

Supplementary Materials

Text:

Supplementary Text S1 Detailed description of GWAS, replication, meta-analysis and study cohorts

Figures:

Supplementary Figure S1 Forrest plot for KLSum-score GWAS tophit rs4764133.

Supplementary Figure S2 Expression and eQTL MGP, ERP27,C12orf60, ART4 and C12orf69 in articular cartilage

Supplementary Figure S3 ASE results MGP

Supplementary Figure S4 : QQplot of GWAS P-values of the discovery cohorts

Tables:

Supplementary Table S1 Quality control and inclusion criteria: See Excel file Supplementary Table S1

Supplementary Table S2 General characteristics of RAAK cartilage samples

Supplementary Table S3 Baseline characteristics of the studies included in the analyses of osteoarthritis of the hand

Supplementary Table S4 Association of rs4764133 and other OA phenotypes Supplementary Table S54 Variants in the ALDH1A2 gene and their association with the KLSum-score measure.

Supplemental Material S1 Text

Detailed Description of Study Cohorts

Ethics Statement

All of the participating studies were approved by medical ethics committees of the participating centers, and all participants gave their written informed consent before entering the participating study.

Genome-wide association study and meta-analysis

We conducted genome-wide association studies to osteoarthritis of the hand, using a semi quantitative phenotype, the KLsum-score. The discovery stage consisted of the Rotterdam Study I (RS-I), Rotterdam Study II (RS-II) and Rotterdam Study III (RS-III), we used linear regression models adjusted for age-, gender and population stratification. The discovery cohorts were quality checked using EASYQC and genomic control correction was applied to the standard errors and P-values. The genomic control inflation factors were low for all three RS groups ($\lambda=1.003$, 1.006 and 1.011, respectively) and we generated a quantile-quantile plot (Supplementary Figure S4), both indicated no substantial population stratification due to cryptic relatedness, population substructure or other biases. We combined the discovery cohorts in a joined meta-analysis using inverse variance weighting with METAL. Manhattan plot and QQ-plot were generated using the joint-meta analysis data from the discovery phase using R and R package qqman. SNPs with a $P\text{-value} \leq 5 \times 10^{-6}$ were selected for replication. The top SNP for each independent locus was taken for replication in three studies ($n=4,011$): Leiden hand OA studies (Genetics osteoARthritis and progression (GARP) study & the Leiden Longevity Study), the TwinsUK (TUK) study cohort, and the Framingham Heart Study (FHS) cohort. Summary level data from the discovery and replication stage were combined in a joined meta-analysis using inverse variance weighting with METAL. For the top SNPs associated with hand osteoarthritis we have plotted the regional association plot using LocusZoom.

Rotterdam study. The Rotterdam Study is a population-based prospective cohort study ongoing since 1990 to study determines of chronic disabling diseases¹. The Rotterdam Study I (RS-I) is the first cohort of 7,983 persons living in the Ommoord district of Rotterdam in the Netherlands. All subjects were aged 55 years and older,

recruitment of participants started in 1990. The Rotterdam Study II (RS-II) started in 1999 when 3,011 participants moved into the study since they became 55 years of age or moved into the study district. The Rotterdam Study III (RS-III) started in 2006 with all 3,932 participants aged 45 years and older from the study district not yet included in the study.

Genotyping of the samples in RS-I and RS-II was carried out with the Illumina HumanHap 550v3 Genotyping BeadChip, for RS-III the Illumina HumanHap610-Quad V1 was used. The Beadstudio GenCall algorithm was used for genotype calling and quality control procedures. The following quality control inclusion filters were applied: call rate $\geq 97,5\%$, $MAF \geq 1\%$, P for Hardy-Weinberg equilibrium $< 1 \cdot 10^{-6}$ (Table S5). The total number of genotyped SNPs that passed these filters was; RS-I=512,349, RS-II=466,389 and RS-III=517,658. Imputation was done against the 1000G V.3 reference panel using the maximum likelihood method implemented in MACH (<http://www.sph.umich.edu/csg/abecasis/MACH/download/1000G.2012-03-14.html>). Analysis of imputed genotype data accounted for uncertainty in each genotype prediction by using the dosage information from MACH. For this analysis we used MACH2QTL via GRIMP², which uses genotype dosage value (0-2, continuous) as a predictor in a linear regression framework and using standardized age-, gender and population stratification adjusted residuals from linear regression. The summary statistics of RS-I, RS-II and RS-III were meta-analysed using METAL applying inverse variance methodology assuming fixed effects with Cochran's Q and 12 metrics used to quantify between-study heterogeneity³. The medical ethics committee of Erasmus University Medical School approved the study and written informed consent was obtained from each participant.

Leiden studies consists of the Genetics ARthrosis and Progression study (GARP) and the Leiden Longevity Study (LSS). **GARP**. The Genetics ARthrosis and Progression study (GARP) study is aimed at identifying (genetic) determinants of osteoarthritis susceptibility and progression. The study consists of sibling pairs with symptomatic OA at multiple joint sites, as outlined in detail previously⁴. The **Leiden Longevity Study** includes independent Caucasian offspring and partners of long-lived siblings that were alive and over 89 years old for males and over 91 years for females as outlined in detail previously⁵. For both studies hand OA scores were determined by experienced readers of the Depts. Radiology and Rheumatology of the LUMC as previously outlined in detail.⁴ Written informed consent was obtained from each subject as approved by the ethical committees of the Leiden

University Medical Center. Genotyping was performed using Illumina Infinium HD Human660W-Quad Beadchips (Illumina, San Diego, CA, USA) and imputation of the GARP and LLS studies was performed simultaneously with 1000G V.3 reference panel using IMPUTE (Table S3).

TwinsUK. The TwinsUK study participants were white monozygotic and dizygotic twin pairs from the TwinsUK adult twin registry, a group used to study the heritability and genetics of age-related diseases⁶. These unselected twins were recruited from the general population through national media campaigns in the United Kingdom. All samples from the TwinsUK cohort for this study were genotyped with the HumanHap610Q (Illumina). The following quality control filters were applied: call rate $\geq 98\%$, MAF $\geq 1\%$, P for Hardy-Weinberg equilibrium $\geq 1 \times 10^{-6}$ (Table S3). The total number of genotyped SNPs that passed these filters was 303,940 SNPs. Imputation was done with reference to 1000G V.3 reference panel using the IMPUTE software package. The St. Thomas' Research Ethics Committee approved the study (EC96/439 TwinsUK), and all participants provided informed written consent

Framingham Heart Study. The original Framingham Study was a population-based sample of adults (ages 28-61 years) that began in 1948⁷. The Framingham Offspring Study is composed of children of the original Framingham Heart Study participants, and the children's spouses⁸. As part of an ancillary study in 1992 to 1995, Offspring (and their spouses) were contacted by mail and telephone call to participate in a visit to assess hand OA. About 1,800 individuals (ages 28-82 years) were examined, representing about 65% of those contacted. Of these individuals, 1,293 participants returned for another hand examination in 2002 to 2005⁹. Individuals underwent bilateral posteranterior hand radiographs, which were read by a trained musculoskeletal radiologist. The bilateral 2nd-5th distal interphalangeal joints (DIPJ), 2nd-5th proximal interphalangeal joints (PIPJ), 1st-5th metacarpophalangeal joints (MCPJ), thumb interphalangeal joint, thumb base (carpometacarpal) joint and wrist joints were scored according to Kellgren-Lawrence (KL) scale¹⁰ with good inter-reader reliability (weighted $\kappa=0.76$).

Genotyping was conducted on the Affymetrix GeneChip Human Mapping 500K Array and the 50K Human Gene Focused Panel. Genotypes were called using the BRLMM algorithm. A total of 549,781 SNPs were genotyped. The following quality-control filters were applied: 1) call rate $< 97\%$, 2), Hardy-Weinberg Equilibrium p-value $< 1 \times 10^{-6}$, 3) minor allele frequency $< 1\%$, and 4) Mendelian errors > 1000 . Additionally, the following filters

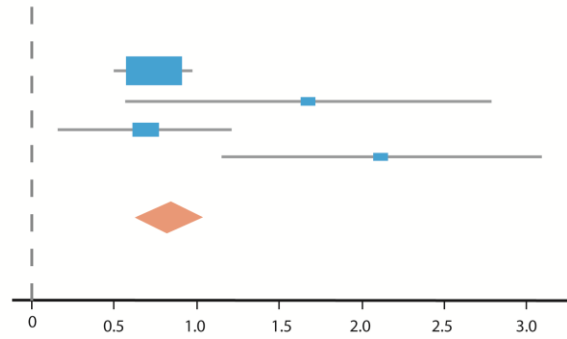
were also applied, 1) SNPs that do not map to NCBI Build 37, 2) duplicate SNPs, and 3) SNPs not on chromosomes 1-22 or X. A total of 137,728 genotyped SNPs were removed, leaving 412,053 SNPs as input to MACH for phasing (Table S3). SNPs were imputed with minimac (<http://genome.sph.umich.edu/wiki/minimac>) using reference haplotypes from 1000 Genomes V.3.

Supplementary Figures

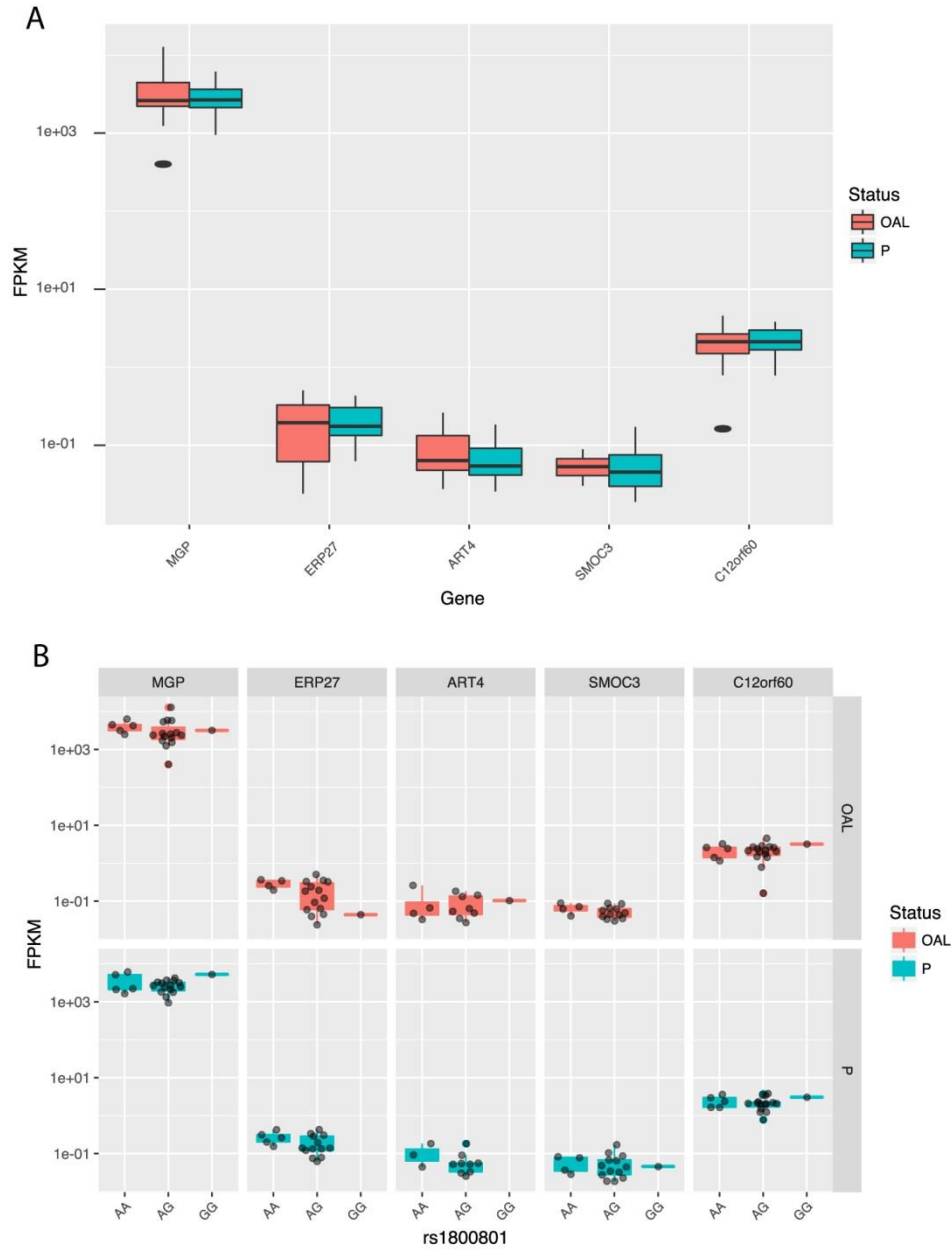
Forrest Plot *rs4764133*

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

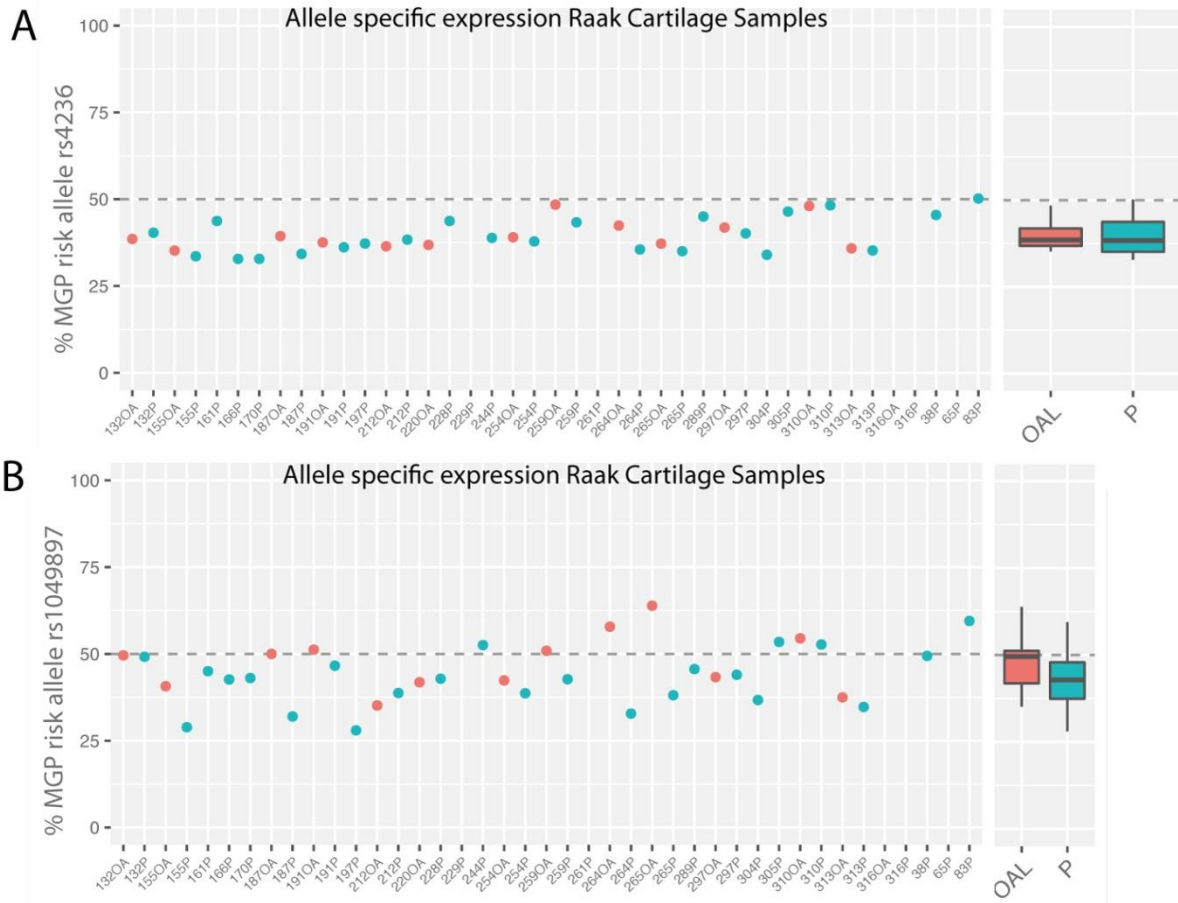
Heterogeneity: $\text{Chi}^2=63.27$, $\text{df} = 3$ ($P = 0.02$), $I^2=69.07$



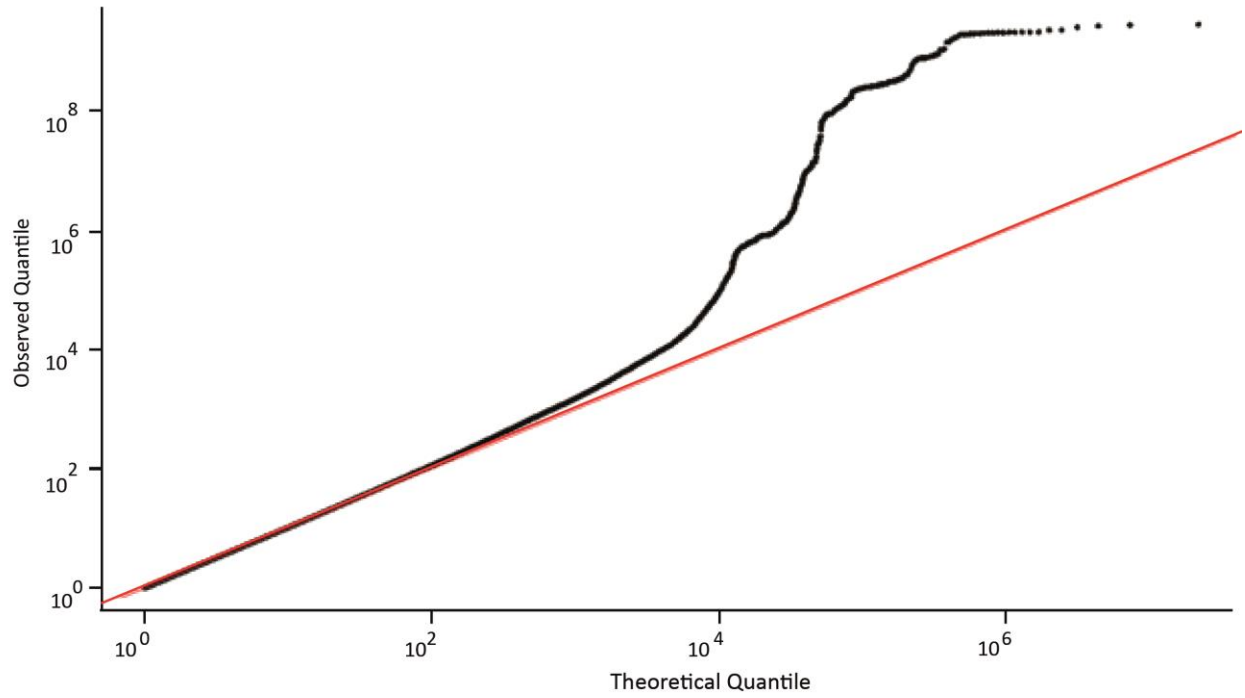
Supplementary Figure S1. Forrest plot for KLsum-score GWAS tophit *rs4764133*.



Supplementary Figure S2. Expression *MGP*, *ERP27*, *C12orf60*, *ART4* and *C12orf69* in articular cartilage (**A**), stratified by preserved (blue) and OA lesioned (red) samples. No significant difference was observed between paired preserved and lesioned cartilage. (**B**) eQTL results for rs1800801 with *MGP*, *ERP27*, *C12orf60*, *ART4* and *C12orf69*, we observed no eQTL effects.



Supplementary Figure S3: Allelic imbalanced expression of *MGP* marked by the alleles among heterozygotes of rs4236 (A) and (B) rs1049897, in the assessed cartilage RNA sequencing dataset.



Supplementary Figure S4: QQplot of GWAS P-values of the discovery cohorts, Rotterdam Study I, Rotterdam Study II and Rotterdam Study III.

Supplementary Tables

Supplementary Table S1

See Excel file Supplementary Table S1

Supplementary Table S2 General characteristics of RAAK cartilage samples

RAAK sample	AGE	SEX	JOINT
132 OA	67	F	Hip
132 P	67	F	Hip
155 OA	61	M	Knee
155 P	61	M	Knee
161 P	67	F	Hip
166 P	79	F	Hip
170 P	49	F	Knee
174 P	69	F	Hip
187 OA	59	F	Knee
187 P	59	F	Knee
191 OA	75	F	Knee
191 P	75	F	Knee
197 P	67	M	Hip
212 OA	69	M	Knee
212 P	69	M	Knee
220 OA	77	F	Knee
228 P	NA	F	Knee
229 P	66	F	Knee
244 P	62	F	Knee
254 OA	79	F	Knee
254 P	79	F	Knee
259 OA	48	F	Hip
259 P	48	F	Hip
261 P	74	F	Knee
264 OA	78	F	Hip
264 P	78	F	Hip
265 OA	79	F	Knee
265 P	79	F	Knee
289 P	69	F	Knee
297 OA	64	F	Knee
297 P	64	F	Knee
304 P	66	M	Hip
305 P	67	M	Hip
310 OA	76	F	Knee
310 P	76	F	Knee
313 OA	67	F	Knee
313 P	67	F	Knee
316 OA	76	M	Knee
316 P	76	M	Knee
32 P	55	F	Hip
38 P	60	M	Hip
65 P	61	F	Hip
83 P	62	F	Hip

P = preserved OA, OA = lesioned/OA affected Cartilage

Table S3. Baseline characteristics of the studies included in the analyses of osteoarthritis of the hand

Cohort Studies	Characteristics*			
	Female (%)	Age (year)	KLsum score	N
RS-I	58%	67.7(7.9)	8.4 (9.8)	4877
RS-II	54%	64.6 (7.9)	6.9 (8.4)	1793
RS-III	57%	56.8 (7.0)	4.6 (6.5)	2073
Twins-UK	100%	54.1 (7.8)	4.6 (7.9)	2000
Leiden Hand OA studies	65%	59.9 (7.3)	11.9(11.7)	836
Framingham Heart Study	56%	65.2 (9.1)	8.9 (12.9)	1175

*Values are expressed as means with standard deviations (SD) or percentages %

Supplementary Table S4. Association of rs4764133 with Hip OA (TREATOA meta-analysis), Knee OA (TREATOA meta-analysis) and bilateral finger OA (y/n finger OA)

Phenotype	SNP	A1	A2	OR	P-value	Case	Control	Cohorts
Finger OA **	rs4764133	T	C	1.25	3.09E-08	2371	5751	RS1, RS2, RS3
Hand OA ***	rs4764133	T	C	1.36	2.80E-02	140	5,801	RS1, RS2, RS3
Hip OA	rs4764133	T	C	0.97	4.93E-01	3,595	6,559	Treat-OA†
Knee OA	rs4764133	T	C	0.96	2.99E-01	5,013	9,132	Treat-OA†
mJSW ⁺	rs1049897	T	A	-0.398	1.28E-02	<i>n</i> = 21,240		5-cohorts ¹²

** fingerOA definition : KL score > 2 in minimal one DIP or PIP joint in both left and right hand

*** Severe handOA definition same as in Styrkarsdottir et al. (2014) Nature Genetics¹¹.

⁺ Minimal Joint Space Width(mJSW), no OR, but Beta Value, quantitative phenotype¹²

‡ Treat-OA consortium¹³

Supplementary Table S5. Variants in the ALDH1A2 gene and their association with the Klscore measure in meta-analysis of discovery cohorts (RS-I, RS-II and RS-III).

SNP	Allele1	Allele2	Freq1	FreqSE	MinFreq	MaxFreq	Effect	StdErr	P-value	Direction	HetISq	HetChiSq	HetDf	HetPVal
rs12907038*	C	G	0.553	0.005	0.548	0.560	-0.434	0.119	2.68E-04	---	0	0.154	2	0.926
rs4238326*	T	C	0.644	0.004	0.641	0.651	-0.475	0.123	1.10E-04	---	0.1	2003.000	2	0.367
rs3204689*	C	G	0.453	0.005	0.447	0.458	0.422	0.118	3.69E-04	+++	0	0.323	2	0.851

* Variants associated with Hand OA by Stykarsdottir et al. (2014) *Nature Genetics*¹¹.

References

1. Hofman, A. *et al.* The Rotterdam Study: 2016 objectives and design update. *Eur. J. Epidemiol.* **30**, 661–708 (2015).
2. Estrada, K. *et al.* GRIMP: A web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics* **25**, 2750–2752 (2009).
3. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
4. 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**, 438–43 (2005).
5. Schoenmaker, M. *et al.* Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. *Eur. J. Hum. Genet.* (2005). doi:10.1038/sj.ejhg.5201508
6. Moayyeri, A., Hammond, C. J., Valdes, A. M. & Spector, T. D. Cohort Profile: TwinsUK and healthy ageing twin study. *Int. J. Epidemiol.* **42**, 76–85 (2013).
7. Dawber, T. R., Meadors, G. F. & Moore, F. E. Epidemiological approaches to heart disease: the Framingham Study. *Am. J. Public Health* **41**, 279–281 (1951).
8. Kannel, W. B., Feinleib, M., McNamara, P. M., Garrison, R. J. & Castelli, W. P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am. J. Epidemiol.* **110**, 281–90 (1979).
9. Haugen, I. K. *et al.* Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. *Ann. Rheum. Dis.* **70**, 1581–1586 (2011).
10. Kellgren, J. H. & Lawrence, J. S. Radiological Assessment of Osteo-Arthrosis. *Ann. Rheum. Dis.* **16**, 494–502 (1957).
11. 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).
12. Castaño-Betancourt, M. C. *et al.* Novel Genetic Variants for Cartilage Thickness and Hip Osteoarthritis. *PLOS Genet.* **12**, e1006260 (2016).
13. Kerkhof, H. J. M. *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.* **19**, 265–71 (2011).