs (**Supplementary Table S7, Figure 1, step 6b**). For the *ESR*, *NAV2* and *WLS*, we observed t-CpGs and SNPs that modulated gene expression jointly.

As example, three notable genes of which we have observed transcriptomic, epigenetic and genetic involvement in OA progression are *VIT* (**Figure 4A-D**), *ROR2* (**Figure 4E-H**) and *WLS* (**Figure 4I-M**). All genes were differentially expressed between preserved and lesioned cartilage (**Figure 4A**, **4E**, **4I**), which was modulated by differential DNA methylation (**Figure 4B**, **4F**, **4J**), also reflected by significant differential methylation at the respective t-CpGs (**Figure 4C**, **4G**, **4K**). Moreover, rs11884419 and rs13292198 influenced gene expression and t-CpG methylation of *ROR2* and *VIT*, respectively (**Figure 4D**, **3H**). Additionally, rs12028757 jointly affected t-CpG methylation and *WLS* gene expression (**Figure 4L-M**).

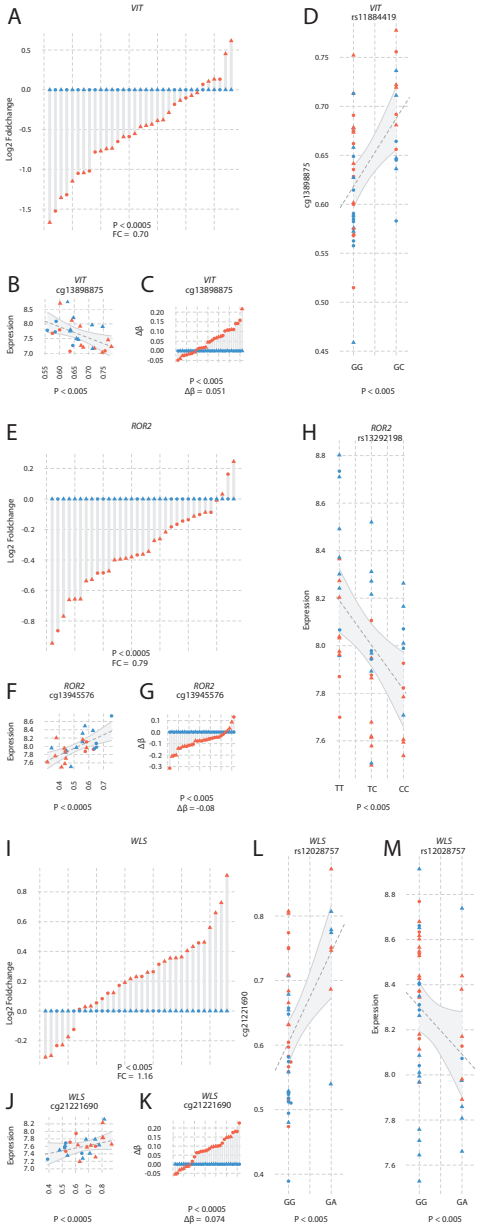


Figure 4. Examples of cumulative evidence for putative causal involvement of *VIT*, *ROR2* and *WLS* in OA pathophysiology. Preserved and lesioned samples are respectively coloured blue and red, while knee and hip samples are respectively depicted as circles and triangles. (**A, E, I**) Significant differential methylation was observed between preserved and lesioned cartilage for all three genes. (**B, F, J**) A significant direct relation between expression and respective CpG methylation was observed. (**C, G, K**) As expected, a significant difference in methylation was observed between preserved and lesioned tissue as well. (**D, H, L, M**) In conjunction with DNA methylation, expression was regulated by proximal SNPs.

DISCUSSION

The here presented study encompasses the first comprehensive multi-level integration of genome wide data to gain a more accurate understanding of OA associated changes in the epigenetically regulated transcriptome of articular chondrocytes. By stepwise integrating transcriptomic and epigenetic data in relation to cartilage OA severity, we identified 70 unique genes with OA responsive t-CpGs likely affecting expression in articular cartilage. Subsequent pathway analyses showed significant enrichment for genes that act within skeletal development. Moreover, we have shown that for 31 and 26 OA cartilage relevant genes, respectively, methylation and expression is additionally affected by genetic variation proximal to these genes.

Although the observed enrichment of OA responsive t-CpGs among genes within developmental pathways either marks disease advancement or an adaptation of the preserved cartilage to the adjacent lesioned tissue, our data shows that changes in epigenetically regulated control of developmental genes and OA progression are markedly linked. This could indicate either specific, dynamic regulation of expression of the genes in these pathways by the challenged articular chondrocytes in an attempt to cope with end-stage OA, or alternatively, chondrocytes at end stage disease have lost their ability to epigenetically control expression of essential genes involved in skeletal development and consequently recuperate growth plate morphology and start cartilage debilitating expression, a well described hallmark of OA. The latter hypothesis is supported by the here observed difference of epigenetic control of skeletal development associated genes with OA, such as *VIT, ROR2* and *WLS*. Markedly, in this respect, are also the results of comprehensive genome wide searches for genetic variants conferring risk for OA that have resulted in robust genome wide significant signals at genes implicated in these developmental pathways (24).

In the current study, we present genetic and epigenetic loci that are functionally relevant for OA responsive t-CpGs and cartilage expressed genes and should ideally now be followed up as candidate genes in large genome wide association (GWA) datasets to investigate whether these variants indeed confer a relatively large number of small effects that are responsible for the missing heritability observed in OA. We have observed a relatively small number of differentially expressed genes that are regulated by t-CpGs and/or SNPs. Although this could be due to statistical power and/or small effect sizes, it unquestionably highlights the importance of combining epigenome data, or gene specific epigenetic data for that matter, with other types of molecular data, to gain a robust understanding of and to biologically interpret the observed differences. Furthermore of note, we did neither observe established genetic OA susceptibility nor OA related epigenetic loci, which implies that our transcriptional, tissue relevant approach offers additional, compelling knowledge about

genes involved in mature articular cartilage homeostasis and late OA when compared to traditional GWA approaches. Even more so, the epigenetic and transcriptional effects of OA susceptibility genes such as *GDF5* (15) and *DIO2* (14) are relatively subtle, whereas we have aimed to select genes with larger effects in late OA. Also, we have possibly missed out on long distance or trans acting t-CpGs or SNPs, as a result of our applied 10kb cut-off. While we do not disregard the possible impact of long distance, transcriptional relevant loci, the measure of effect will likely be inversely correlated with the genomic distance. In order to study these effects accurately, larger sample sizes than our own would be required. Considering the earlier observed gene specific transcriptionally relevant and OA associated differentially methylated CpGs, we were unfortunately unable to address these as the applied methylation array lacks the density and subsequently does not measure these.

The *VIT, ROR2* and *WLS* genes are notable examples for which we here present functional epigenetic, genetic and transcriptional (**Figure 3**) differences depending on the late pathophysiological state of articular cartilage (25;26). *VIT* is a relatively under studied gene in both cartilage biology and osteoarthritis. Nonetheless, proteomic analysis of mouse hip cartilage revealed involvement in cartilage development (27). More specifically, vitrin, the protein product of the *VIT* gene, contains a Von Willebrand factor A domain, and is subsequently involved in ECM integrin signalling (28). Expression of *ROR2* drives chondrocyte expansion (29) and is known to be involved in regulating the TNFRSF11B/TNFSF11 protein (commonly referred to as OPG/RANK) ratio in articular chondrocytes, (30) a well described disrupted pathway in OA. (31) Down regulation of *ROR2* inhibits the chondrocytes regenerative capacities, while disruption of the OPG/RANKL ratio has been shown to induce calcification and bone formation. (29;31;32) Another major player in joint development and cartilage biology is the Wnt pathway (33-36), in which *ROR2* (37;38) and *WLS* (39;40) as well as a number of OA susceptibility genes are situated (34;41;42). While the role of Wnt signalling is evident in cartilage development and OA, *WLS* is specifically involved in the endochondral ossification process (39).

While OA related differences in methylation in articular cartilage have also been reported by others (9;11;12) , our results imply that changes in epigenetic control only lead to expression differences at a limited number of genes. More specifically, at genes involved in either maintaining the chondrocyte phenotype or adversely pursuing the endochondral ossification lineage. Moreover, the detected local SNPs that affected either methylation or gene expression in articular cartilage of epigenetically controlled genes may inherently affect proper cartilage homeostasis and potentially affect OA susceptibility. In this regard, we observed SNPs that influenced DNA methylation at t-CpGs, while no direct relation between the respective genotypes and gene expression was observed. Likely, a large number of factors obscure the direct regulatory mechanism between the local genetic background, t-CpG methylation and gene expression. Of note, these mechanisms likely arbitrate differential expression

among the genes in which we did observe differential t-CpG methylation but no difference in expression. SNPs and t-CpGs that appear to solely affect t-CpG methylation or expression respectively, are still of relevance, however, the transcriptional effects of these variants should be addressed in larger consortia. Further mechanistic studies, such as longitudinal measurements in animal experiments or actively perturbing the relevant genes in cell systems, are required to accurately address the hypothesis.

In conclusion, we have here shown that OA related epigenetic differences need to be integrated with other sources of molecular data, such as genomic and transcriptomic, to enhance our understanding of the pathophysiological processes of OA. Furthermore, by integration of multiple layers of genome wide data we have identified genes, such as *VIT*, *ROR2* and *WLS*, which are likely modulating OA pathophysiology and possibly reflect the loss of the chondrocyte's maturational arrested state. Although targeting DNA methylation seems unlikely to stand at the basis for developing treatments, it serves to deepen our understanding of the complex transcriptomic changes in OA affected articular cartilage.

ACKNOWLEDGMENTS

We thank all participants of the RAAK study. The RAAK studies were supported by the Leiden University Medical Centre, the Dutch Arthritis Association (DAA 101-402 and Reumafonds LRR) and the Centre of Medical System Biology and Netherlands Consortium for Healthy Aging both in the framework of the Netherlands Genomics Initiative (NGI). Furthermore, we acknowledge support by TreatOA and IDEAL, which are funded by the European Union's Seventh Framework Program (FP7/2007-2011) under respective grant agreement nos. 200800 and 259679. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Medical ethics approval LUMC P08.239

REFERENCES

1 Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. The Lancet 2011; 377 (9783):2115-26.
2 Loeser RF, Goldring SR, Scanzello CR, Goldring MB. Osteoarthritis: A disease of the joint as an organ. Arthritis & Rheumatism 2012; 64 (6):1697-707.
3 Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 2013; 8 (2):203-9.
4 Dvir-Ginzberg M, Reich E. Chopping off the chondrocyte proteome. Biomarkers 2014;1-7.
5 Sun MM, Beier F. Chondrocyte hypertrophy in skeletal development, growth, and disease. Birth Defects Res C Embryo Today 2014; 102 (1):74-82.
6 Hosaka Y, Saito T, Sugita S, Hikata T, Kobayashi H, Fukai A et al. Notch signaling in chondrocytes modulates endochondral ossification and osteoarthritis development. Proc Natl Acad Sci U S A 2013; 110 (5):1875-80.
7 Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT et al. Dynamic changes in the human methylome during differentiation. Genome Res 2010; 20 (3):320-31.
8 Slieker RC, Bos SD, Goeman JJ, Bovee JV, Talens RP, van der Breggen R et al. Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. Epigenetics Chromatin 2013; 6 (1):26.
9 Rushton MD, Reynard LN, Barter MJ, Refaie R, Rankin KS, Young DA et al. Characterization of the cartilage DNA methylome in knee and hip osteoarthritis. Arthritis Rheumatol 2014.
10 den Hollander W, Ramos YF, Bos SD, Bomer N, van der Breggen R, Lakenberg N et al. Knee and hip articular cartilage have distinct epigenomic landscapes: implications for future cartilage regeneration approaches. Ann Rheum Dis 2014.
11 Jeffries MA, Donica M, Baker L, Stevenson M, Annan AC, Humphrey MB et al. Genome-wide DNA methylation study identifies significant epigenomic changes in osteoarthritic cartilage. Arthritis Rheumatol 2014.
12 Fernandez-Tajes J, Soto-Hermida A, Vazquez-Mosquera ME, Cortes-Pereira E, Mosquera A, Fernandez-Moreno M et al. Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster of osteoarthritic patients. Ann Rheum Dis 2014; 73 (4):668-77.
13 Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol 2011; 12 (1):R10.
14 Bomer N, den Hollander W, Ramos YFM, Bos SD, van der Breggen R, Lakenberg N et al. Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis. Annals of the Rheumatic Diseases 2014.
15 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; 133 (8):1059-73.
16 Hashimoto K, Oreffo RO, Gibson MB, Goldring MB, Roach HI. DNA demethylation at specific CpG sites in the IL1B promoter in response to inflammatory cytokines in human articular chondrocytes. Arthritis Rheum 2009; 60 (11):3303-13.
17 Hashimoto K, Otero M, Imagawa K, de Andres MC, Coico JM, Roach HI et al. Regulated transcription of human matrix metalloproteinase 13 (MMP13) and interleukin-1beta (IL1B) genes in chondrocytes depends on methylation of specific proximal promoter CpG sites. J Biol Chem 2013; 288 (14):10061-72.
18 Ramos YF, den Hollander W, Bovee JV, Bomer N, van der Breggen R, Lakenberg N et al. Genes Involved in the Osteoarthritis Process Identified through Genome Wide Expression Analysis in Articular Cartilage; the RAAK Study. PLoS ONE 2014; 9 (7):e103056.
19 Schoenmaker M, de Craen AJ, de Meijer PH, Beekman M, Blauw GJ, Slagboom PE et al. Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. Eur J Hum Genet 2006; 14 (1):79-84.
20 Ramos YF, Metrustry S, Arden N, Bay-Jensen AC, Beekman M, de Craen AJ et al. Meta-analysis identifies loci affecting levels of the potential osteoarthritis biomarkers sCOMP and uCTX-II with genome wide significance. J Med Genet 2014; 51 (9):596-604.
21 Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 2014; 46 (11):1173-86.

22 Douglas Bates, Martin Maechler, Ben Bolker. lme4: Linear mixed-effects models using S4 classes. 2012.
23 Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. The Lancet 2001; 380 (9844):815-23.
24 Bos SDl, Slagboom PE, Meulenbelt I. New insights into osteoarthritis: early developmental features of an ageing-related disease. Current Opinion in Rheumatology 2008; 20 (5).
25 Pullig O, Weseloh G, Swoboda B. Expression of type VI collagen in normal and osteoarthritic human cartilage. Osteoarthritis Cartilage 1999; 7 (2):191-202.
26 Soder S, Hambach L, Lissner R, Kirchner T, Aigner T. Ultrastructural localization of type VI collagen in normal adult and osteoarthritic human articular cartilage. Osteoarthritis Cartilage 2002; 10 (6):464-70.
27 Wilson R, Norris EL, Brachvogel B, Angelucci C, Zivkovic S, Gordon L et al. Changes in the chondrocyte and extracellular matrix proteome during post-natal mouse cartilage development. Mol Cell Proteomics 2012; 11 (1):M111.
28 Manabe R, Tsutsui K, Yamada T, Kimura M, Nakano I, Shimono C et al. Transcriptome-based systematic identification of extracellular matrix proteins. Proc Natl Acad Sci U S A 2008; 105 (35):12849-54.
29 DeChiara TM, Kimble RB, Poueymirou WT, Rojas J, Masiakowski P, Valenzuela DM et al. Ror2, encoding a receptor-like tyrosine kinase, is required for cartilage and growth plate development. Nat Genet 2000; 24 (3):271-4.
30 Ermakov S, Trofimov S, Malkin I, Livshits G. A significant association exists between receptor tyrosine kinase-like orphan receptor 2 gene variants and the OPG/RANKL ratio in human plasma. Osteoporos Int 2012; 23 (7):1899-907.
31 Ramos YFM, Bos SD, van der Breggen R, Kloppenburg M, Ye K, Lameijer EW et al. A gain of function mutation in TNFRSF11B encoding osteoprotegerin causes osteoarthritis with chondrocalcinosis. Annals of the Rheumatic Diseases 2014.
32 Liu Y, Bhat RA, Seestaller-Wehr LM, Fukayama S, Mangine A, Moran RA et al. The orphan receptor tyrosine kinase Ror2 promotes osteoblast differentiation and enhances ex vivo bone formation. Mol Endocrinol 2007; 21 (2):376-87.
33 Kruger C, Kappen C. Expression of cartilage developmental genes in Hoxc8- and Hoxd4-transgenic mice. PLoS ONE 2010; 5 (2):e8978.
34 Casta+¦o Betancourt MC, Cailotto F, Kerkhof HJ, Cornelis FMF, Doherty SA, Hart DJ et al. Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. Proceedings of the National Academy of Sciences 2012; 109 (21):8218-23.
35 Leijten JC, Emons J, Sticht C, van GS, Decker E, Uitterlinden A et al. Gremlin 1, frizzled-related protein, and Dkk-1 are key regulators of human articular cartilage homeostasis. Arthritis Rheum 2012; 64 (10):3302-12.
36 Lories RJ, Corr M, Lane NE. To Wnt or not to Wnt: the bone and joint health dilemma. Nat Rev Rheumatol 2013; 9 (6):328-39.
37 Ford CE, Qian Ma SS, Quadir A, Ward RL. The dual role of the novel Wnt receptor tyrosine kinase, ROR2, in human carcinogenesis. Int J Cancer 2013; 133 (4):779-87.
38 Fukuyo S, Yamaoka K, Sonomoto K, Oshita K, Okada Y, Saito K et al. IL-6-accelerated calcification by induction of ROR2 in human adipose tissue-derived mesenchymal stem cells is STAT3 dependent. Rheumatology (Oxford) 2014; 53 (7):1282-90.
39 Lu C, Wan Y, Cao J, Zhu X, Yu J, Zhou R et al. Wnt-mediated reciprocal regulation between cartilage and bone development during endochondral ossification. Bone 2013; 53 (2):566-74.
40 Wang LT, Wang SJ, Hsu SH. Functional characterization of mammalian Wntless homolog in mammalian system. Kaohsiung J Med Sci 2012; 28 (7):355-61.
41 Evangelou E, Chapman K, Meulenbelt I, Karassa FB, Loughlin J, Carr A et al. Large-scale analysis of association between GDF5 and FRZB variants and osteoarthritis of the hip, knee, and hand. Arthritis & Rheumatism 2009; 60 (6):1710-21.
42 Loughlin J, Dowling B, Chapman K, Marcelline L, Mustafa Z, Southam L et al. Functional variants within the secreted frizzled-related protein 3 gene are associated with hip osteoarthritis in females. Proceedings of the National Academy of Sciences of the United States of America 2004; 101 (26):9757-62.

Supplemental figures and tables can be found at doi: 10.1002/art.39162.

6

KNEE AND HIP ARTICULAR CARTILAGE HAVE DISTINCT EPIGENOMIC LANDSCAPES: IMPLICATIONS FOR FUTURE CARTILAGE REGENERATION APPROACHES

W. den Hollander[1,*], Y.F.M. Ramos[1], S.D. Bos[1,2], N. Bomer[1,3], R. van der Breggen[1], N. Lakenberg[1], W.J. de Dijcker[1], Bouke J. Duijnisveld[4], P.E. Slagboom[1,2,3], Rob G.H.H. Nelissen[4], I. Meulenbelt[1,2].

[1]Dept. of Molecular Epidemiology, LUMC, Leiden, The Netherlands. [2]Genomics Initiative, sponsored by the NCHA, Leiden, The Netherlands. [3]IDEAL The Netherlands. [4]Dept. of Orthopaedics, LUMC, Leiden, The Netherlands.

http://dx.doi.org/10.1136/annrheumdis-2014-205980

ABSTRACT

Objectives To elucidate the functional epigenomic landscape of articular cartilage in osteoarthritis (OA) affected knee and hip joints in relation to gene expression.

Methods Using Illumina Infinium HumanMethylation450 BeadChip arrays genome wide DNA methylation was measured in 31 preserved and lesioned cartilage sample pairs (14 knees and 17 hips) from patients who underwent a total joint replacement due to primary OA. Using previously published genome wide expression data of 33 pairs of cartilage samples, of which 13 pairs were overlapping with the current methylation dataset, we assessed gene expression differences in Differentially Methylated Regions (DMRs).

Results Principal Component Analysis of the methylation data revealed distinct clustering of knee and hip samples, irrespective of OA pathophysiology. A total of 6272 CpG dinucleotides were differentially methylated between the two joints, comprising a total of 357 DMRs containing 1817 CpGs and 245 unique genes. Enrichment analysis of genes proximal of the DMRs revealed significant enrichment for developmental pathways and homeobox (HOX) genes. Subsequent transcriptomic analysis of DMR genes exposed distinct knee and hip expression patterns.

Conclusions Our findings reveal consistent DMRs between knee and hip articular cartilage that marked transcriptomic differences among HOX genes, which were not reflecting the temporal sequential HOX expression pattern during development. This implies distinct mechanisms for maintaining cartilage integrity in adulthood, thereby contributing to our understanding of cartilage homeostasis and future tissue regeneration approaches.

INTRODUCTION

Articular cartilage (AC) is highly specialized and characteristic tissue in all synovial joints at the ends of longitudinal bones. Its main function is to facilitate protection of subchondral bone against heavy loads, while maintaining smooth locomotor function of the articular joint. (1) AC of the load bearing knee and hip joints are morphologically similar, both in health and disease. (2-4) Furthermore, histological assessment and expression profiling of preserved and osteoarthritic AC have revealed generic processes and pathways to be involved in osteoarthritis (OA) pathophysiology, independent of the affected joint. (5-7) Nevertheless, epidemiological studies and genome wide approaches have respectively shown distinct prevalence patterns and genetic risk factors for OA at different joints. (8, 9) Moreover, although pathway analysis of gene expression data has revealed transcriptomic commonalities between knee and hip AC in OA, individual gene expression differences have been reported. (7)

In general, tissue identity is marked by the epigenetic landscape of respective cells and is, among others, reflected in the DNA methylation profile. (10) DNA methylation, in which the cytosine residue in cytosine-phosphate-guanine dinucleotides (CpGs) acquires a methyl group, is known to regulate gene expression upon environmental changes such as age and disease. Overall differences in the methylome on the tissue level are commonly reflected in differentially methylated regions (DMRs), while single CpGs do usually not harbor this property (10, 11) and possibly only mark environmental, stochastic or individual differences. These observations raise the question whether knee and hip AC are either epigenetically distinct or similar tissues. Although some differences on the epigenetic level between knee and hip joints have been reported, (12) it is currently unknown whether the observed differences in DNA methylation have any functional properties in terms of regulating expression of putative joint specific genes.

In the current study, pairwise preserved and lesioned AC from knee and hip joints was sampled from patients undergoing joint replacement surgery due to primary OA. Genome wide DNA methylation was measured to assess communalities and discrepancies of the AC methylome in knee and hip joints and with respect to preserved and lesioned AC. Furthermore, we have subsequently combined epigenomic and transcriptomic data to gain a functional understanding of the observed methylation differences. To our knowledge this is the first study in which highly similar tissues, being AC from either knee or hip joints, are compared comprehensively on the epigenomic and transcriptomic level.

MATERIALS AND METHODS

THE RAAK COHORT AND SAMPLING.

Ethical approval was obtained from the medical ethics committee of the LUMC (P08.239) and informed consent was obtained from all participants. (6) Participant details are listed in **Supplementary Table S1**. For sampling details see online supplemental methods and (13). Macroscopically preserved as well as macroscopically lesioned cartilage was sampled from patients who underwent a total joint replacement due to primary OA of either the knee (N=14) or hip (N=17). From an additional 3 knee and 3 hip joints healthy cartilage was sampled.

DNA ISOLATION.

DNA was isolated using the Promega Wizard Genomic DNA Purification kit according to the manufacturer's protocol.

METHYLATION ARRAYS.

DNA was bisulphite treated using the ZymoResearch EZ DNA Methylation kit. DNA methylation was assessed using Illumina Infinium HumanMethylation450 BeadChips. All methylation values are reported as fractions between 0 and 1, commonly known as the β value. For additional details see online supplemental methods.

EXPRESSION DATA.

Normalized expression data was downloaded from GEO (GSE57218) (13).

STATISTICAL ANALYSES.

Principal Component Analysis (PCA) and statistical procedures were carried out in R-3.0.2. All analyses were corrected for technical covariates as well as sex, disease status, age and BMI. A random effect for patient ID was included to correct for putative correlations between preserved and OA affected AC from the same joint. Additional details are listed in the supplemental methods.

RESULTS

KNEE AND HIP AC SHOW DISTINCT METHYLATION PROFILES.

Genome wide DNA methylation profiling was performed in all samples, consisting of both macroscopically preserved and lesioned AC derived from 14 knee and 17 hip joints. By means of PCA we observed two distinct clusters of samples, reflecting the joint type from which the cartilage was

sampled (**Figure 1A**). Although in knee samples preserved AC tended to cluster apart from lesioned AC, clustering by joint type was largely independent of the OA affection status (**Figure 1C-F**). This was further emphasized when 6 truly healthy samples (3 knees, 3 hips) were included in the PCA, which showed that irrespective of age and disease samples clustered according their joint type (**Figure 1B**). Next, to elucidate the specific CpGs driving this distinct clustering by joint type, we fitted a linear mixed model to identify the specific CpGs that were differentially methylated between hip and knee cartilage, while correcting for sex, age, BMI and OA affection status. After adjustment for multiple testing (Benjamini-Hochberg) 6272 CpGs were significantly differentially methylated between knee and hip AC (P<0.05) by at least 0.1β, covering a total of 2726 unique genes (**Supplementary Table S2**).

Differentially methylated regions among homeobox containing genes. Next, we applied a sliding window algorithm (10) to distinguish inherent tissue differences in the methylation data from possible environmental, stochastic or individual differences and thereby observed 357 DMRs, consisting of 1817 CpGs and 245 unique genes (**Supplementary Table S3**). Pathway analysis revealed significant enrichment among the constructed DMRs mostly for developmental pathways (such as limb development and skeletal system morphogenesis) and, more specifically, homeodomain containing genes (**Supplementary Table S4**). Rather strikingly in this respect is the presence of 42 DMRs in all four canonical homeobox (HOX) clusters, comprising over 10% of the observed DMRs. Here again we observed no major distinction between OA affected joints and healthy joints, confirming that the DMRs are highly joint specific (**Supplementary Figure S1**). Visual inspection of representative CpGs in DMRs ratified the distinct and consistent differences in methylation between the two joint types, while this was less evident between preserved and OA affected AC (**Supplementary Figure 2**).

Putative functionality of DMRs in adult articular cartilage. Finally, to investigate the putative functionality of the observed DMRs, in terms of respective mRNA expression of proximal genes, we assessed the expression patterns of the 245 DMR associates genes (**Supplementary Table S3**). Previously, gene expression was quantified for 33 pairs of preserved and lesioned knee (N=11) and hip (N=22) AC, of which 13 pairs were overlapping with the methylation data (GSE57218) (13). Hereby, differential expression analysis of the entire GSE57218 dataset revealed that independent of the OA affection status, 28 out of 245 genes were differentially expressed between knee and hip AC (**Table 1**). Respectively, 6 and 11 genes were only expressed in either knee or hip AC, while 11 genes were expressed significantly different between knee and hip AC. Among the 28 differentially expressed genes, genes from all four HOX clusters were present, as well as multiple HOX containing co-factors (such as *PITX1*, *MEIS2*, *DLX5* and *IRX3*).

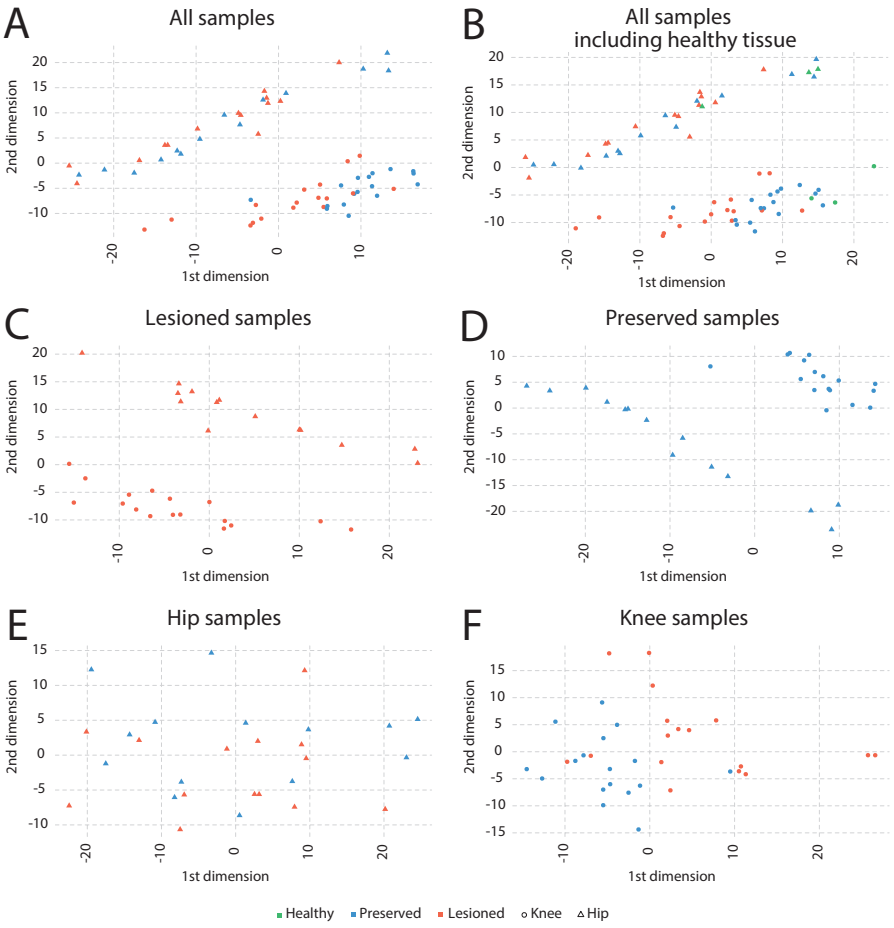


Figure 1. PCA of DNA methylation data. Red dots indicate lesioned samples, blue dots indicate preserved samples. Knee and hip samples are depicted as circles and triangles respectively. (**A**) PCA of all OA affected cartilage samples (N=62) revealed clear clustering based on joint type. (**B**) PCA of all OA affected samples and an additional 6 healthy (3 knees, 3 hips) samples. Again clustering is evidently according to joint type, whereas the disease state of the joint does not influence clustering markedly. (**C**) PCA of only lesioned samples. (**D**) PCA of only preserved samples. (**E**) PCA of only OA affected hip samples, both preserved and lesioned. (**F**). PCA of only OA affected knee samples, both preserved and lesioned.

Table 1. Significant differentially expressed DMR associated genes between knee (N=11) and hip (N=22) AC in the entire GSE57218 dataset (13). Empty fields indicate that expression was not detected (ND) in either knee or hip samples, therefore no fold change or P-value could be calculated. Mean expression values are reported as normalized log transformed intensity signals, while Fold difference indicates the difference in relative expression.

Joint	Gene	Mean Expression		Fold	Adjusted
		Knee	Hip	difference	P-value
Knee	*SIX1*	7.55	ND		
	HOXA3	7.50	ND		
	WWOX	7.49	ND		
	HOXD13	7.46	ND		
	C10orf41	7.46	ND		
	HOXA11	7.41	ND		
Both	*LOC375295*	10.3	9.14	2.25	1.51^{-14}
	PITX1	10.6	9.57	2.02	2.66^{-7}
	DLX5	10.9	10.2	1.66	1.26^{-4}
	GMDS	11.7	11.3	1.36	1.21^{-2}
	SCNN1A	8.71	8.28	1.35	7.95^{-3}
	HOXB2	7.87	8.40	0.69	9.64^{-4}
	HOXC8	7.82	8.40	0.67	5.34^{-13}
	MEIS2	7.74	8.34	0.66	4.93^{-8}
	SGK	10.3	11.2	0.53	4.23^{-3}
	IRX3	7.92	8.86	0.52	9.64^{-4}
	HOXC6	8.62	9.65	0.49	1.79^{-18}
Hip	*HOXA4*	ND	7.40		
	KLHL26	ND	7.41		
	HOXC9	ND	7.41		
	FGF9	ND	7.45		
	HOXA10	ND	7.46		
	SAMD11	ND	7.53		
	HAND2	ND	7.54		
	HAND2-AS1	ND	7.56		
	IRX5	ND	7.58		
	GSC	ND	7.62		
	HOXC4	ND	7.69		

DISCUSSION

In the current study, we report on differences in the epigenetic landscapes between knee and hip AC. Based on the entire DNA methylation landscape, knee and hip AC show distinct epigenomic profiles independent of the tissue disease state (**Figure 1**). Subsequent in-depth analysis of the CpG dinucleotides conferring these distinct profiles (**Figure 2**, **Supplementary Table S2**) revealed significant enrichment for developmental genes such as the canonical homeotic clusters and HOX co-factors (**Supplementary Table S4**). Furthermore, integration of epigenomic and transcriptomic data revealed significant differences in expression among these enriched loci between knee and hip AC, mediated by tissue specific DMRs (**Table 1**).

In order to identify generic and subsequently functional joint related changes, as opposed to possible stochastic, environmental or individual related differences, we have constructed DMRs, since DMRs are known to consistently reflect the tissue of origin (10, 11). We here report on the fact that despite the morphological and functional similarities between knee and hip AC (1-4), they contain inherently distinct cellular phenotypes based on their functional epigenomic landscape. Nonetheless, our data shows that although methylation profiles at DMRs are highly tissue specific, they do not necessarily correlate to gene expression, as only a minority of DMRs appear to be associated with joint specific gene expression differences (28 out of 245 genes). This emphasizes the need for comprehensive integration of multiple levels of genome wide data, such as transcriptomics, for the interpretation of epigenomic studies in OA.

Although specific HOX gene functions remain partly elusive, in part due to complex interactions with HOX co-factors, increasingly more developmentally distinct functions are being ascribed to the various HOX genes. (14, 15) In adult tissues, however, regulation of HOX gene expression and their respective function remains largely unknown and is likely tissue specific. (16, 17) Multiple studies have reported on distinct expression patterns of homeotic genes and related HOX co-factors in adult tissues, reflecting the collinear embryonic HOX code. (16-18) However, up to date not much is known about AC in this respect. If the observed differences in this study are due to retainment of the embryonic HOX code, i.e. the spatiotemporal expression pattern during development, then across the four canonical HOX clusters similar differences in methylation between knee and hip AC were to be expected. Here however, functional differences in DNA methylation across the HOX clusters were observed, while the embryonic HOX code or colinearity were absent, as is reflected by unique knee and hip methylation patterns observed across the four HOX clusters (**Supplementary Table S1**). This observation suggests specific functional roles for the basal HOX transcription factors and likely marks differences in cellular identity between chondrocytes residing in either knee or hip AC.

In the field of tissue engineering major efforts are made to understand cartilage homeostasis, thereby contributing to the development of novel therapeutic approaches for treatment of degenerative joint diseases, including OA. (19) The results presented in the current study, suggesting differences in cellular identity between chondrocytes residing in either knee or hip AC independent of OA pathophysiology, could putatively have implications for future regenerative approaches. As HOX genes are crucially involved in AC development, (20, 21) the observed epigenomic and transcriptomic differences in this study could indicate that directing articular chondrocytes into extracellular matrix production and/or active remodeling of damaged AC could hypothetically be achieved only via distinct mechanisms, depending on the joint type of a cartilage lesion. Moreover, it has been shown that expression of certain sets of HOX genes regulate the regenerative propensity of neural crest cells, (22) presumptively indicating that chondrocytes originating from either knee or hip AC exhibit unequal regenerative capacities.

To our knowledge, we are the first to report on inherent differences between knee and hip AC by virtue of the joint specific epigenetically regulated transcriptomic landscape of HOX gene clusters and related co-factors. Recent studies on genome wide DNA methylation in AC have focused on either the knee (12, 23) or hip (12, 24) joints and more primarily on comparing OA affected to control tissues. Nevertheless, Rushton *et al.* (12) did report epigenetic differences between knee and hip AC at specific CpGs. However, they have only analyzed single CpGs as opposed to DMRs, did not reported on the specific HOX loci nor for that matter did they study the functionality of the reported differences in terms of transcriptomic regulation. We have here consequently shown that especially the latter is crucial to comprehend the results from epigenetic studies.

Although we here have jointly analyzed preserved and lesioned cartilage originating from the same joint, as shown in **Figure 2** there exists only little epigenetic variation within sample pairs, while the joint specific differences are markedly present. This is even further pronounced when non OA samples, derived from healthy joints, cluster tightly according their joint specific epigenome (**Figure 1** and **Supplementary Figure S1**). Furthermore, putative confounding due to pooling is corrected for in all statistical analyses by including a random effect for patient ID. Of note, we do neither rule out nor disregard the presence of relevant gene specific epigenetic differences between preserved and lesioned AC. In the current study, however, we have restricted our focus on inherent epigenetic tissue differences between knee and hip joints.

In conclusion, we have observed consistent DMRs between knee and hip joints among HOX domain containing genes, both in the four canonical homeotic clusters as well as HOX co-factors. They were found to mark differential expression of genes residing in or near these DMRs in AC of knee and hip

joints. The different methylation profiles of knee and hip AC likely mark distinct cellular identities, which could have relevant implications for the field of AC tissue engineering. Together these findings contribute to our understanding of cartilage homeostasis and future repair strategies.

ACKNOWLEDGMENTS

We thank all participants of the RAAK study. The RAAK studies were supported by the Leiden University Medical Centre, the Dutch Arthritis Association (DAA 101-402 and Reumafonds LRR) and the Centre of Medical System Biology and Netherlands Consortium for Healthy Aging both in the framework of the Netherlands Genomics Initiative (NGI). Furthermore, we acknowledge support by TreatOA and IDEAL, which are funded by the European Union's Seventh Framework Program (FP7/2007-2011) under respective grant agreement nos. 200800 and 259679. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing Interests None

Patient consent Obtained

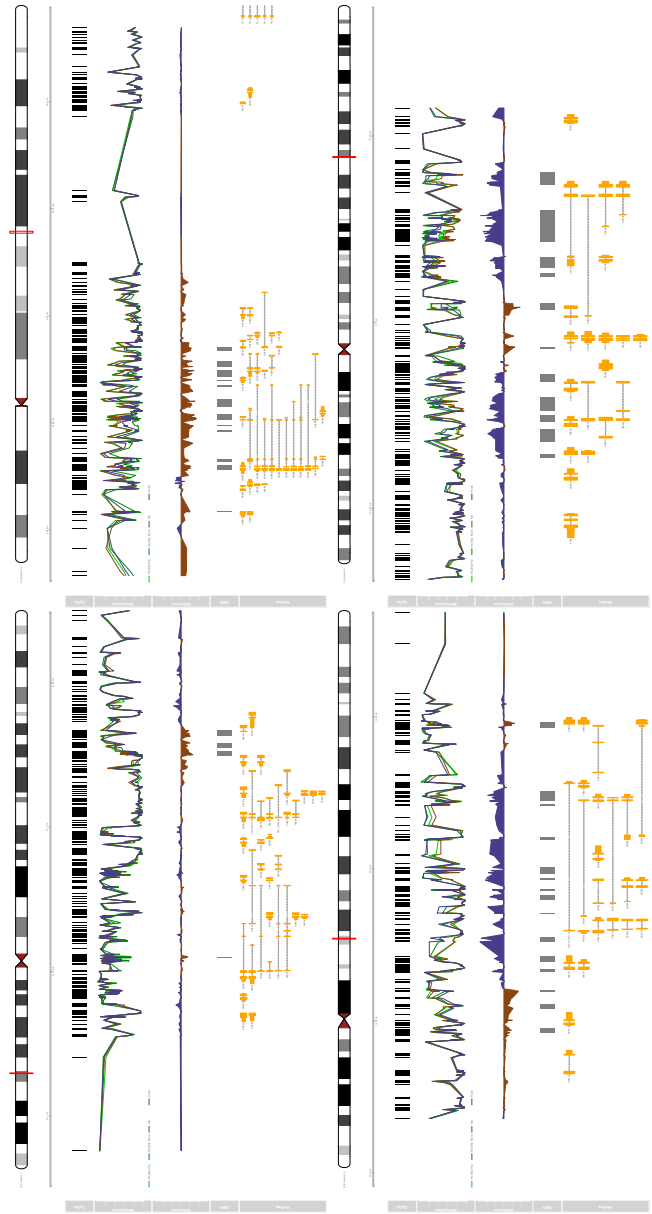
Ethics approval LUMC (P08.239)

REFERENCES

1 Hosseini SM, Wu Y, Ito K, van Donkelaar CC. The importance of superficial collagen fibrils for the function of articular cartilage. *Biomech Model Mechanobiol* 2014;13:41-51.
2 Shepherd DE, Seedhom BB. Thickness of human articular cartilage in joints of the lower limb. *Ann Rheum Dis* 1999;58:27-34.
3 Hosseininia S, Lindberg LR, Dahlberg LE. Cartilage collagen damage in hip osteoarthritis similar to that seen in knee osteoarthritis; a case-control study of relationship between collagen, glycosaminoglycan and cartilage swelling. *BMC Musculoskelet Disord* 2013;14:18.
4 Khan HI, Aitken D, Zhai G, Ding C, Pelletier JP, Pelletier JM, *et al.* Association between hip and knee cartilage measured using radiographs and magnetic resonance imaging: the Tasmanian Older Adult Cohort Study. *Rheumatology* (*Oxford*) 2013;52:2009-15.
5 Bos SD, Bov+¬e JVMG, Duijnisveld BJ, Raine EVA, van Dalen WJ, Ramos YFM, *et al.* Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. *Annals of the Rheumatic Diseases* 2012;71:1254-8.
6 Bomer N, den Hollander W, Ramos YFM, Bos SD, van der Breggen R, Lakenberg N, *et al.* Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis. *Annals of the Rheumatic Diseases* 2014.
7 Xu Y, Barter MJ, Swan DC, Rankin KS, Rowan AD, Santibanez-Koref M, *et al.* Identification of the pathogenic pathways in osteoarthritic hip cartilage: commonality and discord between hip and knee OA. *Osteoarthritis Cartilage* 2012;20:1029-38.
8 Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. *Bull World Health Organ* 2003;81:646-56.
9 Reynard LN, Loughlin J. The genetics and functional analysis of primary osteoarthritis susceptibility. *Expert Reviews in Molecular Medicine* 2013;15:null.
10 Slieker RC, Bos SD, Goeman JJ, Bovee JV, Talens RP, van der Breggen R, *et al.* Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. *Epigenetics Chromatin* 2013;6:26.
11 Rakyan VK, Down TA, Thorne NP, Flicek P, Kulesha E, Graf S, *et al.* An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res* 2008;18:1518-29.
12 Rushton MD, Reynard LN, Barter MJ, Refaie R, Rankin KS, Young DA, *et al.* Characterization of the cartilage DNA methylome in knee and hip osteoarthritis. *Arthritis Rheumatol* 2014.
13 Ramos YF, den HW, Bovee JV, Bomer N, van der Breggen R, Lakenberg N, *et al.* Genes Involved in the Osteoarthritis Process Identified through Genome Wide Expression Analysis in Articular Cartilage; the RAAK Study. *PLoS ONE* 2014;9:e103056.
14 Ladam F, Sagerstrom CG. Hox regulation of transcription: more complex(es). *Dev Dyn* 2014;243:4-15.
15 Cerda-Esteban N, Spagnoli FM. Glimpse into Hox and tale regulation of cell differentiation and reprogramming. *Dev Dyn* 2014;243:76-87.
16 Takahashi Y, Hamada J, Murakawa K, Takada M, Tada M, Nogami I, *et al.* Expression profiles of 39 HOX genes in normal human adult organs and anaplastic thyroid cancer cell lines by quantitative real-time RT-PCR system. *Exp Cell Res* 2004;293:144-53.
17 Chang HY, Chi JT, Dudoit S, Bondre C, van de Rijn M, Botstein D, *et al.* Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc Natl Acad Sci U S A* 2002;99:12877-82.
18 Tsumagari K, Baribault C, Terragni J, Chandra S, Renshaw C, Sun Z, *et al.* DNA methylation and differentiation: HOX genes in muscle cells. *Epigenetics Chromatin* 2013;6:25.
19 Kock L, van Donkelaar CC, Ito K. Tissue engineering of functional articular cartilage: the current status. *Cell Tissue Res* 2012;347:613-27.
20 Kruger C, Kappen C. Expression of cartilage developmental genes in Hoxc8- and Hoxd4-transgenic mice. *PLoS ONE* 2010;5:e8978.
21 Yueh YG, Gardner DP, Kappen C. Evidence for regulation of cartilage differentiation by the homeobox gene Hoxc-8. *Proc Natl Acad Sci U S A* 1998;95:9956-61.
22 Le Douarin NM, Creuzet S, Couly G, Dupin E. Neural crest cell plasticity and its limits. *Development* 2004;131:4637-50.
23 Fernandez-Tajes J, Soto-Hermida A, Vazquez-Mosquera ME, Cortes-Pereira E, Mosquera A, Fernan-

dez-Moreno M, *et al.* Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster of osteoarthritic patients. *Ann Rheum Dis* 2014;73:668-77.
24 Jeffries MA, Donica M, Baker L, Stevenson M, Annan AC, Humphrey MB, *et al.* Genome-wide DNA methylation study identifies significant epigenomic changes in osteoarthritic cartilage. *Arthritis Rheumatol* 2014.

Supplementary data can be found at doi: 10.1136/annrheumdis-2014-205980



Supplementary Figure S1. Overview of the methylation status of the four HOX clusters.

7

DNA METHYLATION IN OSTEOARTHRITIS

Wouter den Hollander[1,*] & Ingrid Meulenbelt[1]

[1]*Deptartment of Molecular Epidemiology, LUMC, Leiden, The Netherlands*

https://dx.doi.org/10.2174%2F1389202916666150817212711

ABSTRACT

Osteoarthritis (OA) is a prevalent disease of articular joints and primarily characterized by degradation and calcification of articular cartilage. Presently, no effective treatment other than pain relief exists and patients ultimately need to undergo replacement surgery of the affected joint. During disease progression articular chondrocytes, the single cell type present in articular cartilage, show altered transcriptional profiles and undergo phenotypic changes that resemble the terminal differentiation route apparent in growth plate chondrocytes. Hence, given its prominent function in both regulating gene expression and maintaining cellular phenotypes, DNA methylation of CpG dinucleotides is intensively studied in the context of OA. An increasing number of studies have been published that employed a targeted approach on genes known to play a role in OA pathophysiology. As of such, it has become clear that OA responsive DNA methylation changes seem to mediate disease associated aberrant gene expression. Furthermore, established OA susceptibility alleles such as *GDF5* and *DIO2* appear to confer OA risk via DNA methylation and respective pathophysiological expression changes. In more recent years, genome wide profiling of DNA methylation in OA affected articular cartilage has emerged as a powerful tool to address the epigenetic changes in their entirety, which has resulted in the identification of putative patient subgroups as well as generic OA associated pathways.

Box 1. List of abbreviations.

OA	Osteoarthritis
ECM	Extracellular matrix
CpG	Cytosine-guanine dinucleotide
SNP	Single nucleotide polymorphism
eQTM	Expression quantitative trait methylation
eQTL	Expression quantitative trait locus
mQTL	Methylation quantitative trait locus
AZA	5-aza-2'-deoxycytidine
DMR	Differentially methylated region
HOX	Homeobox

INTRODUCTION

Osteoarthritis (OA) is currently the most prevalent arthritic disease among the elderly population (1). Patients are subject to impaired mobility, joint stiffness, pain and a significant decrease in quality of life. Presently, no cure exists and patients with end-stage disease ultimately need to undergo a joint replacement surgery of the affected joint. Subsequently, as disease incidence is increasing with the ongoing ageing population, the societal burden both in terms of disabling patients and economic concerns will continue to rise (2). Pathophysiologically, the disease is primarily characterized by progressive degradation and calcification of cartilage in the articular joints, although in recent years it has become apparent that other tissues such as subchondral bone (3-5) and synovium (6-8) play substantial roles in OA pathology as well. The articular cartilage contains a hyaline type extracellular matrix (ECM) made up of collagens, proteoglycans and other structural proteins (9). Articular chondrocytes, the single cell type present in articular cartilage, reside here and maintain tissue homeostasis by remodelling the ECM upon stresses and microtraumas (10). To ensure cartilage integrity and cope with the challenges throughout life, the maturational arrested articular chondrocytes need to continuously employ, possibly error prone, adaptions to changes in the environment (11). In order to facilitate these adaptations, articular chondrocytes are required to dynamically adjust expression of the appropriate genes, while maintaining their specific cellular phenotype (11). In this regard, articular chondrocytes present in OA affected cartilage however, show increased expression of catabolic enzymes (12-14) and have seemingly lost their maturational arrested state (15-17), as they proliferate and regain growth-plate morphology (17), while degrading and calcifying the ECM of the articular cartilage (18,19).

A substantial number of mechanisms are known that regulate gene expression (20) and cell fate persistence (21-23), commonly referred to as epigenetics. While there exists a considerable number of epigenetic layers, such as histone modifications, microRNAs and long non-coding RNAs, the most studied in OA is decidedly DNA methylation. Partly due to its seemingly, relatively stable epigenetic mark on gene expression and partly due to the readily available techniques to measure it. DNA methylation is the phenomenon in which the cytosine nucleotide that is located in cytosine-guanine (CpG) residue pairs gets methylated. Whether CpGs get methylated depend on the local genetic sequence (24), the cellular requirements (21) and environmental factors such the putative pathophysiological state of the respective tissue (24,25). The presence of methyl groups on CpGs is believed to interfere with binding of proteins to the DNA and subsequently affects transcription (26), which is marked by the correlation between the fraction of CpG methylation and gene expression (27), commonly referred to as expression quantitative trait methylation (eQTM) (24,27). As of such, tissue specific methylation profiles are assumed to maintain the respective transcriptional character and identity

of distinct cell types. Interestingly, disease associated single nucleotide polymorphisms (SNPs) that influence aberrant expression, so called expression quantitative trait loci (eQTL), frequently seem to modulate their transcriptomic properties via DNA methylation quantitative trait loci (mQTL) (28,29).

DNA METHYLATION IN OSTEOARTHRITIS

The involvement of DNA methylation in OA pathophysiology is becoming increasingly evident, reflected by the growing body of literature on the subject (13,28-40). However, given the cross-sectional nature of the reports that studied DNA methylation in OA affected articular cartilage, it is currently unclear whether the observed epigenetic differences precede disease onset or are merely consequence of the environmental changes that articular chondrocytes are subject to in OA. As of such, we here discuss two hypotheses that have been proposed to attribute distinct roles of epigenetics in OA pathophysiology.

Firstly, it is hypothesized that individuals with unfavourable epigenetic profiles will be more prone to develop OA and/or progress faster (41,42). As the local genetic sequence significantly influences the DNA methylation state of CpGs, these putative profiles likely arise due to the presence of OA susceptibility SNPs (43). This would be in concordance with the identification of susceptibility genes in complex diseases such as OA, which not uncommonly assert their susceptibility via altered DNA methylation, depending on the presence of associated risk alleles (28,29). For example, OA risk alleles of SNPs in *DIO2* and *GDF5* appear to generate OA predisposing epigenetic profiles and modulate disease associated, epigenetically associated gene expression. The OA risk allele C of the rs225014 T>C SNP, which has been identified in a combined genome wide linkage and association study by our own group (44), marked not only allele specific expression of *DIO2* (45), it additionally mediated epigenetic regulation of the gene (29). More specifically, carriers of the risk allele showed increased transcriptional activity upon hypermethylation at a CpG located in a distal regulatory element. The risk allele T of the rs143383 C>T SNP (46-48), located in the 5' untranslated region of *GDF5*, disrupts a CpG dinucleotide, of which has been shown that it modulates absolute *GDF5* expression in articular cartilage (28). Furthermore, the reported allele specific expression of *GDF5* in heterozygous carriers marked by the rs143383 alleles (49) is affected by the extent of methylation of the respective CpG. Although increasing additional evidence is reported about OA susceptibility alleles, the exact biological mechanisms that confer OA susceptibility is unclear. More specifically, it is unclear whether epigenetic regulation of aberrant gene expression brought about by genetic variation is involved in disease aetiology or disease progression. We can however, conclude that epigenetic regulation of gene expression modulates OA susceptibility, at least at the *DIO2* and *GDF5* loci.

Alternatively, it has been suggested that age-related loss of epigenetic control (50) mediates the loss of the articular chondrocyte's phenotype with ongoing OA, as life-long stresses and adaptations are expected to leave their mark on the epigenome (11). Furthermore, given the altered cellular phenotype chondrocytes acquire in OA and given the role DNA methylation fulfils in maintaining cellular phenotypes, we advocate here that loss of epigenetic control causes reactivation of developmental pathways among articular chondrocytes that are subject to OA and subsequently mediate the morphological changes that are associated with affected articular chondrocytes.

EPIGENETIC CHANGES ACCOMPANY OSTEOARTHRITIS

A growing body of literature reports on altered DNA methylation at specific genes involved in the OA disease process, commonly at precise CpG sites located in regulatory elements near the respective genes (**Table 1**). Frequently, a cross-sectional study design is utilized to study OA associated methylation differences, either between healthy and affected or pairwise between preserved and macroscopically lesioned articular cartilage. Among others, catabolic and developmentally associated genes like *MMP13*, *GDF5*, *SOX9, DIO2* and *ADAMTS4* were shown to be differentially expressed between control and affected tissue, presumably mediated by differences in DNA methylation (13,28-34). Although it is not quite clear to what extent DNA methylation changes in OA development contribute to disease onset or progression, a growing number of studies investigates the mechanism by which such changes may affect chondrocyte function.

Table 1. Overview of gene targeted DNA methylation studies in osteoarthritic articular cartilage.

Gene	Methylation in OA	Expression in OA	Sample size (OA, Control)	Joint	AZA	CpG vector	Reference
COL9	Increased	Down	12, 10	Hip	yes	Yes	(31)
GDF5	Decreased	Up	24, 19	Knee and hip	yes	Yes	(28)
DIO2	Increased	Up	52, 52	Knee and hip	yes	No	(29)
IL1β	Decreased	Up	18, 12	Hip	yes	Yes	(30)
MMP13	Decreased	Up	17, 12	Hip	yes	Yes	(30)
iNOS	Decreased	Up	13, 15	Hip	no	Yes	(34)
SOX9	Increased	Down	9, 9	Hip	yes	No	(32)
ADAMTS4	Decreased	Up	4, 1	Hip	no	No	(33)
ADAMTS4	Decreased	Up	16, 10	Hip	no	No	(13)
MMP13	Decreased	Up	16, 10	Hip	no	No	(13)
MMP3	Decreased	Up	16, 10	Hip	no	No	(13)
MMP9	Decreased	Up	16, 10	Hip	no	No	(13)

Consequently, various experimental set ups have been applied to confirm the mechanistic relation between DNA methylation and gene expression in the chondrocyte (**Table 1**). A commonly applied experiment that aims to validate observed eQTM loci, as has our own group, is the addition of 5-aza-2'-deoxycytidine (AZA) to the medium of cultured chondrocytes, either using immortalized cell lines or primary chondrocyte cultures (28-30,32,35). AZA is a chemical agent that interferes with the addition of methyl groups on a genome wide scale to newly formed DNA strands during replication. Although valuable information can be gained from such experiments in a global sense, negative outcomes should not per se be considered as experimental falsifications of the earlier observed correlation between DNA methylation and expression, which is likely cell type and locus specific. As of such, not being able to validate earlier observed eQTM loci can very well arise from the fact that the entire genome gets demethylated upon prolonged AZA treatment, of which the transcriptomic consequences potentially overshadow the locus specific relation. Additionally, culturing cells *in-vitro* forces the cells to adapt to an artificial environment, which possibly only resembles the original tissues in a broad sense and subsequently disrupts the regulatory properties of cell type specific eQTM loci.

Another type of validation experiment being applied is cloning the genomic sequence, in which differential methylation was observed, into a CpG-free vector (51) containing a luciferase gene downstream of the multiple cloning site (30,31,34,52). Next, the vector is methylated *in-vitro* only at the cloned region and transfected into chondrocyte like cell-lines. Luciferase activity now depends on the transcriptional activity of the cloned region, which in turn depends on the methylated state of

that respective region. Although using a CpG-free vector has great advantages over AZA treatment, as it does not induce genome-wide altered methylation, the technique will potentially only work for proximal promotors, as long-distance three dimensional genomic structures, which are often seen in gene expression regulation (53-55), depend on the complex interplay of the distant and proximal regulatory elements of genes. Again, as cell culturing is required to apply the technique, measuring promoter activity using a CpG free vector might be influenced by cellular adaptations to the artificial environment.

STRATIFICATION OF OSTEOARTHRITIS PATIENTS BY GENOME WIDE DNA METHYLATION PROFILING

In more recent years, multiple studies have reported on genome wide DNA methylation profiles of articular cartilage in the context of OA (**Table 2**), not least due to development of affordable genome wide DNA methylation arrays, such as the Illumina Infinium HumanMethylation450k BeadChip array. Genome wide profiling of DNA methylation in OA affected cartilage has revealed the presence of multiple OA patient strata, which are characterized by their respective methylation profiles. Firstly, Rushton *et al*. (37) and our group (38) have reported on the distinct genome wide DNA methylation profiles of knee and hip articular cartilage. Secondly, Rushton *et al*. (37) and Fernandez-Tajes *et al.* (36) have reported on a subgroup of patients defined by altered DNA methylation at inflammatory related genes.

Table 2. Overview of current genome wide DNA methylation experiments in osteoarthritic articular cartilage.

Sample size (OA, Control)	Joint	Platform	OA associated pathways	Reference
20, 25	Knee	Illumina 27K	Inflammation, transcriptional regulation, ECM homeostasis	(36)
21, 96	Knee and hip	Illumina 450k	ECM homeostasis, ossification, inflammation, angiogenesis	(37)
24, 24	Hip	Illumina 450k	Development, angiogenesis, inflammation	(39)
16, 16	Knee	Agilent 244k	Development, catalytic activity	(40)
31, 31	Hip and knee	Illumina 450k		(38)
31, 31	Hip and knee	Illumina 450k	Development, ECM homeostasis	(56)

With respect to the first bifurcation of OA patients, the two types of joint specific articular cartilage are distinguished by large differentially methylated regions (DMRs), primarily at genes involved in development and cellular differentiation. Notably, over 10% of DMRs were observed in the four canonical homeobox (HOX) clusters while the spatiotemporal pattern present at the HOX clusters during development was absent, indicating that joint specific DMRs likely bear distinct functionality in adult articular cartilage (38). Both studies that compared knee and hip articular cartilage reported

on joint specific methylation patterns at the four HOX clusters and several homeotic cofactors, such as *IRX3*, *SIX1*, *MEIS2* and *DLX5*. By analysis of transcriptomic data our group subsequently revealed that joint specific DMRs potentially mediate distinctive regenerative capacities of articular chondrocytes residing in different joints, given the developmental and regenerative nature of joint specific expression of genes located in the developmentally associated DMRs (38). Although reports on joint specific methylation profiles might not reveal loci that are directly relevant for OA onset or progression, they undeniably expose the heterogeneity of the disease. These joint specific DNA methylation and associated transcriptomic profiles are concurrent with the reports of joint specific genetic OA susceptibility loci (57). Henceforth, we can conclude that although a substantial amount of common features are present between knee and hip OA, the disease process and subsequent mode of action of putative therapeutic interventions might be different between the two joints.

With respect to the second, another dichotomy is observed by cluster analysis of genome wide DNA methylation profiles of osteoarthritic articular cartilage (36,37). More specifically, putative patient subgroups have been defined by DNA methylation profiles of the affected cartilage that are enriched for differentially methylated CpGs located in or near genes involved in inflammation. The study of Fernandez-Tajes *et al*. used a sparse methylation array and a relatively small sample size consisting of only knee OA patients, which possibly explains why Rushton *et al*. observed substantially more differentially methylated CpGs (1,357 vs. 5,769) that separated the inflammatory cluster of patients in both knee and hip articular cartilage. While enrichment analysis of both studies revealed broadly similar enriched GO terms, being the inflammatory response and cytokine production, in the study of Fernandez-Tajes *et al.* this was only apparent among the hypomethylated CpGs and not among hypermethylated CpGs. Among the consistent inflammatory signals identified by the two studies among knee OA patients were multiple interleukin genes such as *IL10*, *IL16* and *IL19*, but also developmental and ECM related genes such as *RUNX2*, *FGFR1*, *COL6A3* and *COL18A1*. Noteworthy, in the study of Rushton *et al*. stratification by the inflammatory profile, using both hypo- and hypermethylated CpGs, of both knee and hip OA patients is observed, presumably mediated by overlapping pathways, albeit that only 3,496 out of 15,239 (23%) differentially methylated CpGs that separated OA hip patients overlapped with differentially methylated CpGs that separated OA knee patients (37). Very recently, the group of Rushton *et al*. has further investigated the hypomethylated genes that are apparent among the inflammatory profiles in their hip OA patient cohort and report on specific zinc reporter genes that potentially mediate the patient stratification. (58)

Intriguingly, when we performed GO term analysis on the genes that separated hip and knee OA patients consistently in the initial study of Rushton *et al*. and Fernandez-Tajes *et al*., we observed significant enrichment for ECM maintenance pathways (data not shown). Thus, while one of the

reported clusters of OA patients is presumably characterized by an epigenetic inflammatory profile, a common underlying mechanisms appears defined by epigenetic regulation of ECM related genes, such as *COL6A3*, *RUNX2*, *MMP13* and *ADAMTS5*. This proposition is additionally reflected by the fact that all studies report on the enrichment of ECM related pathways (**Table 2**) and by the analysis Fernandez-Tajes *et al*. performed on solely the hypermethylated CpGs, in which they also observe enrichment among ECM maintenance pathways. Finally, the study of Fernandez-Tajes *et al.* grossly compared methylation and additionally expression profiles between the OA subgroups. However, it is evident that the transcriptional consequences of these profiles need to be precisely elucidated, as modulating unfavourable epigenetic but subsequent transcriptomic profiles specifically can potentially attenuate disease onset or progression and might therefor serve as valuable therapeutic targets for the putative subtypes of OA. It should, however, be noted that of the five published genome wide DNA methylation studies in OA, only Fernandez-Tajes *et al.* and Rushton *et al.* have observed separate clustering of OA patients, while Jeffries *et al*. (39), Moazedi-Fuerst *et al*. (40) and our group (56) have not, warranting further research hereinto. Furthermore, not only discovery driven epigenomic profiling of articular cartilage is required to understand possible OA patient heterogeneity. In light of proposed OA subtypes in the literature, epigenetic interrogation of specific pathways by means of burden analyses in pathways such as those in estrogenic sensitivity (59) or apoptosis (60), might be a more powerful approach in the context of the large number of differentially methylated loci in OA reported by increasingly larger studies. In parallel, genome wide DNA methylation profiles of other joint tissues might additionally confer the proposed different OA subtypes.

CONSISTENT DNA METHYLATION DIFFERENCES BETWEEN OSTEOARTHRITIS AFFECTED AND CONTROL ARTICULAR CARTILAGE

Despite the putative segregation of distinct methylation profiles among OA affected joints (37,38) and OA patients (36,37), major communalities have been observed across the different genome wide studies (**Table 2**). Specifically, consistent enrichment of differential methylation among genes involved in development, as well as in collagen synthesis and other ECM maintenance pathways is reported by all studies, particularly among genes from the RUNX, COL and MMP families. Developmental processes thus appear entangled with OA associated degradation and calcification of articular cartilage during disease progression. Although our group has shown that the majority of differentially methylated CpGs do not associate *in-cis* with gene expression, enrichment analysis of OA responsive CpGs that did correlate with gene expression (87 CpGs, 70 genes) also revealed enrichment for ECM maintenance and developmental pathways, consisting of genes such as *ROR2*, *WLS*, *VIT* and *SPP1* (56). Nonetheless, genome wide DNA methylation profiles and the inherent differences between OA and control tissue need thus be interpreted with cause and preferably coupled with additional molecular measurements, such as gene expression data, to clarify the biological

consequence. The 76% of genes of which expression did not correlate with *in-cis* DNA methylation could be regulated additionally by other epigenetic mechanisms, such as histon modifications or miRNA mediated silencing or alternatively, might possibly reflect remnants of early developmental or past wound healing processes.

In parallel, as reflected by OA risk alleles that affect DNA methylation mediated gene expression (28,29) and the influential role of the genomic sequence on DNA methylation, partly reflected by the 40% of OA associated CpGs that are affected by the alleles of proximal SNPs (56), it is not unlikely that the total combination of minor genetic variants among OA patient contributes to putative unfavourable epigenomic and subsequent transcriptomic profiles. In this respect, the putative OA inflammatory subtype and the tissue specific mQTLs reported by our own group, could be the consequence of an inherent unfavourable epigenomic profile preceding disease onset. In line with the observed lack of heritability explained by traditional GWA approaches, these putative minor genetic variants might make up for the larger part of heritability in OA.

CONCLUSIONS

Genome wide DNA methylation profiling of OA affected articular cartilage has revealed widespread differences between OA and control tissues. Although the majority of CpGs do not associate with gene expression, CpGs of which we can consequently not conclude whether they play an active role in disease progression, reactivation of developmental pathways due to changes in epigenetic landscape is apparent in OA affected joints. Thus, we hypothesize here that to maintain healthy articular cartilage homeostasis throughout life, the respective chondrocytes utilize epigenetic mechanisms to transcriptomically adapt to the changing environment. Furthermore, the fact that gene expression changes occur via both hyper- and hypomethylation in OA affected articular cartilage, indicates that disease associated differential methylation is unlikely to be the product of a solely passive process. Lifelong adjustment of regulatory mechanisms, such as DNA methylation, is likely subject to stochastic error and subsequent accumulation of epigenetic modifications at developmental pathways, either via inaccurate restoration of the chondrocytes' steady state or via an increasing number of epigenetic adaptations, seem to force chondrocytes towards terminal differentiation (**Figure 1**). This proposition is additionally reflected by the morphological changes that occur in articular cartilage with ongoing OA, as reflected by degradation and calcification, mechanisms also observed among growth plate chondrocytes during development.

In conclusion, studying genome wide DNA methylation in OA has proven to serve as an excellent proxy to probe the underlying pathological cellular processes in OA. In light of the inherent genetically and consequent epigenetically complex nature of OA pathology, the genome wide efforts of recent years that have reported on the DNA methylation landscape of OA affected articular cartilage have delivered valuable insight, insight that would not have become apparent by mere gene targeted approaches.

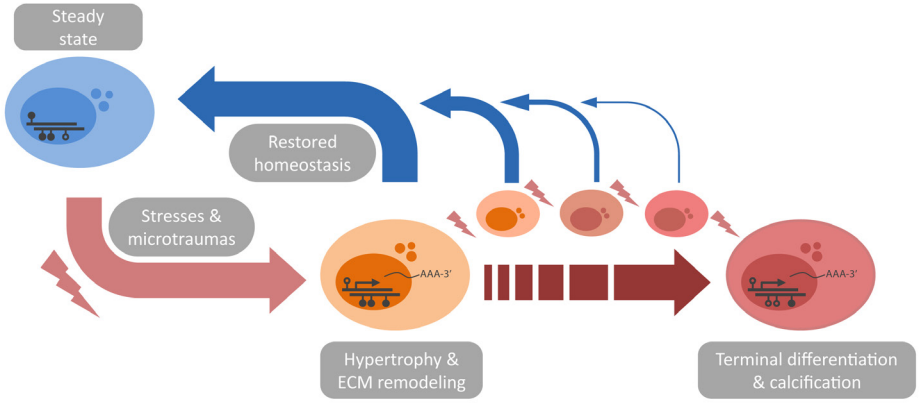


Figure 1. Schematic overview of the proposed loss of epigenetic control among articular chondrocytes. Articular cartilage is subject to lifelong challenges which requires the respective chondrocytes to dynamically adapt gene expression in order to return to and maintain homeostasis and subsequent tissue integrity. This is seemingly accomplished by using epigenetic modifications, reflected by OA associated DNA methylation differences, which are likely not reversed flawlessly upon returning to a steady state. We propose that accumulation of these epigenetic adaptations will eventually lead to altered cellular phenotypes which are unable to return to a healthy, steady state chondrocyte. In parallel, articular chondrocytes of OA patients might already bear an unfavorable epigenomic profile preceding disease onset, which implies that these chondrocytes are more prone to enter the active ECM remodeling state and/or might be less able to return to a steady state. Finally, independent of OA associated epigenetic changes in articular chondrocytes, joint and patient specific epigenomic profiles seem to modulate our proposed mechanism.

FUTURE PROSPECTS OF EPIGENETICS IN OSTEOARTHRITIS

Given that differences in the methylome of OA affected compared to healthy or preserved articular cartilage do not imply downstream effects, reflected by the relatively small number of *in-cis* eQTM loci reported by our group, stresses the need for integration of DNA methylation data with other types of molecular profiling. It is of great interest to the field that the disease relevant, tissue specific epigenomic and transcriptomic QTL profiles are characterised, as they might reveal the predisposing, complex genetic architecture that underlies OA susceptibility. Our own group has undertaken the first step herein and has reported on the effects of SNPs on OA associated DNA methylation and

transcriptional differences, as well as on the direct relation between methylation and expression on a genome wide scale (56). Moreover, the reported genetic variants will potentiate the power of GWA studies, as the multiple testing penalty is substantially reduced when only the functional SNPs, in terms of regulation of transcription, are addressed as opposed to the entire genome. Likely a larger number of SNPs are of relevance, as we have not addressed long distance effects and larger sample sizes are possibly required to gain additional, robust understanding of the reported differences and associations, a point also raised by others (61). In order to test whether indeed loss of epigenetic control confers the pathophysiological changes are apparent with OA progression, longitudinal studies that address the contribution of ageing in altered DNA methylation are required, preferably specifically at genes of which transcriptional changes potentially modulate the disease process. Confirmation of this hypothesis might potentially reveal the genetic drivers of OA. _

In light of the systemic, but not unidirectional differences in DNA methylation on a genome wide scale that are apparent with OA progression in affected cartilage, it seems that systemically targeting DNA methylation for clinical purposes in OA is farfetched. Moreover, the widespread epigenetic differences that accompany OA suggestively affect expression of genes primarily involved in developmental processes, such as endochondral ossification. Aside from the substantial scientific challenge to locus specifically modulate DNA methylation, it seems more pragmatic to directly address the mRNA or protein molecules of the respective genes and pathways. While this is a difficult task on its own, specifically and locus specifically targeting DNA methylation in our view will unlikely serve a clinical purpose for OA in the near future.

However, aside from the putative limited role of epigenetics in curing OA, it may serve as an important biomarker when measured in clinically available tissues. As has also been shown for a number of molecular markers in blood or serum (62-64), it is apparent that transcriptomic data can purposely be used to identify symptomatic OA patients using mRNA extracted from peripheral blood mononuclear cells (60). For example, the blood transcriptomic profiles of OA patients are enriched for genes involved in apoptosis, which were subsequently shown to reflect the pathophysiological state of the articular cartilage (60). Hence, in light of the responsiveness of the epigenome to environmental changes and its relationship with gene expression it is expected epigenetics can fulfil a similar purpose. Moreover, hereby not only clinical associations can be constructed, but also important insight is given into the complex disease process. Preliminary results from our own group indicate that indeed DNA methylation might serve as a powerful biomarker for OA progression. More specifically, as little as four CpGs were needed to distinguish fast progressing from non-progressing OA patients with 76% accuracy (65).

It should be noted that we have here primarily discussed DNA methylation in articular cartilage in the context of OA, not in the last place due to the fact that degradation of articular cartilage is the primary feature of OA. However, it is clear that other joint tissues are also involved in the disease process as a whole. Other disease relevant tissues, such as the synovium (66-68) and to some extent the subchondral bone (69), have been addressed in light of other muscoskeletal or rheumatic pathologies and indicate compelling prospects for OA research. In this regard, preliminary data from a small study revealed DNA methylation differences between subchondral bone adjacent to varying degrees of damaged articular cartilage of knee OA patients, indicating that epigenetic regulation is likely involved in the pathophysiological interplay between subchondral bone and articular cartilage (69).

Considering the proposed subgroups of OA patients, the initial genome wide DNA methylation studies discussed here have reported on putative stratification of OA patients based on inflammatory profiles present in articular cartilage only. However, given the increasingly important role of secondary tissues in OA, it is possible that subgroups of patients can be defined by genome wide DNA methylation profiling of those respective tissues. Consequently, the OA research field might elucidate other proposed OA patient subgroups, such as distinct differential estrogenic responses advocated by Herrero-Beaumont *et al*. (59).

REFERENCES

1. Bijlsma J.W., Berenbaum F., and Lafeber F.P. (2011). Osteoarthritis: an update with relevance for clinical practice**.** The Lancet *377***:** 2115-2126.
2. Woolf A.D., Erwin J., and March L. (2012). The need to address the burden of musculoskeletal conditions**.** Best. Pract. Res. Clin. Rheumatol. *26***:** 183-224.
3. Thysen S., Luyten F.P., and Lories R.J. (2015). Loss of Frzb and Sfrp1 differentially affects joint homeostasis in instability-induced osteoarthritis**.** Osteoarthritis. Cartilage. *23***:** 275-279.
4. Cui Z., Xu C., Li X., Song J., and Yu B. (2015). Treatment with recombinant lubricin attenuates osteoarthritis by positive feedback loop between articular cartilage and subchondral bone in ovariectomized rats**.** Bone *74C***:** 37-47.
5. Binks D.A., Gravallese E.M., Bergin D., Hodgson R.J., Tan A.L., Matzelle M.M., McGonagle D., and Radjenovic A. (2015). Role of vascular channels as a novel mechanism for subchondral bone damage at cruciate ligament entheses in osteoarthritis and inflammatory arthritis**.** Ann. Rheum. Dis. *74***:** 196-203.
6. van Lent P.L., Blom A.B., Schelbergen R.F., Sloetjes A., Lafeber F.P., Lems W.F., Cats H., Vogl T., Roth J., and van den Berg W.B. (2012). Active involvement of alarmins S100A8 and S100A9 in the regulation of synovial activation and joint destruction during mouse and human osteoarthritis**.** Arthritis Rheum. *64***:** 1466-1476.
7. Lories R.J., Corr M., and Lane N.E. (2013). To Wnt or not to Wnt: the bone and joint health dilemma**.** Nat. Rev. Rheumatol. *9***:** 328-339.
8. Ford C.E., Qian Ma S.S., Quadir A., and Ward R.L. (2013). The dual role of the novel Wnt receptor tyrosine kinase, ROR2, in human carcinogenesis**.** Int. J. Cancer *133***:** 779-787.
9. Hosseini S.M., Wu Y., Ito K., and van Donkelaar C.C. (2014). The importance of superficial collagen fibrils for the function of articular cartilage**.** Biomech. Model. Mechanobiol. *13***:** 41-51.
10. Lotz M. and Loeser R.F. (2012). Effects of aging on articular cartilage homeostasis**.** Bone *51***:** 241-248.
11. Goldring M.B. and Marcu K.B. (2012). Epigenomic and microRNA-mediated regulation in cartilage development, homeostasis, and osteoarthritis**.** Trends Mol. Med. *18***:** 109-118.
12. Struglics A., Larsson S., Pratta M.A., Kumar S., Lark M.W., and Lohmander L.S. (2006). Human osteoarthritis synovial fluid and joint cartilage contain both aggrecanase- and matrix metalloproteinase-generated aggrecan fragments**.** Osteoarthritis. Cartilage. *14***:** 101-113.
13. Roach H.I., Yamada N., Cheung K.S., Tilley S., Clarke N.M., Oreffo R.O., Kokubun S., and Bronner F. (2005). Association between the abnormal expression of matrix-degrading enzymes by human osteoarthritic chondrocytes and demethylation of specific CpG sites in the promoter regions**.** Arthritis Rheum. *52***:** 3110-3124.
14. Hashimoto K., Otero M., Imagawa K., de Andres M.C., Coico J.M., Roach H.I., Oreffo R.O., Marcu K.B., and Goldring M.B. (2013). Regulated transcription of human matrix metalloproteinase 13 (MMP13) and interleukin-1beta (IL1B) genes in chondrocytes depends on methylation of specific proximal promoter CpG sites**.** J. Biol. Chem. *288***:** 10061-10072.
15. Aigner T. and Dudhia J. (1997). Phenotypic modulation of chondrocytes as a potential therapeutic target in osteoarthritis: a hypothesis**.** Ann. Rheum. Dis. *56***:** 287-291.
16. Facchini A., Borzi R.M., Olivotto E., Platano D., Pagani S., Cetrullo S., and Flamigni F. (2012). Role of polyamines in hypertrophy and terminal differentiation of osteoarthritic chondrocytes**.** Amino. Acids *42***:** 667-678.
17. Sun M.M. and Beier F. (2014). Chondrocyte hypertrophy in skeletal development, growth, and disease**.** Birth Defects Res. C. Embryo. Today *102***:** 74-82.
18. Bantsimba-Malanda C., Cottet J., Netter P., Dumas D., Mainard D., Magdalou J., and Vincourt J.B. (2013). Chondrocalcin is internalized by chondrocytes and triggers cartilage destruction via an interleukin-1beta-dependent pathway**.** Matrix Biol. *32***:** 443-451.
19. van C.A., Blaney D.E., Garcia d., V, van G.E., van den Berg W., Goumans M.J., Ten D.P., and van der Kraan P. (2015). The high affinity ALK1-ligand BMP9 induces a hypertrophy-like state in chondrocytes that is antagonized by TGFbeta1**.** Osteoarthritis. Cartilage.
20. Jaenisch R. and Bird A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals**.** Nat. Genet. *33 Suppl***:** 245-254.
21. Slieker R.C., Bos S.D., Goeman J.J., Bovee J.V., Talens R.P., van der Breggen R., Suchiman H.E., Lameijer E.W., Putter H., van den Akker E.B. et al. (2013). Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array**.** Epigenetics. Chromatin. *6***:** 26.

22. Laurent L., Wong E., Li G., Huynh T., Tsirigos A., Ong C.T., Low H.M., Kin Sung K.W., Rigoutsos I., Loring J. et al. (2010). Dynamic changes in the human methylome during differentiation. Genome Res. *20*: 320-331.
23. Lister R., Pelizzola M., Dowen R.H., Hawkins R.D., Hon G., Tonti-Filippini J., Nery J.R., Lee L., Ye Z., Ngo Q.M. et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*: 315-322.
24. Gutierrez-Arcelus M., Lappalainen T., Montgomery S.B., Buil A., Ongen H., Yurovsky A., Bryois J., Giger T., Romano L., Planchon A. et al. (2013). Passive and active DNA methylation and the interplay with genetic variation in gene regulation. Elife. *2*: e00523.
25. Ramos Y.F., den H.W., Bovee J.V., Bomer N., van der Breggen R., Lakenberg N., Keurentjes J.C., Goeman J.J., Slagboom P.E., Nelissen R.G. et al. (2014). Genes Involved in the Osteoarthritis Process Identified through Genome Wide Expression Analysis in Articular Cartilage; the RAAK Study. PLoS. ONE. *9*: e103056.
26. Wang H., Maurano M.T., Qu H., Varley K.E., Gertz J., Pauli F., Lee K., Canfield T., Weaver M., Sandstrom R. et al. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome Res. *22*: 1680-1688.
27. Wagner J.R., Busche S., Ge B., Kwan T., Pastinen T., and Blanchette M. (2014). The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. Genome Biol. *15*: R37.
28. Reynard L.N., Bui C., Canty-Laird E.G., Young D.A., and Loughlin J. (2011). Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. Hum. Mol. Genet. *20*: 3450-3460.
29. Bomer N., den Hollander W., Ramos Y.F.M., Bos S.D., van der Breggen R., Lakenberg N., Pepers B.A., van Eeden A.E., Darvishan A., Tobi E.W. et al. (2014). Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis. Annals of the Rheumatic Diseases.
30. Hashimoto K., Otero M., Imagawa K., de Andres M.C., Coico J.M., Roach H.I., Oreffo R.O., Marcu K.B., and Goldring M.B. (2013). Regulated transcription of human matrix metalloproteinase 13 (MMP13) and interleukin-1beta (IL1B) genes in chondrocytes depends on methylation of specific proximal promoter CpG sites. J. Biol. Chem. *288*: 10061-10072.
31. Imagawa K., de Andres M.C., Hashimoto K., Itoi E., Otero M., Roach H.I., Goldring M.B., and Oreffo R.O. (2014). Association of reduced type IX collagen gene expression in human osteoarthritic chondrocytes with epigenetic silencing by DNA hypermethylation. Arthritis Rheumatol. *66*: 3040-3051.
32. Kim K.I., Park Y.S., and Im G.I. (2013). Changes in the epigenetic status of the SOX-9 promoter in human osteoarthritic cartilage. J. Bone Miner. Res. *28*: 1050-1060.
33. Cheung K.S., Hashimoto K., Yamada N., and Roach H.I. (2009). Expression of ADAMTS-4 by chondrocytes in the surface zone of human osteoarthritic cartilage is regulated by epigenetic DNA de-methylation. Rheumatol. Int. *29*: 525-534.
34. de Andres M.C., Imagawa K., Hashimoto K., Gonzalez A., Roach H.I., Goldring M.B., and Oreffo R.O. (2013). Loss of methylation in CpG sites in the NF-kappaB enhancer elements of inducible nitric oxide synthase is responsible for gene induction in human articular chondrocytes. Arthritis Rheum. *65*: 732-742.
35. Loeser R.F., Im H.J., Richardson B., Lu Q., and Chubinskaya S. (2009). Methylation of the OP-1 promoter: potential role in the age-related decline in OP-1 expression in cartilage. Osteoarthritis. Cartilage. *17*: 513-517.
36. Fernandez-Tajes J., Soto-Hermida A., Vazquez-Mosquera M.E., Cortes-Pereira E., Mosquera A., Fernandez-Moreno M., Oreiro N., Fernandez-Lopez C., Fernandez J.L., Rego-Perez I. et al. (2014). Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster of osteoarthritic patients. Ann. Rheum. Dis. *73*: 668-677.
37. Rushton M.D., Reynard L.N., Barter M.J., Refaie R., Rankin K.S., Young D.A., and Loughlin J. (2014). Characterization of the cartilage DNA methylome in knee and hip osteoarthritis. Arthritis Rheumatol.
38. den Hollander W., Ramos Y.F., Bos S.D., Bomer N., van der Breggen R., Lakenberg N., de Dijcker W.J., Duijnisveld B.J., Slagboom P.E., Nelissen R.G. et al. (2014). Knee and hip articular cartilage have distinct epigenomic landscapes: implications for future cartilage regeneration approaches. Ann. Rheum. Dis. *73*: 2208-2212.
39. Jeffries M.A., Donica M., Baker L., Stevenson M., Annan A.C., Humphrey M.B., James J.A., and Sawalha A.H. (2014). Genome-wide DNA methylation study identifies significant epigenomic changes in osteoarthritic cartilage. Arthritis Rheumatol.
40. Moazedi-Fuerst F.C., Hofner M., Gruber G., Weinhaeusel A., Stradner M.H., Angerer H., Peischler D., Lohberger B., Glehr M., Leithner A. et al. (2014). Epigenetic differences in human cartilage between mild and

severe OA. J. Orthop. Res. *32*: 1636-1645.
41. Reynard L.N. and Loughlin J. (2013). The genetics and functional analysis of primary osteoarthritis susceptibility. Expert Reviews in Molecular Medicine *15*: null.
42. Reynard L.N. and Loughlin J. (2012). Genetics and epigenetics of osteoarthritis. Maturitas *71*: 200-204.
43. Rakyan V.K., Down T.A., Balding D.J., and Beck S. (2011). Epigenome-wide association studies for common human diseases. Nat. Rev. Genet. *12*: 529-541.
44. Meulenbelt I., Min J.L., Bos S., Riyazi N., Houwing-Duistermaat J.J., van der Wijk H.-J., Kroon H.M., Nakajima M., Ikegawa S., Uitterlinden A.G. et al. (2008). Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. Human Molecular Genetics *17*: 1867-1875.
45. Bos S.D., Bov+¬e J.V.M.G., Duijnisveld B.J., Raine E.V.A., van Dalen W.J., Ramos Y.F.M., van der Breggen R., Nelissen R.G.H.H., Slagboom P.E., Loughlin J. et al. (2012). Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. Annals of the Rheumatic Diseases *71*: 1254-1258.
46. Southam L., Rodriguez-Lopez J., Wilkins J.M., Pombo-Suarez M., Snelling S., Gomez-Reino J.J., Chapman K., Gonzalez A., and Loughlin J. (2007). An SNP in the 5'-UTR of GDF5 is associated with osteoarthritis susceptibility in Europeans and with in vivo differences in allelic expression in articular cartilage. Hum. Mol. Genet. *16*: 2226-2232.
47. Chapman K., Takahashi A., Meulenbelt I., Watson C., Rodriguez-Lopez J., Egli R., Tsezou A., Malizos K.N., Kloppenburg M., Shi D. et al. (2008). A meta-analysis of European and Asian cohorts reveals a global role of a functional SNP in the 5' UTR of GDF5 with osteoarthritis susceptibility. Hum. Mol. Genet. *17*: 1497-1504.
48. Valdes A.M., Evangelou E., Kerkhof H.J., Tamm A., Doherty S.A., Kisand K., Tamm A., Kerna I., Uitterlinden A., Hofman A. et al. (2011). The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. Ann. Rheum. Dis. *70*: 873-875.
49. Egli R.J., Southam L., Wilkins J.M., Lorenzen I., Pombo-Suarez M., Gonzalez A., Carr A., Chapman K., and Loughlin J. (2009). Functional analysis of the osteoarthritis susceptibility-associated GDF5 regulatory polymorphism. Arthritis Rheum. *60*: 2055-2064.
50. Talens R.P., Christensen K., Putter H., Willemsen G., Christiansen L., Kremer D., Suchiman H.E., Slagboom P.E., Boomsma D.I., and Heijmans B.T. (2012). Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. Aging Cell *11*: 694-703.
51. Klug M. and Rehli M. (2006). Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector. Epigenetics. *1*: 127-130.
52. Reynard L.N., Bui C., Syddall C.M., and Loughlin J. (2014). CpG methylation regulates allelic expression of GDF5 by modulating binding of SP1 and SP3 repressor proteins to the osteoarthritis susceptibility SNP rs143383. Hum. Genet. *133*: 1059-1073.
53. Zuin J., Dixon J.R., van der Reijden M.I., Ye Z., Kolovos P., Brouwer R.W., van de Corp, van de Werken H.J., Knoch T.A., van IJcken W.F. et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc. Natl. Acad. Sci. U. S. A *111*: 996-1001.
54. Smyk M., Szafranski P., Startek M., Gambin A., and Stankiewicz P. (2013). Chromosome conformation capture-on-chip analysis of long-range cis-interactions of the SOX9 promoter. Chromosome. Res. *21*: 781-788.
55. Sanyal A., Lajoie B.R., Jain G., and Dekker J. (2012). The long-range interaction landscape of gene promoters. Nature *489*: 109-113.
56. den Hollander W., Ramos Y.F.M., Bomer N., Elzinga S., van der Breggen R., Lakenberg N., de Dijcker W.J., Suchiman E.D., Duijnisveld B.J., Houwing-Duistermaat J.J. et al. (2015). Transcriptional associations of osteoarthritis mediated loss of epigenetic control in articular cartilage. Arthritis & Rheumatology n/a.
57. arcOGEN Consortium and arcOGEN Collaborators (2012). Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. Lancet *380*: 815-823.
58. Rushton M.D., Young D.A., Loughlin J., and Reynard L.N. (2015). Differential DNA methylation and expression of inflammatory and zinc transporter genes defines subgroups of osteoarthritic hip patients. Annals of the Rheumatic Diseases.
59. Herrero-Beaumont G., Roman-Blas J.A., Castaneda S., and Jimenez S.A. (2009). Primary osteoarthritis no longer primary: three subsets with distinct etiological, clinical, and therapeutic characteristics. Semin. Arthritis Rheum. *39*: 71-80.
60. Ramos Y.F., Bos S.D., Lakenberg N., Bohringer S., den Hollander W.J., Kloppenburg M., Slagboom P.E., and Meulenbelt I. (2014). Genes expressed in blood link osteoarthritis with apoptotic pathways. Ann. Rheum.

Dis. *73*: 1844-1853.
61. Loughlin J. and Reynard L.N. (2015). Osteoarthritis: Epigenetics of articular cartilage in knee and hip OA. Nat. Rev. Rheumatol. *11*: 6-7.
62. Meulenbelt I., Kloppenburg M., Kroon H.M., Houwing-Duistermaat J.J., Garnero P., Hellio Le Graverand M.P., Degroot J., and Slagboom P.E. (2006). Urinary CTX-II levels are associated with radiographic subtypes of osteoarthritis in hip, knee, hand, and facet joints in subject with familial osteoarthritis at multiple sites: the GARP study. Ann Rheum Dis *65*: 360-365.
63. Meulenbelt I., Kloppenburg M., Kroon H.M., Houwing-Duistermaat J.J., Garnero P., Hellio-Le Graverand M.P., Degroot J., and Slagboom P.E. (2007). Clusters of biochemical markers are associated with radiographic subtypes of osteoarthritis (OA) in subject with familial OA at multiple sites. The GARP study. Osteoarthritis. Cartilage. *15*: 379-385.
64. Van Spil W.E., Jansen N.W., Bijlsma J.W., Reijman M., Degroot J., Welsing P.M., and Lafeber F.P. (2012). Clusters within a wide spectrum of biochemical markers for osteoarthritis: data from CHECK, a large cohort of individuals with very early symptomatic osteoarthritis. Osteoarthritis. Cartilage. *20*: 745-754.
65. Ramos Y.F., Hollander W.d., Lakenberg N., van der Breggen R., Bomer N., Kroon H., Kloppenburg M., Slagboom P., and Meulenbelt I. (2015). Risk prediction using epigenetic profiles in blood of osteoarthritis patients. Osteoarthritis and Cartilage *23, Supplement 2*: A73-A74.
66. Karouzakis E., Trenkmann M., Gay R.E., Michel B.A., Gay S., and Neidhart M. (2014). Epigenome analysis reveals TBX5 as a novel transcription factor involved in the activation of rheumatoid arthritis synovial fibroblasts. J. Immunol. *193*: 4945-4951.
67. Nakano K., Boyle D.L., and Firestein G.S. (2013). Regulation of DNA methylation in rheumatoid arthritis synoviocytes. J. Immunol. *190*: 1297-1303.
68. Park S.H., Kim S.K., Choe J.Y., Moon Y., An S., Park M.J., and Kim D.S. (2013). Hypermethylation of EBF3 and IRX1 genes in synovial fibroblasts of patients with rheumatoid arthritis. Mol. Cells *35*: 298-304.
69. Zhang Y., Fukui N., Yahata M., and Lee M. (2015). Genome-wide DNA methylation profiling of osteoarthritic subchondral bone. Osteoarthritis and Cartilage *23, Supplement 2*: A195.

8

GENERAL DISCUSSION

MAIN AIMS

The main aims of this thesis were to address several challenges OA research is currently opposed to. Firstly, given that the apparent link between OA associated genetic variation and gene expression, possibly mediated by CpG methylation, is frequently reported on by candidate gene studies (1–5), we set out to identify novel genetic variation that marks AI in articular cartilage on a transcriptome wide scale. In **chapter 2**, we performed genetic association analyses for OA with SNPs that both marked AI and were differentially expressed between preserved and paired OA affected cartilage. By doing so, we exploited the decreased multiple testing penalty that inherently obstructs canonical GWA studies and reported on novel SNPs that appear to predispose for OA. Further expansion on these *in-silico* derived results by *in-vitro* experiments revealed a potential causally degenerative mechanism by which genetic variation at and upregulation of the *CRLF1* gene predisposes for OA. Furthermore, in **chapter 3** we showed that the generated database of articular cartilage AI SNPs can contribute to canonical GWA studies, delivering increased mechanistic insight of OA susceptibility alleles.

Secondly, whereas it appears eminent that dysregulation of gene expression is linked with OA pathophysiology in articular cartilage, it remains unclear how these altered levels of transcription are brought about. Hence, in light of the described relation between the established OA susceptibility SNP rs225014 (6) and allele dependent transcription of *DIO2* in articular cartilage (5), we explored the regulatory properties of the gene and elaborate on the link between genetic variation, local DNA methylation levels and their influence on *DIO2* expression in an OA context in **chapter 4** (7).

Thirdly, given that DNA methylation appears to be involved in mediating the relation between OA associated genetic variation and gene expression levels (8–10), the methylome has gained increased interest among OA researchers (2,11–14). Additionally, OA associated methylomic profiles might pose an arguable candidate for maintaining the pathological transcriptomic profiles observed in OA affected articular cartilage (15–20). This is reflected by elaborate papers that have reported on extensive numbers of CpGs that mark the ongoing disease process by differential methylation between OA lesioned and preserved or healthy articular cartilage (21–23). Nonetheless, at the beginning of the research project described in this thesis, it remained unclear how to interpret these methylation differences in context of OA related transcriptomic profiles. By integration of (semi) genome wide SNP, methylation and expression data in **chapter 5** we reported on CpGs and SNPs that show presumptive evidence for (dys)regulating gene expression in OA affected articular cartilage. During these academic pursuits we encountered distinctly differing methylomic profiles between hip and knee articular cartilage, irrespective of the tissue's OA affection state. While considerable overlapping methylomic OA related differences were observed among both knee and hip samples, the presence

of marked joint specific profiles might indicate that the actual tissues should be viewed as distinct. These observations and possible implications are described in **chapter 6**.

GENETIC VARIATION AFFECTING GENE EXPRESSION IN ARTICULAR CARTILAGE

Despite our increased mechanistic understanding of established OA susceptibility genes in relation to the disease, a significant challenge that remains unaddressed is the substantial amount of missing heritability (24–27). In part, this can be explained by the existence of considerable phenotypic heterogeneity among OA patients. While this could be addressed by deeper phenotyping of possible OA subtypes, another approach to increase the likelihood of detecting unknown OA associated genetic variation is to adequately reduce the number of statistical tests by *a priori* discarding SNPs that are deemed less likely to contribute to OA pathophysiology, as described in **chapter 2**.

Having acquired whole transcriptome RNA-sequencing data from articular cartilage of primary OA patients (chapter 2 **Supplementary Table S1**), we exploited the mRNA sequence information to identify SNPs that mark imbalanced expression of the respective genes they are located in. Among a total of 13853 transcribed heterozygous SNPs (>25 counts and present in at least two individuals), we identified 2070 located in 1031 unique genes that significantly did so. Next, we analyzed whether these 1031 genes were in addition be subject to AI also significantly differentially expressed between preserved and paired OA affected cartilage. Indeed, 32 genes appeared to do so and we subsequently performed genetic association analysis for OA with the top ten AI SNPs (single most significant SNP per gene). Having substantially reduced the multiple testing correction penalty implied in canonical GWA studies, we were able to identify an otherwise discarded association signal, located in the *CRLF1* gene.

In light of the acquired regulatory, transcriptional and genetic evidence as well as having these technically and/or biologically replicated, we initiated mechanistic *in-vitro* studies addressing the suspected relation between *CRLF1* and OA. The respective protein products of *CRLF1* and its reported binding partner *CLCF1* assert their function through the ciliary neurotrophic factor receptor, encoded by the *CNTFR* gene. We treated primary human chondrocyte cultures with the heterodimerized protein complex and observed a marked upregulation of anabolic genes, while catabolic genes remained unaffected. Arguably paradoxical, Tsuritani *et al.* (2010) reported downregulation of anabolic and no effect on catabolic genes in similar experiments using a chondrogenic mouse cancer cell line (ATDC5) (28). Whether these apparent opposing results are due to species differences and/or a consequence of utilizing distinct cell types remains subject of future research, but it appears *CRLF1* is involved in establishing and/or maintaining OA associated, aberrant cartilage homeostasis.

Apart from the exploratory nature of chapter 2, the RNA sequencing dataset generated therein also complements ongoing GWA approaches, as we describe in **chapter 3**. In a GWA study and combined meta-analyses for replication on bilateral hand OA, we identified several coding variants within the *MGP* gene. By lack of hand articular cartilage samples, we interrogated the knee and hip articular cartilage RNA sequencing dataset and found significant AI of the *MGP* gene, marked by alleles of the OA associated SNP. Custom targeted assays were performed and validated this observation, not only in articular cartilage, but also in subchondral bone. Given that *MGP* is reported to be an inhibitor of cartilage calcification (29–32), high expression in articular cartilage needs to be warranted in order to maintain healthy tissue homeostasis. In concordance with the OA predisposing property of the identified risk allele, we found lower expression of *MGP* marked by this allele among heterozygotes, suggesting that carriers are more prone to develop OA due to their inability to halt cartilage calcification, a well described pathological manifestation of ongoing OA. Of note, in light of the observed imbalanced *MGP* expression marked by OA risk alleles in articular cartilage of knee and hip joints, one could hypothesize that these alleles are therefor likely to harbor a genetic association with other types of OA as well, in addition to the initial quantitative association with bilateral hand OA. While a lookup in hip and knee OA GWA data did not satisfy these expectations, one could argue that the observed AI is systemically present among multiple tissues, but exerts its downstream effects only in those that are susceptible to it. In parallel, the phenotypic heterogeneity among hip and knee OA patients might be larger when compared to patients in which bilateral hand OA was measured quantitatively.

In addition to the RNA sequencing dataset's successful applications described above, we were unable to replicate some other reported AI SNPs from literature in the context of OA affected articular cartilage. While we were able to verify AI of rs3204689 (risk allele C marked significant lower expression of *ALDH1A2*, Styrkarsdottir *et al*. (2015) (45)) as well as suggestive AI of rs225014 (3/4 heterozygotes revealed higher *DIO2* expression of the risk allele C, Bos *et al.* (2012) (5)), other genes reported to be subject to AI in articular cartilage such as *COL11A1* (rs1676486) (1), *GNL3* (rs11177) (3) and *SPCS1* (rs6617) (3) were not observed as such in our dataset. The absence of those SNPs among our results might have arisen due multiple reasons. Firstly, we were bound to address AI with the amount of respective heterozygotes present in our data, whereas the initial reports had gathered larger samples sizes. Secondly, the alleles of AI SNPs are not by definition the mechanistically driving source leading towards the observed transcriptional imbalance. In other words, the actual genetic variation that interferes with the transcription machinery might be positioned in relatively high LD elsewhere, for example in close or distal regulatory elements. If LD between the addressed SNP and respective regulatory element is low, however, one could expect inconsistent results and consequentially face difficulty in replicating earlier observations.

OA ASSOCIATED DNA METHYLATION PATTERNS

Given the large number of OA associated SNPs that seem to exert predisposition for OA through AI in cartilage, as well as the marked number if differentially expressed genes in OA affected cartilage when compared to control, we set out to interrogate possible regulatory mechanisms that might mediate these observations.

Although (semi-)genome wide technologies and respective analysis tools are rapidly emerging as affordable and accessible (33–37), the OA research field has also gained substantial knowledge from gene targeted measurements of regulatory mechanisms and expression in recent years. In this regard, the rs225014 T>C SNP located within the *DIO2* gene might confer susceptibility towards OA by multiple mechanisms, as it both marks an amino change (threonine to alanine), as well as AI of *DIO2* in articular cartilage as shown by Bos *et al*. (2012) (5). Given that the resulting amino acid change induced by the rs225014 OA susceptibility allele C did not directly result in altered turnover of inactive (T4) into active thyroid hormone (T3) (29), susceptibility seems mediated through increased *DIO2* expression, presumably leading to higher protein levels, marked by the rs225014 risk allele. Additionally, it was reported that *DIO2* is significantly higher expressed in articular cartilage among OA patients when compared to that of healthy controls (5). Hence, in **chapter 4** we set out to elucidate how *DIO2* expression is regulated in articular cartilage. Publically available CHiP-Seq data showed that CTCF, a DNA binding protein and regulator of gene expression, binds at the rs225014 locus in multiple cell types (ENCODE consortium) (38). Hence, we assessed whether the rs225014 alleles influence binding of CTCF in cultured primary human chondrocytes. We observed that CTCF protein does bind the DNA at the rs225014 locus in chondrocytes, however no relation between the respective alleles and the amount of bound CTCF was observed. Nonetheless, while allele dependent CTCF binding did not appear to explain the observed *DIO2* AI in articular cartilage, we did observe that three CpG dinucleotides were significantly differentially methylated between macroscopically preserved and paired lesioned articular cartilage of 52 OA patients (29 knee, 22 hip), independent of gender, age and affected joint. Among these, a single CpG dinucleotide >2kb upstream of the *DIO2* transcription start site associated significantly with *DIO2* expression. Interestingly, whereas overall this CpG is hypermethylated in OA affected compared to paired preserved cartilage, OA associated hypermethylation among carriers of the rs225014 risk allele C is significantly more pronounced when compared to homozygous wildtype carriers. Accordingly, *DIO2* expression was expressed higher in OA articular cartilage in general, but even more so among rs225014 risk allele carriers. These observations imply that both genetic and epigenetic variation at OA susceptibility loci, as well as the apparent interaction between the two, are mechanisms worthwhile to assess in order to deepen our understanding of OA associated gene expression in articular cartilage. This statement is further strengthened by efforts from other academic research groups studying these apparent regulatory

relationships. For example, Roach *et al.* (2005) reported that OA associated upregulation of multiple ECM degrading MMPs was marked by demethylation of CpGs located in respective promotor regions (39), Reynard *et al*. (2013, 2015) showed that DNA methylation associated with both basal expression as well as AI of *GDF5* marked by alleles of the OA susceptibility rs143383 (C>T) SNP (2,11) and Takahashi *et al*. (2015) has revealed that OA associated upregulation of *IL8* is accompanied by differences in DNA methylation at the gene locus (40). As such, probing the relation between genetic, epigenetic and transcriptional features, has delivered valuable insight into mechanisms of pathological gene expression observed in OA affected articular cartilage.

It deservers mentioning that when descriptive measurements are performed as described above it is inherently challenging to assert causality towards DNA methylation in terms of it leading to differences in expression. In addition, it remains unclear as to whether upregulation of *DIO2*, or other disease responsive genes for that matter, is causing OA associated cartilage degradation as opposed to *vice versa*. Due these matters, which are inherent to paired cross-sectional study designs, we will also be unable to detect genes involved in the disease's onset. This could arguably be addressed to certain extent by comparing articular cartilage transcriptional and/or methylation profiles across affected and unaffected individuals. This would, however, oppose another challenge by means of the distinct genetic backgrounds of the respective individuals. Mendelian randomization might additionally offer insight here into, but unfortunately requires substantially larger samples sizes than we were able to address here.

METHYLOMIC PROFILING OF OA AFFECTED ARTICULAR CARTILAGE

The development of affordable omics technologies (*e.g.* microarrays and next generation sequencing) has substantially influenced the rate at which new molecular biology discoveries are being reported. The OA research field is hereto no exception, and in **chapter 5** we have measured and integrated DNA methylation and gene expression data on semi genome wide scales among preserved and paired OA lesioned cartilage samples, followed by inspection of the respective genetic environment.

Given that the majority of CpGs located at a gene locus do not appear to be associated with expression of the respective gene, we first set out to identify CpGs in articular cartilage that do correlate with gene expression levels. Although a large number of statistical tests were performed, possibly leading to increased false positive results, we allowed ourselves to lessen the canonical multiple testing correction penalty (Bonferroni) when testing for correlations between methylation and expression (41). As opposed to reduce the significance threshold **α** by the total number of statistical tests, we reduced **α** separately for each gene, depending on the number of CpGs annotated to it. As such, we observed 3748 CpGs of which the extent of methylation correlated with expression of 2324 genes

in articular cartilage, independent of age, gender, affected joint and OA status. Interestingly, where increased methylation was canonically associated with decreased expression, we observed CpGs to harbor transcriptional associations in both directions with expression of proximal genes. Specifically we reported on 1741 negative (i.e. decreased expression alongside increased methylation) and 2007 positive correlations (i.e. increased expression alongside increased methylation). These observations indicate that, albeit challenging to assert definite causality, DNA methylation appears involved in general in mediating gene expression levels in articular cartilage. In parallel, however, we additionally observed that less than 3% (N=87) of these CpGs were also differentially methylated between preserved and paired OA affected samples. Interestingly, the genes near these specific CpGs are enriched for developmental pathways, not unlike those observed during endochondral ossification during fetal development. While this might still be a mere reflection of the ongoing pathophysiology, it does support the proposition that in an OA context, chondrocytes appear to dedifferentiate and start dividing, form columnar structures and actively start calcifying the ECM. Whereas chondrocytes are presumably required to actively remodel the ECM upon everyday stresses and potential microtraumas, this remodeling requires dynamic and timely regulation of both catabolic genes as well as ECM anabolic genes, to remain homeostasis. However, perpetual adjustment of DNA methylation is possibly prone to errors and might therefore affect gene expression inappropriately, leading to pathological remodeling of the ECM. Nonetheless, even if we are to refrain from interpreting these results in a directional context, we can state which actively expressed genes and/or pathways seem over- or underrepresented among chondrocytes residing in OA affected articular cartilage. This insight might aid in dissecting possible future therapeutic strategies that aim to revert chondrocytes residing OA affected cartilage back towards a healthy state.

Other groups have also reported on extensive numbers of CpGs that reflect the ongoing OA process in articular cartilage. Fernández-Tajes *et al.* (2013) were among the first to utilize microarrays (Illumina Infinium HumanMethylation27 BeadChip) to address DNA methylation differences in a non-targeted fashion in articular cartilage (22). They observed 91 CpGs that were significantly differentially methylated between knee OA cartilage and articular cartilage derived from fractured hip joints. Interestingly, upon principal component analysis (PCA) knee samples grouped together in two distinct clusters, which appear to be driven by 1357 CpGs that were enriched for inflammatory pathways. Interestingly, 450 genes located in close proximity of those CpGs revealed differential expression between the observed clusters, albeit by utilizing a partially overlapping microarray dataset in terms of samples. Also previous to our efforts, Rushton *et al.* (2014) considerably expanded on the work of Fernández-Tajes *et al.* by addressing over 450.000 CpGs (Illumina Infinium HumanMethylation450 BeadChip) in cartilage from both OA knee and OA hip joints, and also compared the acquired methylation levels to cartilage from fractured hip joints (21). While Rushton *et al.* did not

integrate their findings with gene expression data, they were able to replicate inflammatory enriched clusters, however, only among hip and not among knee OA patients. Furthermore, they observed that independent of the pathological state, hip and knee cartilage seem to harbor joint specific methylation profiles, reflected by 5547 differentially methylated CpGs between the two. Finally, methylation levels of 5322 CpGs comprised a distinct OA hip cartilage profile, when compared to control hip cartilage. These initial reports by Fernández-Tajes *et al.* and Rushton *et al.* have assigned a prominent role to DNA methylation in OA pathophysiology. Nonetheless, these elaborate studies did leave us challenges that we have aimed to address in the current thesis. Firstly, as Rushton *et al.* uncovered the presence of joint specific profiles independent of pathological status, the observed differentially methylated CpGs by Fernández-Tajes *et al.* might partially be driven by inherent joint related differences. Secondly, the presence of mQTLs will interfere when DNA methylation levels of OA cartilage are compared to separate, unrelated controls. And thirdly, although Fernández-Tajes *et al.* did intersect their results with gene expression data, it remained unclear whether there existed a direct correlation between the reported OA associated methylation profiles and respective *in-cis* expression.

Lastly, we integrated our methylation and transcription data with SNP data in order to identify to what extent the local genetic context contributes hereto. By doing so, we identified SNPs of which the alleles correlated with DNA methylation (N=36), transcription (N=26) or both (N=3). Given that mQTL and eQTL analyses generally require ample sample sizes, we cannot rule out the presence of possible false positives, but do propose that these SNPs are enriched for association signals that confer OA susceptibility. Perhaps more surprisingly, we were unable to observe alleles of SNPs that associated with DNA methylation and/or transcription levels as described in chapter 5, as well as additionally mark AI of the respective genes as described in chapter 2. While the lack of overlap could have arisen by a multitude of technical, analytical or even biological causes, and most likely has so due to a combination of them, a clear impediment in comparing the results from these two chapters is the lack of overlapping samples.

JOINT SPECIFIC ARTICULAR CARTILAGE DNA METHYLATION PROFILES

In our efforts to describe general OA related methylomic and associated transcriptomic profiles, we have corrected our analyses performed in chapter 5 for a number of variables, among them the joint from which articular cartilage was sampled. Correcting for joint was deemed necessary, as GWA studies increasingly report on SNPs that predispose for joint specific OA or related phenotypes (42–51). Even more so, in **chapter 6** we describe how hip and knee articular cartilage appear to harbor highly joint specific DNA methylation profiles. Exploratory analysis by means of dimension reduction (principal component analysis, PCA) of generated DNA methylation data revealed that

knee and hip articular cartilage samples formed tight clusters, independent of the tissues' disease state. As such, we applied a sliding window algorithm developed by Slieker *et al.* (52), to determine robust tissue specific DNA methylation profiles.

While stretches of differentially methylated CpGs were found near 245 genes, most strikingly were the substantial differences in homeobox (HOX) clusters and respective co-factors. HOX genes are well known for their role in development and specifically for controlling segmentation of the body plan along the head-tail axis (53,54). Expression of HOX genes during development is regulated in a spatiotemporal pattern, *i.e.* along the head-tail axis the four different HOX clusters express a different number of HOX genes in a linear pattern when development eventuates (55). While in adult tissues this spatiotemporal expression pattern has been described to remain to some extent (56), we found that in adult articular cartilage there is little to no linear correlation between the four HOX clusters. In other words, it appears as though somewhere during development and/or life, specific HOX genes have been activated differently between knee and hip articular cartilage and remain as such. Despite generic processes in knee and hip OA, it thus appears that alongside joint specific OA predisposing SNPs, joint specific DNA methylation and associated gene expression profiles might substantially contribute to OA pathophysiology differences between joints. In this regard, it might be a worthwhile pursuit to address these patterns in articular cartilage from other joints that are frequently affected by OA, such as the spine and hands. Also, these apparent joint related differences in methylation and expression of key developmental genes, raises the question whether efforts to generate cartilage *in-vitro* need to be tailored for the joint they aim to address.

FUTURE PERSPECTIVES

In the last years, we have seen considerable improvements within the biomedical scientific community to scale up experimental setups, both in terms of sample size as well as in the number of measured features, and it can be expected that this will continue on in the near future. It is eminent that these developments will not only aid researchers in directly identifying new disease associated genes, whether they do so by interrogating and/or integrating (epi)genetic variation, altered transcription or other type of omics platforms, but will also allow more complex analyses. Although complex analyses as such might not be the point by which the OA research field will benefit from the most, it is clear that the enormous amount of data that is currently produced across numerous research groups does contain information that is currently not exploited. Consequently, further research into developing appropriate bioinformatic tools to integrate multiple layers of data and sources is required to accurately interpret the descriptive data the field are currently generating, and will additionally open up the possibility to address other (epigenetic) layers of information as well. When able to do so, we might finally be able to thoroughly start investigating the proposed combined

effect of putative common genetic variants that are associated to OA by only relatively small extent. It should be noted, however, that if that is indeed the direction the field is heading towards, we need to accept that mechanistic follow-up experiments will become exceedingly complicated. Both *in-vitro* as well as *in-vivo* studies aimed at elucidating the combinatorial effect of multiple genetic variants or genes will be prone to noise, due to the fact that the background in which the proposed experiment will be performed (e.g. cell type, mouse strain, genetic background, etc.) will have a relatively large impact on the outcome, when compared to single, but large-effect perturbations.

However, seen from the patient's perspective, these proposed developments might not be the most obvious goals to pursue. Increasing our understanding of the complex pathological process of OA will undoubtedly help generate novel and better treatments on the middle to long term, and is a very worthwhile academic pursuit in its own. Nonetheless, dissecting the molecular genetic signals that are associated with OA might just not be the most opportunistic approach at the current time, in the context of developing treatments in the short term. In recent years, we have seen a substantial increase in our ability to generate cell cultures *in-vitro* that resemble the actual tissue *in-vivo*, using patient derived differentiated cells and/or induced pluripotent stem cells (iPSCs) (57,58). From a highly pragmatic point of view, if we are able to (re-)generate articular cartilage, using cells from the actual patient, we might not need a thorough understanding of the complex interplay between environment, genetics and molecular changes that appear to underlie and/or drive OA. Of course, given that we are currently unable to do so, efforts into generating *in-vitro* cell cultures that harbor the desired chondrocytic phenotype will benefit from descriptive studies that report on the dynamic states of articular chondrocytes *in-vivo*, not entirely unlike the chapters in this thesis.

REFERENCES

1. Raine E V, Dodd AW, Reynard LN, Loughlin J. Allelic expression analysis of the osteoarthritis susceptibility gene COL11A1 in human joint tissues. BMCMusculoskeletDisord. 2013;14(1471–2474):85.
2. 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. HumGenet. 2014 May 27;(1432–1203).
3. Gee F, Clubbs CF, Raine E V, Reynard LN, Loughlin J. Allelic expression analysis of the osteoarthritis susceptibility locus that maps to chromosome 3p21 reveals cis-acting eQTLs at GNL3 and SPCS1. BMCMedGenet. 2014;15(1471–2350):53.
4. Raine E V, Wreglesworth N, Dodd AW, Reynard LN, Loughlin J. Gene expression analysis reveals HBP1 as a key target for the osteoarthritis susceptibility locus that maps to chromosome 7q22. AnnRheumDis. 2012 Dec;71(1468–2060):2020–7.
5. Bos SD, Bovee J V, Duijnisveld BJ, Raine E V, van Dalen WJ, Ramos YF, et al. Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. AnnRheumDis. 2012 Jul;71(1468–2060):1254–8.
6. Meulenbelt I, Min JL, Bos S, Riyazi N, Houwing-Duistermaat JJ, van der Wijk HJ, et al. Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. HumMolGenet. 2008 Jun 15;17(1460–2083):1867–75.
7. Bomer N, den Hollander W, Ramos YFM, Bos SD, van der Breggen R, Lakenberg N, et al. Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis. Ann Rheum Dis. 2014;
8. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M. Mapping complex disease traits with global gene expression. Nature . 2009;10 (3):184–94.
9. Garnier S, Truong V, Brocheton J, Zeller T, Rovital M, Wild PS, et al. Genome-wide haplotype analysis of cis expression quantitative trait loci in monocytes. PLoSGenet. 2013;9(1553–7404):e1003240.
10. Westra HJ, Franke L. From genome to function by studying eQTLs. BiochimBiophysActa. 2014 May 4;(0006–3002 (Print)).
11. Reynard LN, Bui C, Canty-Laird EG, Young DA, Loughlin J. Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. HumMolGenet. 2011 Sep 1;20(1460–2083):3450–60.
12. Goldring MB, Marcu KB. Epigenomic and microRNA-mediated regulation in cartilage development, homeostasis, and osteoarthritis. Trends MolMed. 2012 Feb;18(1471–499X):109–18.
13. Barter MJ, Young DA. Epigenetic mechanisms and non-coding RNAs in osteoarthritis. CurrRheumatolRep. 2013 Sep;15(1534–6307):353.
14. Loughlin J. Genetic contribution to osteoarthritis development: current state of evidence. CurrOpinRheumatol. 2015 May;27(1531–6963):284–8.
15. Dongen J Van, Nivard MG, Willemsen G, Hottenga J, Helmer Q, Dolan C V, et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. Nat Commun . 2016;7:1–13.
16. Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. PLoSGenet. 2012;8(1553–7404):e1002629.
17. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. NatGenet. 2003 Mar;33 Suppl(1061–4036 (Print)):245–54.
18. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol . 2011;12 (1):R10.
19. Feinberg AP. Genome-scale approaches to the epigenetics of common human disease. Virchows Arch. 2010 Jan;456(1432–2307):13–21.
20. Witzmann SR, Turner JD, Meriaux SB, Meijer OC, Muller CP. Epigenetic regulation of the glucocorticoid receptor promoter 1 (7) in adult rats. Epigenetics. 2012 Nov;7(1559–2308):1290–301.
21. Rushton MD, Reynard LN, Barter MJ, Refaie R, Rankin KS, Young DA, et al. Characterization of the cartilage DNA methylome in knee and hip osteoarthritis. Arthritis Rheumatol. 2014 May 16;(2326–5205).
22. Fernandez-Tajes J, Soto-Hermida A, Vazquez-Mosquera ME, Cortes-Pereira E, Mosquera A, Fernandez-Moreno M, et al. Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster

of osteoarthritic patients. AnnRheumDis. 2014 Apr;73(1468–2060):668–77.
23. Zhang Y, Fukui N, Yahata M, Katsuragawa Y, Tashiro T, Ikegawa S, et al. Genome-wide DNA methylation profile implicates potential cartilage regeneration at the late stage of knee osteoarthritis. Osteoarthr Cartil. 2016;24 (5):835–43.
24. Maher B. Personal genomes: The case of the missing heritability. Nature. 2008 Nov 6;456(1476–4687):18–21.
25. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009 Oct 8;461(1476–4687):747–53.
26. Slatkin M. Epigenetic inheritance and the missing heritability problem. Genetics. 2009;182 (3):845–50.
27. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. Proc Natl Acad Sci U S A . 2012;109 (4):1193–8.
28. Tsuritani K, Takeda J, Sakagami J, Ishii A, Eriksson T, Hara T, et al. Cytokine receptor-like factor 1 is highly expressed in damaged human knee osteoarthritic cartilage and involved in osteoarthritis downstream of TGF-beta. CalcifTissue Int. 2010 Jan;86(1432–0827):47–57.
29. L.M. B, T.a. O, M.S. T, R. B, G.S. S, J.B. L. Developmental expression and hormonal regulation of the rat matrix GLA protein (MGP) gene in chondrogenesis and osteogenesis. J Cell Biochem . 1991;46 (4):351–65.
30. Marulanda J, Gao C, Roman H, Henderson JE, Murshed M. Prevention of arterial calcification corrects the low bone mass phenotype in MGP-deficient mice. Bone. 2013;57 (2):499–508.
31. Borr??s T, Comes N. Evidence for a calcification process in the trabecular meshwork. Vol. 88, Experimental Eye Research. 2009. p. 738–46.
32. Luo G, Ducy P, McKee MD, Pinero GJ, Loyer E, Behringer RR, et al. Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein. . Vol. 386, Nature. 1997. p. 78–81.
33. Van Iterson M, Tobi EW, Slieker RC, Den Hollander W, Luijk R, Slagboom PE, et al. MethylAid: Visual and interactive quality control of large Illumina 450k datasets. Bioinformatics. 2014;30 (23):3435–7.
34. Ramos YFM, Den Hollander W, Bovée JVMG, Bomer N, Van Der Breggen R, Lakenberg N, et al. Genes involved in the osteoarthritis process identified through genome wide expression analysis in articular cartilage; the RAAK study. PLoS One. 2014;9 (7).
35. Ramos YFM, Bos SD, Lakenberg N, Böhringer S, den Hollander WJ, Kloppenburg M, et al. Genes expressed in blood link osteoarthritis with apoptotic pathways. Ann Rheum Dis. 2013;
36. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics. 2008 Jul 1;24(1367–4811):1547–8.
37. Goeman JJ, van de Geer SA, de KF, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics. 2004 Jan 1;20(1367–4803 (Print)):93–9.
38. Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, et al. ENCODE Data in the UCSC Genome Browser: Year 5 update. Nucleic Acids Res. 2013;41(D1).
39. Roach HI, Aigner T, Kouri JB. Chondroptosis: a variant of apoptotic cell death in chondrocytes? Apoptosis. 2004 May;9(1360–8185 (Print)):265–77.
40. Takahashi A, de Andrés MC, Hashimoto K, Itoi E, Oreffo ROC. Epigenetic regulation of interleukin-8, an inflammatory chemokine, in osteoarthritis. Osteoarthr Cartil. 2015;23 (11):1946–54.
41. Dunn OJ. Multiple Comparisons Among Means. J Am Stat Assoc. 1961;56 (293):52–64.
42. Valdes AM, Styrkarsdottir U, Doherty M, Morris DL, Mangino M, Tamm A, et al. Large scale replication study of the association between HLA class II/BTNL2 variants and osteoarthritis of the knee in European-descent populations. PLoSOne. 2011;6(1932–6203):e23371.
43. Valdes AM, Spector TD. The genetic epidemiology of osteoarthritis. CurrOpinRheumatol. 2010 Mar;22(1531–6963):139–43.
44. Zeggini E, Panoutsopoulou K, Southam L, Rayner NW, Day-Williams AG, Lopes MC, et al. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. Lancet. 2012 Sep 1;380(1474–547X):815–23.
45. Styrkarsdottir U, Thorleifsson G, Helgadottir HT, Bomer N, Metrustry S, Bierma-Zeinstra S, et al. Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene and with rare variants at 1p31. NatGenet. 2014 May;46(1546–1718):498–502.
46. Valdes AM, Spector TD, Doherty S, Wheeler M, Hart DJ, Doherty M. Association of the DVWA and GDF5 polymorphisms with osteoarthritis in UK populations. AnnRheumDis. 2009 Dec;68(1468–2060):1916–20.
47. Valdes AM, Spector TD, Tamm A, Kisand K, Doherty SA, Dennison EM, et al. Genetic variation in the SMAD3

gene is associated with hip and knee osteoarthritis. Arthritis Rheum. 2010 Aug;62(1529–0131):2347–52.
48. Ingvarsson T, Hagglund G, Lohmander LS. Prevalence of hip osteoarthritis in Iceland. AnnRheumDis. 1999 Apr;58(0003–4967 (Print)):201–7.
49. Valdes AM, Hart DJ, Jones KA, Surdulescu G, Swarbrick P, Doyle D V, et al. Association study of candidate genes for the prevalence and progression of knee osteoarthritis. Arthritis Rheum. 2004 Aug;50(0004–3591 (Print)):2497–507.
50. Kerkhof HJ, Meulenbelt I, Akune T, Arden NK, Aromaa A, Bierma-Zeinstra SM, et al. Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. OsteoarthritisCartilage. 2011 Mar;19(1522–9653):254–64.
51. Ramos YF, Metrustry S, Arden N, Bay-Jensen AC, Beekman M, de Craen AJ, et al. Meta-analysis identifies loci affecting levels of the potential osteoarthritis biomarkers sCOMP and uCTX-II with genome wide significance. JMedGenet. 2014 Sep;51(1468–6244):596–604.
52. Slieker RC, Bos SD, Goeman JJ, Bovee J V, Talens RP, van der Breggen R, et al. Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. EpigeneticsChromatin. 2013;6(1756–8935):26.
53. Ladam F, Sagerstr??m CG. Hox regulation of transcription: More complex(es). Vol. 243, Developmental Dynamics. 2014. p. 4–15.
54. Cerd??-Esteban N, Spagnoli FM. Glimpse into Hox and tale regulation of cell differentiation and reprogramming. Vol. 243, Developmental Dynamics. 2014. p. 76–87.
55. Richardson MK. The Hox complex- An interview with Denis Duboule. Vol. 53, International Journal of Developmental Biology. 2009. p. 717–23.
56. Chang HY, Chi J-TJ, Dudoit S, Bondre C, van de Rijn M, Botstein D, et al. Diversity, topographic differentiation, and positional memory in human fibroblasts. Proc Natl Acad Sci . 2002;99 (20):12877–82.
57. Diederichs S, Shine KM, Tuan RS. The promise and challenges of stem cell-based therapies for skeletal diseases: stem cell applications in skeletal medicine: potential, cell sources and characteristics, and challenges of clinical translation. Bioessays. 2013 Mar;35(1521–1878):220–30.
58. Bomer N, den HW, Suchiman H, Houtman E, Slieker RC, Heijmans BT, et al. Neo-cartilage engineered from primary chondrocytes is epigenetically similar to autologous cartilage, in contrast to using mesenchymal stem cells. OsteoarthritisCartilage. 2016 Mar 17; (1522–9653).

9

Nederlandse Samenvatting

Dankwoord

Curriculum vitae

Publicatielijst

NEDERLANDSE SAMENVATTING

Wat betreft de kennisoverdracht naar collega's hebben wetenschappers een min-of-meer goed werkend systeem, namelijk het schrijven, onderling toetsen en uiteindelijk publiceren van resultaten in academische tijdschriften en het presenteren ervan op conferenties. Tegelijkertijd zijn wetenschappers lang niet altijd even goed in staat om uit te leggen waarom wij doen wat we doen. En dat terwijl ik het wel ethisch noodzakelijk acht dat wetenschappelijke bevindingen, welke niet tot stand had kunnen komen zonder geld en middelen uit de samenleving, hun weg terug moet vinden daarnaartoe. Daarom zal ik in deze sectie pogen om het gedane werk en de daaruit volgende resultaten, beschreven in het wetenschappelijke deel van dit proefschrift, zodanig uiteen te zetten opdat ik een graag geziene gast blijf op niet-wetenschappelijke bijeenkomsten.

Artrose is een aandoening van de gewrichten, gekarakteriseerd door vernauwing van de gewrichtsspleet, ontsteking en onomkeerbare afbraak en verbening van het kraakbeen. Gedurende het leven zijn kraakbeen cellen (chondrocyten) in staat om beschadigingen in het kraakbeen te herstellen door actief het extracellulaire weefsel af te breken en weer op te bouwen. Het vermogen van chondrocyten om het gewrichtskraakbeen te herstellen is echter niet onuitputtelijk en na verloop van tijd zullen beschadigingen zich ophopen in het kraakbeen. Hierdoor ervaren patiënten een geleidelijke afname van functionaliteit van het gewricht, wat zich uit in verlies van bewegingsvrijheid en uiteindelijk in pijn. Ondanks relatief succesvolle behandelmethoden voor patiënten in het eindstadium van de ziekte, namelijk het vervangen van het gewricht door een prothese, zijn we niet in staat om de ziekte in een vroeg, symptoomvrij stadium te herkennen en/of te voorkomen dat er verdere schade aan het gewricht optreedt. Dit proefschrift richt zich op het in kaart brengen van de processen in chondrocyten die bijdragen aan mechanismen waarmee deze cellen kraakbeenschade proberen te herstellen; mechanismen die op latere leeftijd verloren lijken te gaan.

Eerder is beschreven dat chondrocyten in artrotisch kraakbeen een afwijkend genexpressie profiel hebben; genen betrokken bij het afbreken en opbouwen van het weefsel lijken niet meer in balans. Eén van de mechanismen waarmee cellen hun genexpressie profiel dynamisch aanpassen in reactie op veranderingen in hun omgeving, is het aanbrengen of verwijderen van methyl-groepen op specifieke locaties in het genoom. Het (de)methyleren van deze cytosine-guanine nucleotide paren (CpGs) gaat samen met verschillen in expressie van lokale genen, ogenschijnlijk doordat DNA-gebonden methylgroepen interfereren met het binden van genexpressie regulerende eiwitten aan het DNA (transcriptie factoren).

Artrose gerelateerde verschillen in genexpressie kunnen echter het gevolg zijn van een groot aantal mechanismen. Wanneer men uitsluitend kijkt naar genen die verschillend tot expressie komen samen met een afwijkend lokaal methylatie profiel, blijken dit genen te zijn die met name betrokken zijn bij de ontwikkeling van bot tijdens de embryonale ontwikkeling, ook wel calcificatie genoemd. Deze observatie is in lijn met artrose geassocieerde veranderingen van het kraakbeen, namelijk de afbraak en calcificatie van het weefsel. De verrichtte metingen en analyses kunnen geen onderscheid maken tussen oorzaak en gevolg, daar het niet duidelijk is of permante kraakbeen schade voorafgaat aan verschillen in methylatie en/of genexpressie of juist het gevolg ervan is. Er kan wel gesteld worden dat chondrocyten in artrotisch kraakbeen niet langer in staat zijn om het omringende weefsel terug te brengen in de functionele staat. Het is waarschijnlijk dat chondrocyten in deze permanente, kraakbeen afbrekende staat terecht zijn gekomen doordat gedurende het leven de cellen hun methylatie profiel niet volledig kunnen terugbrengen van een actieve naar een rustende staat nadat opgedane schade is hersteld.

Doordat in ons onderzoek kraakbeen is bestudeerd uit verschillende soorten gewrichten, namelijk knieën en heupen, konden we bestuderen of er naast artrose geassocieerde methylatie verschillen ook locaties op het genoom bestaan die gewricht specifieke methylatie patronen laten zien. Ondanks de vele overeenkomsten tussen knie en heup kraakbeen, in zowel gezond als ziek weefsel in zowel micro- als macroscopisch opzicht, blijken met name de homeobox (HOX) genen en hun geassocieerde cofactoren opvallend verschillend gemethyleerd te zijn tussen de twee gewrichten. Dat is opmerkelijk, omdat van HOX-processen wordt aangenomen dat ze een cruciale rol spelen gedurende de ontwikkeling van het embryo, maar minder relevant zijn gedurende het leven. Ongeacht of deze verschillen tussen gewrichten betrokken zijn bij het artrose proces, zou dit kunnen betekenen dat toekomstige therapieën gericht op kraakbeen regeneratie zich zullen moeten toespitsen op het type aangedane gewricht.

Naast de invloed van externe factoren en weefselspecifieke eigenschappen, is de mate van methylatie ook onderhevig aan de lokale genetische context. Met andere woorden, de aanwezigheid van genetische variatie, over het algemeen bestudeerd als puntmutaties of single nucleotide polymorphisms (SNPs), heeft invloed op de methylering van het genoom en daarmee ook op de expressie van lokale genen. Deze eigenschap lijkt ook in artrose een substantiële rol te spelen. SNPs die geassocieerd zijn met een verhoogd risico op het krijgen van artrose, lijken bepalend te zijn voor de mate van expressie van genen die dichtbij liggen op het genoom, ogenschijnlijk gemedieerd door methylatie. Eén van de uitdagingen waar onderzoekers voor staan in de zoektocht naar nieuwe genetische variatie die de aanleg voor artrose verhoogt is van technische aard. De significantie van toetsen die de aanwezigheid van specifieke genetische variatie vergelijken tussen groepen artrose patiënten en gezonde

mensen is in grote mate afhankelijk van het aantal genetische varianten dat wordt getest. Als men alle SNPs in het genoom zou toetsen (genoom wijd) zouden SNPs significant kunnen verschillen puur op basis van kans en niet omdat de genetische variant echt bij artrose betrokken is. Daarom selecteren wij op voorhand uit de genoom wijde SNP set die variatie die een effect heeft op methylering en/of genexpressie in kraakbeen (potentieel functionele SNPs). Vervolgens wordt deze selectie SNPs getoetst op de relatie met artrose (meer of minder aanwezig in patienten dan in controles). Door deze zogenaamde kanskapitalisatie te verkleinen hebben we nieuwe genen kunnen identificeren die een rol lijken te spelen bij het ontstaan en/of ziekteverloop van artrose.

Voor het bestuderen van het effect van SNPs op lokale methylatie en/of genexpressie profielen zijn grote studiepopulaties nodig. Dit komt onder meer omdat deze profielen sterk tussen mensen verschillen ongeacht de aanwezigheid van genetische variatie of artrtose. Met nieuwe sequencing technologieën kunnen zowel de hoeveelheid als de volgorde van nucleotiden van de tot expressie komende genen kwantitatief worden bepaald in tegenstelling tot enkel een kwalitatieve maat van de hoeveelheid. Daardoor is men tegenwoordig in staat om te detecteren of genetische variatie invloed heeft op genexpressie in een enkel individu. Van alle genen draagt ieder mens twee kopieën (één van beide ouders) en wanneer slechts een enkele van deze twee kopieën een genetische variant herbergt, oftewel de drager ervan heterozygoot is, kan men bepalen in welke mate de twee kopieën tot expressie komen in verhouding tot elkaar.

Door alle varianten die ongelijk tot expressie komen in het kraakbeen van meerdere artrose patiënten te vergelijken met de lijst met genen waarvan we zien dat ze verschillend tot expressie komen tussen artrotisch en normaal kraakbeen (ongeacht of deze verschillen samengaan met afwijkende methylatie profielen), pogen we een selecte groep genetische varianten te presenteren die naar alle waarschijnlijkheid een invloed hebben op de kraakbeenhuishouding.

Eén van de genen die uit deze analyses naar voren kwam waarvooralsnog weinig over bekend was in relatie tot artrose en kraakbeenonderhoud, lijkt een cruciale rol te spelen in de progressie van artrose: het cytokine recptor-like factor 1 (*CRLF1*) gen. Dit gen blijkt significant hoger tot expressie te komen in artrotisch kraakbeen. Bovendien komt een genetische variant die lagere expressie teweeg brengt in heterozygote dragers minder vaak voor bij artrose patiënten vergeleken met de gezonde populatie. Gekweekte chondrocyten bleken na het toevoegen van het CRLF1 eiwitproduct een afwijkend expressie patroon te laten zien van genen die betrokken zijn bij de kraakbeenhuishouding. Dit is slechts een enkel voorbeeld van de ontdekking van een artrose gen dat hoogstwaarschijnlijk niet gevonden had kunnen worden als niet op voorhand zou zijn geselecteerd op potentieel functionele SNPs zoals hierboven beschreven. Ook bij de interpretatie van gevonden genen in genoom

wijde genetische associatie studies, waarin miljoenen SNPs worden getest op associatie met artrose zonder een a priori selectie, is deze lijst met potentiele kraakbeen functionele SNPs van toevoegende waarde. We vonden ook artrose geassocieerde SNPs waarvan eerder al is aangetoond dat heterozygote dragers de twee kopieën van het gen niet gelijk tot expressie brengen (aldehyde dehydrogenase 1 family memer A2, *ALDH1A2*), en SNPs gelegen in genen waarvan we weten dat ze het risico op artrose verhogen, maar waarvan in eerste instantie niet bekend was hoe die associatie mechanistisch tot stand zou kunnen komen (matrix Gla protein, *MGP*).

DANKWOORD

Zo, het heeft even mogen duren, maar het mag dan na een kleine zes jaar een naam hebben. Ik zou men de illusie kunnen voorhouden dat het eindproduct enkel door mijn toedoen tot stand is gekomen, maar hiermee zou ik de mensen om mij heen onnodig tekortdoen.

Allereerst wil ik Eva bedanken voor wie ze is. Ik had me dit academische avontuur niet kunnen voorstellen zonder jou aan mijn zijde. Als er iemand is van wie ik veel heb geleerd, ben jij het wel.

Twijfelend of ik het dankwoord zou ophangen aan een academische, ludieke dan wel sociale kapstok of een combinatie daarvan, is er bij uitstek één iemand die aan alle beschrijvingen voldoet. Nils, de marmelade kon je niet mondig genoeg. Bovendien hebben de hoofdstukken in dit proefschrift dankzij jouw wet-lab affiniteit aanmerkelijk meer shine gekregen en wil ik je bedanken voor alle goede herinneringen, uit elk van de drie domeinen, die voor mij onlosmakelijk aan dit proefschrift zijn verbonden.

Ik wil mijn promotoren, Ingrid en Eline, bedanken om in mij het vertrouwen te hebben gesteld het promotietraject tot een goed einde te breien. Ingrid, dankzij jouw advies en sturing heb ik mij in de afgelopen jaren kunnen ontwikkelen tot zelfstandig wetenschapper met een kritische blik. Met name je pragmatische, *cut-to-the-case* mentaliteit vond ik een inspirerende manier om wetenschap mee te bedrijven. Eline, het is vaker dan eens voorgekomen dat ik met trots de laatste versie van een manuscript aan je heb voorgelegd, waarin je vervolgens binnen mum van tijd de nodige struikelpunten wist te benoemen. Eenmaal aan de slag met je commentaren, bij tijd en wijle schoorvoetend, bleken ze echter consequent de manuscripten naar een hoger niveau te tillen. Eveneens mag de rol van mijn copromotor niet worden onderschat. Yolande, ik vond het ontzettend leuk om te wetenschappen met iemand in hetzelfde schuitje, maar met beduidend meer ervaring. Bovendien hebben we er veel bij gelachen, waar ik je voor wil bedanken.

In de jaren bij Moleculaire Epidemiologie ben ik in een vroeg stadium het lab uitgegaan, om me te kunnen storten op de toenemende digitale noodzakelijkheden binnen de levenswetenschappen. Het knallen van getallen heeft me als jonge bioloog niet meteen kunnen bekoren, tot ik in aanraking kwam met *in-silico* enthousiastelingen zoals Erik, Matthijs & Elmar. Zonder meastRo's als jullie in de directe nabijheid, zou het me een stuk minder snel en minder leuk zijn afgegaan om mezelf de bioinformatische kneepjes eigen te maken. Tegelijkertijd zijn er natuurlijk data nodig om die kneepjes op los te kunnen laten, en wil ik bij deze de analisten Ruudje, Eka, Wesley en Nico, alsmede de studenten Annelies en Peter bedanken voor het meedenken over en het genereren van. Noemenswaardig is ook de secretariële inzet van Inge, welke ervoor heeft gezorgd dat ik altijd op een vers wachtwoord

kon rekenen als die weer eens verlopen was. Evengoed verdienen ook de afdeling orthopedie en haar patiënten een plek in het dankwoord; zonder hen zou de afstand tussen de voor- en achterkant van dit proefschrift een stuk kleiner zijn geweest. In het verlengde van die voor- en achterkant, wil ik John bedanken voor het ontwerp ervan.

Ik wil de begeleider van mijn afstudeerstage bedanken, zonder wiens enthousiasme ik niet zou zijn begonnen aan een promotieonderzoek. Peter uit Amsterdam, door jouw nuchtere, en tegelijkertijd extravagante verschijning binnen de wetenschap realiseerde ik me dat er na de studietijd nog veel meer academische pret in het verschiet zou liggen. Laat de totstandkoming van dit proefschrift daar het bewijs van zijn.

De leden van de promotiecommissie wil ik bedanken voor het lezen van en het wisselen van gedachte over het proefschrift.

Als laatste wil ik alle niet-bij-naam-genoemde (ex-)collega's, vrienden en familie bedanken welke tijdens werkbesprekingen, vakanties, verjaardagen, congressen, koffiepauzes, vriendenweekenden en/of borrels mijn tijd als promovendus zo mooi hebben gekleurd en beïnvloed zoals jullie dat hebben gedaan.

CURRICULUM VITAE

On the 26th of September 1987, Wouter den Hollander was born in the coastal village of Noordwijk located in the Netherlands. He attended Northgo College, acquiring his Voorbereidend Wetenschappelijk Onderwijs degree in 2005, which allowed him to pursue his academic interests in biology. Having acquired a bachelor's degree in biology at Leiden University in 2009, his enthusiasm got sparked by molecular biology and genetics. These interests led him to obtain a master's degree in molecular cell biology at Leiden University in 2011, during which he worked on the identification of genes involved in regulation of the hypothalamic-pituitary-adrenal axis. Fascinated by the apparent molecular mechanisms that underlie regulation of gene expression, he continued his academic journey studying the relation between the genome, methylome and transcriptome in the context of osteoarthritis. Having gotten inspired by the advances in bioinformatics, he ventured out of the laboratory into a seat behind the computer at the Medical Statistics and Bioinformatics department of the Leiden University Medical Center. Being a PhD student from 2011 onward and situated in the Molecular Epidemiology group, supervised by Prof. Dr. Ingrid Meulenbelt and Prof. Dr. P Eline Slagboom, he swiftly adapted the computational aptitude to dissect the aforementioned molecular relation systemically. The consequent academic output gained notable attention in the osteoarthritis research field, resulting in often cited peer-reviewed articles and international scientific presentations. Hereafter, Wouter obtained a post-doctoral position in 2016 at the Drug Safety Sciences group of Prof. Dr. Bob van de Water, embedded in the Leiden Academic Center for Drug Research. While further expanding the analytical skillset he attained during his PhD project, he developed algorithms that predict clinical drug induced liver injury through *in vitro* high-density transcriptomic and microscopy data. As of 2018, Wouter is employed as a research data analist at the Trimbos Instituut, combining his data analysis expertise with his latent interest in psychoactive substance use.

PUBLICATIELIJST

Genome-wide association and functional studies identify a role for matrix Gla protein in osteoarthritis of the hand.
den Hollander W*, Boer CG*, Hart DJ, Yau MS, Ramos YFM, Metrustry S, Broer L, Deelen J, Cupples LA, Rivadeneira F, Kloppenburg M, Peters M, Spector TD, Hofman A, Slagboom PE, Nelissen RGHH, Uitterlinden AG, Felson DT, Valdes AM, Meulenbelt I, van Meurs JJB.
Annals of the Rheumatic Diseases. (*2017*)

The transcriptional landscape of age in human peripheral blood.
Peters MJ, Joehanes R, Pilling P, Schurmann C, Conneely KN, Powell J, Reinmaa E, Sutphin GL, Zhernakova A, Schramm K, Wilson YA, Kobes S, Tukiainen T, NABEC/UKBEC Consortium, Ramos YFM, Göring HHH, Fornage M, Liu Y, Gharib SA, tranger BE, De Jager PL, Aviv A, Levy D, Murabito JM, Munson PJ, Huan T, Hofman A, Uitterlinden AG, Rivadeneira F, van Rooij J, Stolk L, Broer L, Verbiest MMPJ, Jhamai M, Arp P, Metspalu A, Tserel L, Milani L, Samani NJ, Peterson P, Kasela S, Codd V, Peters A, Ward-Caviness CK, Herder C, Waldenberger M, Roden M, Singmann P, Zeilinger S, Illig T, Homuth G, Grabe HJ, Völzke H, Steil L, Kocher T, Murray A, Melzer D, Yaghootkar H, Bandinelli S, Moses EK, Kent JW, Curran JE, Johnson MP, Williams- Blangero S, Westra H, McRae AF, Smith JA, Kardia SLR, Hovatta I, Perola M, Ripatti S, Salomaa V, Henders AK, Martin NG, Smith AK, Mehta D, Binder EB, Nylocks KM, Kennedy EM, Klengel T, Ding J, Suchy-Dicey AM, Enquobahrie DA, Brody J, Rotter JI, Chen YI, Houwing-Duistermaat J, Kloppenburg M, Slagboom PE, Helmer Q, **den Hollander W**, Bean S, Raj T, Bakhshi N, Wang QP, Oyston LJ, Psaty BM, Tracy RP, Montgomery GW, Turner ST, Blangero J, Meulenbelt I, Ressler KJ, Yang J, Franke L, Kettunen J, Visscher PM, Neely GG, Korstanje R, Hanson RL, Prokisch H, Ferrucci L, Esko T, Teumer A, van Meurs JBJ, Johnson AD.
Nature Communications. (*2015*)

DNA methylation in osteoarthritis.
den Hollander W, Meulenbelt I.
Current Genomics. (*2015*)

Transcriptional associations of osteoarthritis-mediated loss of epigenetic control in articular cartilage.
den Hollander W, Ramos YF, Bomer N, Elzinga S, van der Breggen R, Lakenberg N, de Dijcker WJ, Suchiman HE, Duijnisveld BJ, Houwing-Duistermaat JJ, Slagboom PE, Bos SD, Nelissen RG, Meulenbelt I.
Arthritis & Rheumatology. (*2015*)

The effect of forced exercise on knee joints in Dio2-/- mice: type II iodothyronine deiodinase- deficient mice are less prone to develop OA-like cartilage damage upon excessive mechanical stress.
Bomer N, Cornelis FM, Ramos YF, **den Hollander W**, Storms L, van der Breggen R, Lakenberg N, Slagboom PE, Meulenbelt I, Lories RJ.
Annals of the Rheumatic Diseases. (*2014*)

Knee and hip articular cartilage have distinct epigenomic landscapes: implications for future cartilage regeneration approaches.
den Hollander W, Ramos YF, Bos SD, Bomer N, van der Breggen R, Lakenberg N, de Dijcker WJ, Duijnisveld BJ, Slagboom PE, Nelissen RG, Meulenbelt I.
Annals of the Rheumatic Diseases. (*2014*)

MethylAid: visual and interactive quality control of large Illumina 450k datasets.
van Iterson M, Tobi EW, Slieker RC, **den Hollander W**, Luijk R, Slagboom PE, Heijmans BT.
Bioinformatics. (*2014*)

Genes involved in the osteoarthritis process identified through genome wide expression analysis in articular cartilage; the RAAK study.
Ramos YF*, **den Hollander W***, Bovée JV, Bomer N, van der Breggen R, Lakenberg N, Keurentjes JC, Goeman JJ, Slagboom PE, Nelissen RG, Bos SD, Meulenbelt I.
PLoS One. (*2014*)

Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis.
Bomer N*, **den Hollander W***, Ramos YF*, Bos SD, van der Breggen R, Lakenberg N, Pepers BA, van Eeden AE, Darvishan A, Tobi EW, Duijnisveld BJ, van den Akker EB, Heijmans BT, van Roon-Mom WM, Verbeek FJ, van Osch GJ, Nelissen RG, Slagboom PE, Meulenbelt I.
Annals of the Rheumatic Diseases. (*2014*)

Genes expressed in blood link osteoarthritis with apoptotic pathways.
Ramos YF, Bos SD, Lakenberg N, Böhringer S, **den Hollander W**, Kloppenburg M, Slagboom PE, Meulenbelt I.
Annals of the Rheumatic Diseases. (*2014*)

* Equal contribution / shared first author