OPEN ACCESS

EPIDEMIOLOGICAL SCIENCE

Meta-analysis of erosive hand osteoarthritis identifies four common variants that associate with relatively large effect

Unnur Styrkarsdottir,[1] Lilja Stefansdottir,[2] Gudmar Thorleifsson,[2] Olafur A Stefansson,[1] Saedis Saevarsdottir,[1,3] Sigrun H Lund,[2] Thorunn Rafnar,[1] Kazuyuki Hoshijima,[4] Kendra Novak,[5] Natividad Oreiro,[6] Ignacio Rego-Perez,[6] Channing Hansen,[7] Nikolas Kazmers,[5] Lambertus A Kiemeney,[8] Francisco J Blanco,[6,9] Tyler Barker,[5,10,11] Margreet Kloppenburg,[12] Michael J Jurynec,[4] Daniel F Gudbjartsson,[2,13] Helgi Jonsson,[3,14] Unnur Thorsteinsdottir,[1,3] Kari Stefansson[3,15]

Handling editor Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/ard-2022-223468).

For numbered affiliations see end of article.

Correspondence to
Dr Unnur Styrkarsdottir, Population Genomics, deCODE genetics / Amgen Inc, Reykjavik, Iceland; unnur.styrkarsdottir@decode.is and Kari Stefansson, deCODE genetics, Reykjavik, Iceland; kstefans@decode.is

Received 11 October 2022
Accepted 25 February 2023

Check for updates

To cite: Styrkarsdottir U, Stefansdottir L, Thorleifsson G, *et al*. *Ann Rheum Dis* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/ard-2022-223468

ABSTRACT

Objectives Erosive hand osteoarthritis (EHOA) is a severe subset of hand osteoarthritis (OA). It is unclear if EHOA is genetically different from other forms of OA. Sequence variants at ten loci have been associated with hand OA but none with EHOA.

Methods We performed meta-analysis of EHOA in 1484 cases and 550 680 controls, from 5 populations. To identify causal genes, we performed eQTL and plasma pQTL analyses, and developed one zebrafish mutant. We analysed associations of variants with other traits and estimated shared genetics between EHOA and other traits.

Results Four common sequence variants associated with EHOA, all with relatively high effect. Rs17013495 (*SPP1*/*MEPE*, OR=1.40, $p=8.4\times10^{-14}$) and rs11243284 (6p24.3, OR=1.35, $p=4.2\times10^{-11}$) have not been associated with OA, whereas rs11631127 (*ALDH1A2*, OR=1.46, $p=7.1\times10^{-18}$), and rs1800801 (*MGP*, OR=1.37, $p=3.6\times10^{-13}$) have previously been associated with hand OA. The association of rs1800801 (*MGP*) was consistent with a recessive mode of inheritance in contrast to its additive association with hand OA (OR homozygotes vs non-carriers=2.01, 95% CI 1.71 to 2.37). All four variants associated nominally with finger OA, although with substantially lower effect. We found shared genetic components between EHOA and other OA measures, grip strength, urate levels and gout, but not rheumatoid arthritis. We identified *ALDH1A2*, *MGP* and *BMP6* as causal genes for EHOA, with loss-of-function *Bmp6* zebrafish mutants displaying EHOA-like phenotypes.

Conclusions We report on significant genetic associations with EHOA. The results support the view of EHOA as a form of severe hand OA and partly separate it from OA in larger joints.

WHAT IS ALREADY KNOWN ON THIS TOPIC

No genetic associations have been reported for erosive hand osteoarthritis (EHOA).

WHAT THIS STUDY ADDS

This study finds the first genetic association with EHOA at four loci that all confer relatively high risk of the disease, identifies candidate causal genes at three loci: *ALDH1A2*, *MGP* and *BMP6*, and strong candidates at one locus: *SPP1*, *IBSP* and *MEPE*.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

This study highlights EHOA as somewhat separate from osteoarthritis in the larger joints and points to potential drug targets for the disease.

INTRODUCTION

Erosive hand osteoarthritis (EHOA) is a severe form of hand osteoarthritis (OA), one of the most prevalent forms of OA.[1–3] The clinical burden of EHOA is higher than for other types of hand OA (nodal hand OA or OA in the thumb base). It is characterised by abrupt onset with inflammation, radiographic features of central erosions and collapse of the subchondral bone, and rapid progression. Markers of inflammation and bone resorption are higher in EHOA patients than in other forms of hand OA. This can make it challenging to differentiate clinically from erosive rheumatoid arthritis (RA) and erosive gout in the small joints of the hand, two disease entities that have specific effective therapies on the market, while no disease-modifying drugs are yet available for EHOA. Between 5% and 20% of patients with symptomatic hand OA have EHOA which, as other OA types, predominantly affects females (reviewed in ref 3).[3]

Although EHOA is phenotypically different from nodal hand OA in the distal and proximal interphalangeal joints, it is not clear if EHOA represents a genetically distinct form of hand OA. Several studies have identified a genetic or familial component to EHOA[4 5] and a few candidate genes and loci, such as HLA alleles and the *IL1B* gene, have been suggested.[6–8]

There is, however, no genome-wide association study (GWAS) of EHOA that has been reported, but ten loci have been described for hand/finger/thumb

OA.[9–12] The first and only meta-analysis of hand OA, which included 20 901 individuals with hand OA from 9 populations,[12] found associations at the previously reported *ALDH1A2*,[9] *MGP*[10 11] and *WNT9A*[11] loci, as well as at seven additional loci. None of the earlier studies separated EHOA from finger or hand OA, that is, EHOA patients were included in these analyses.

Here, based on five independent EHOA study populations, we identified four genetic loci that associate with EHOA. Two of these loci were previously associated with hand OA overall, at *ALDH1A2* and *MGP*. We also discovered two new loci with candidate causal genes involved in bone biology, *BMP6* and *SPP1/MEPE/IBSP*. Our data indicate that EHOA has substantial genetic overlap with finger OA, yet displays risk alleles that are associated with susceptibility of EHOA over that of finger or hand OA and of OA in other joints.

METHODS

Details on the study populations and the methods used are given in online supplemental material to this publication.

Study populations

Iceland: EHOA (918 cases) was diagnosed from conventional dorsopalmar radiographs taken of individuals with provisional diagnosis of hand OA and compared with 109 249 controls. The proximal and distal interphalangeal joints were scored according to Verbruggen-Veys (VV)[13] and patients with at least one joint in the E phase (erosive) or R phase (remodelled) were classified as having EHOA. Individuals diagnosed with RA were excluded.

The Netherlands: The EHOA cases (N=145) were derived from the Hand OSTeoArthritis in Secondary care study,[14] and the controls (N=5102) from the Nijmegen Biomedical Study.[15] EHOA cases were classified according to VV,[13] excluding RA.

UK: The UK Biobank resource (http://www.ukbiobank.ac.uk) includes data from 500 000 volunteers who were recruited between the age of 40 and 69 years in 2006–2010 across the United Kingdom. EHOA included those with the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD10) code M15.4, excluding RA (63 EHOA cases/430 875 controls).

USA: The EHOA cases (N=145) included those with ICD-10 code M15.4, excluding RA, in the Utah Population Database[16] and the Intermountain Healthcare HerediGene: Population Study (Utah, USA), compared with 5308 controls.

Spain: The EHOA cases (N=218) were derived from the PROspective COhort of A Coruña cohort,[5 17] and the controls (N=164) were from other projects at A Coruña University Hospital who had not been diagnosed with hand OA on radiographs. EHOA cases were scored according to VV.[13]

All participants in this study were genetically determined to be of European descent.

Genotyping and association analysis

All the samples, except UK Biobank, were genotyped at deCODE genetics, using various Illumina chips, while UK Biobank genotyping used a custom-made Affimetrix chip. Imputation of all datasets was performed at deCODE genetics. Association analysis was done using logistic regression, adjusting for age, sex and principal components.

EHOA meta-analysis

We meta-analysed GWAS summary results from the additive model using a fixed-effects inverse variance method,[18] including variants with info >0.8 and present in at least two datasets (N=46 million). For GWS thresholds we used the weighted Holm-Bonferroni method to allocate familywise error rate of 0.05 equally between five annotation-based classes of sequence variants.[19] For the EHOA associated variants we also tested the recessive model, and the full genotype model.

Polygenic Risk Score and phenotype correlation analysis

We used Polygenic Risk Score (PRS) analysis based on a EHOA meta-analysis of Icelandic, Dutch, Spanish and US GWASs to investigate its correlation with about 5000 quantitative and case/control traits in the UK Biobank dataset. The PRSs was calculated using genotypes for about 600 000 autosomal markers included on the Illumina SNP chips to avoid uncertainty due to imputation quality.[20]

Genetic correlations

Using cross-trait linkage disequilibrium (LD) score regression method,[21] we estimated the genetic correlation between EHOA and other OA subtypes in the Genetics of Osteoarthritis (GO) consortium dataset,[12] and with other traits identified as correlated with EHOA in the PRS analysis in data from UK biobank, or associated with the EHOA variants, and RA (see online supplemental material for description of these phenotypes). In this analysis, we used results for about 1.2 million well-imputed variants, and for LD information, we used precomputed LD scores for European populations (downloaded from: https://data.broadinstitute.org/alkesgroup/LDSCORE/eur_w_ld_chr.tar.bz2.

Phenoscan of public datasets

Associations of EHOA variants with other phenotypes was assessed using the Open Targets Genetics website (https://genetics.opentargets.org/), and a diverse set of phenotypes in UK Biobank that were generated at deCODE genetics. Associations with the lead EHOA variants, and variants in LD with the EHOA variants ($r^2>0.8$), and $p<1\times10^{-6}$ were evaluated.

Functional annotation of sequence variants and enrichment of association signals

We determined if the lead sequence variant or correlated variants ($r^2>0.80$) were located within candidate cis-regulatory elements (cCRE)[22] or tissue-specific regulatory regions[23] and looked for association signals in enhancer elements defined in EpiMap. We also determined their location within tissue-specific regulatory regions.[23]

Co-localisation of GWA signals with expression quantitative trait loci (eQTL) and protein quantitative trait loci (pQTL) signals

We analysed co-localisation of the EHOA associations with variation in gene transcription (eQTL) or variations in protein levels in plasma (plasma pQTL).[24] For the eQTLs analysis, we used data from the publicly available Genotype-Tissue Expression (GTEx) project (https://www.gtexportal.org/), and deCODE genetics RNA sequence data from whole blood of 13 175 Icelanders and subcutaneous adipose tissue from 700 Icelanders.[25] For plasma pQTL analysis, we used the dataset described in Ferkingstad *et al*,[26] which tested association of 27.2 million variants with levels of 4719 proteins (adjusted and standardised levels) in plasma samples from 35 559 Icelanders.

Plasma protein levels

The dataset used for analysis of plasma protein levels is the same as for the plasma proteomics, restricted to those EHOA patients who had their sample taken within a year (±1 year)

Table 1 Characteristics of the study subjects

		N (% female)	Age, mean (±SD)	BMI, mean (±SD)
Iceland	EHOA	918 (79)	75.0 (11.2)	27.3 (4.9)
	Controls	109 249 (46)	66.5 (14.0)	26.8 (5.3)
UK Biobank	EHOA	63 (79)	61.3 (6.6)	28.6 (6.4)
	Controls	430 875 (54)	57.4 (8.0)	27.4 (4.8)
USA	EHOA	145 (82)	68.9 (12.1)	27.5 (6.1)
	Controls	5308 (60)	56.3 (18.2)	29.6 (6.9)
Spain	EHOA	218 (84)	61.1 (8.7)	28.1 (5.3)
	Controls	164 (32)	58.9 (12.6)	27.5 (4.6)
The Netherlands	EHOA	139 (82)	64.3 (8.4)	27.5 (4.7)
	Controls	5102 (53)	54.9 (18.2)	25.2 (4.0)

BMI, body mass index; EHOA, erosive hand osteoarthritis.

from the radiograph that was used to diagnose EHOA. Association between protein levels and EHOA was tested with logistic regression (R V.3.6.3), adjusting for age, sex and body mass index. Results are represented as OR of having EHOA per SD increase in standardised plasma protein levels.

Zebrafish experiments

The zebrafish (*Danio rerio*) Tu strain was used in all experiments. The generation of F0 and germline zebrafish lacking *bmp6* gene function is described in detail in online supplemental material and shown schematically in online supplemental figure S1. Cartilage and bone staining was performed on 14 days post fertilisation (dpf) larvae.

Patient and public involvement statement

This research was done without direct patient involvement.

RESULTS

GWAS and meta-analysis

To search for sequence variants that contribute to EHOA, we performed GWAS in samples from Iceland, The Netherlands, Spain, UK and USA (table 1), and subsequently meta-analysed the results from 1484 subjects with EHOA and 550 680 controls.

We found four independent associations which satisfied our GWS criteria (table 2, online supplemental table S1, figure 1 and online supplemental material): rs17013495 (4q22.1, between *SPP1* and *MEPE*), rs11243284 (6p24.3), rs1800801 in 5'UTR of *MGP* (12p12.3) and rs11631127 (15q21.3, in *ALDH1A2*).

The associations at *MGP* and *ALDH1A2* have previously been reported for hand OA,[9–11] whereas rs17013495 (*SPP1/MEPE*) and rs11243284 (6p24.3) have not, nor with any other forms of OA. Rs11243284 at 6p24.3 is not correlated with the recently identified association of rs12190551 with spine OA (r^2=0.002).[27]

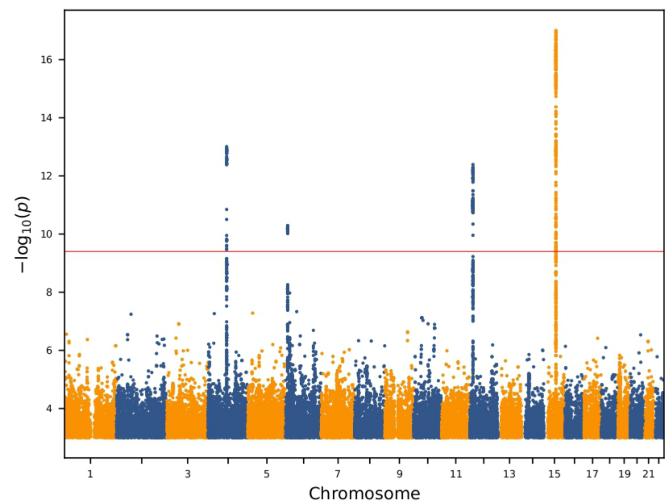


Figure 1 Manhattan plot of the genome-wide analysis of erosive hand osteoarthritis The p values (–log10) are plotted against their respective positions on each chromosome. Results are shown for all variants with significance level p<0.001 and imputation information greater than 0.8.

Rs1800801 in the 5' UTR in *MGP* associated stronger with EHOA under a recessive model (OR=1.85 (95% CI 1.59 to 2.14), p=3.7 × 10^{-16}), than under an additive/multiplicative model (OR=1.37 (1.26, 1.49), p=3.6 × 10^{-13}) (online supplemental table S2). In the full genotype model, which assesses risk of heterozygous and homozygous genotypes compared with the homozygous wild-type, the OR for the heterozygotes (TC) was smaller than expected for the additive model, OR_{het}=1.15 (1.00, 1.32), p=0.047, while the OR for the homozygotes (TT) was larger, OR_{hom}=2.01 (1.71, 2.37), p=1.1 × 10^{-16}. The full model fits significantly better than the additive model for rs1800801 (p=0.0011) (online supplemental table S2), demonstrating the recessive nature of this association. As opposed to the association of rs1800801 with EHOA, the association of rs1800801 with hand, finger and thumb OA was consistent with the additive model rather than the recessive model (online supplemental table S3).

For the other three EHOA-associated variants, we did not observe deviation from the additive/multiplicative model for the genotype risk (online supplemental table S2).

Functional annotation of the EHOA-associated variants

We annotated the EHOA variants according to location in ENCODE's encyclopaedia of cCRE,[22] their tissue specificity,[23] co-localisation with mRNA expression (eQTL) in various tissues and co-localisation with protein expression (pQTL) in plasma. We specifically note that bone, cartilage or other joint tissues are not available for eQTL/pQTL analysis in any public dataset.

Table 2 Genome wide significant associations with erosive hand osteoarthritis

Variant	Chr:position	EA/NEA	Freq%	Closest gene	VA	P value	OR
rs17013495	4:87 885 460	T/C	59.6	*SPP1/MEPE*	Intergenic	8.40E-14	1.40 (1.28, 1.53)
rs11243284	6:8 945 086	C/T	28.9		Intergenic	4.20E-11	1.35 (1.23, 1.48)
rs1800801	12:14 885 854	T/C	37.2	*MGP*	5'UTR	3.60E-13	1.37 (1.26, 1.49)
rs11631127	15:57 977 811	C/G	57.6	*ALDH1A2*	Intron	7.10E-18	1.46 (1.34, 1.59)

Results are shown from the meta-analysis of the Icelandic, Dutch, Spanish, UK and US sets. Results for individual sample sets are shown in online supplemental table S1. Chr is chromosome, Pos is the position in build GRCh38, EA designate the effect allele (EA) and NEA the other allele (non effect allele). Freq. is the allelic frequency of the effect allele. Gene refers to the nearest gene and VA is variant annotation. 5'UTR is the 5 prime untranslated region. P values are two sided and derived from a likelihood ratio test.

The EHOA-associated variants (the lead variant or highly correlated variants, $r^2>0.8$) at all four loci reside in enhancer-like sequences (online supplemental table S4), and the variants at *MGP* and *ALDH1A2* also overlap with promoter-like sequences, suggesting a regulatory role of these variants in expression of nearby genes. The 12p12.3 (*MGP*), 15q21.3 (*ALDH1A2*) and 4q22.1 (*SPP1*/*MEPE*) signals are in cCREs found in many different tissues, whereas the 6p24.3 signal is restricted to few tissue types (online supplemental table S5), possibly suggesting tissue specific activity. Consistent with this observation we found co-localisation of the EHOA variants and/or mRNA expression or protein levels in plasma, at three of the loci: *SPP1* at 4q22.1, *MGP* at 12p12.3 and *ALDH1A2* at 15q21.3 (online supplemental table S6 and S7). *MGP* and *ALDH1A2* are also predicted target genes in the EpiMap resource[28] (online supplemental table S8). Furthermore, all of the four EHOA loci are within tissue-specific regulatory regions for vascular/endothelial cells which we estimate is 2.8-fold higher than expected by chance alone (expected overlap=35%; 95% CI 0% to 75%), but, as we tested for enrichment within 16 different tissue-specific groups,[23] the enrichment was only nominally significant (p=0.011, online supplemental table S9).

The EHOA risk allele of rs11631127 co-localised with reduced expression of *ALDH1A2* in cultured fibroblasts (online supplemental table S6), consistent with previous results in cartilage and other joint tissues,[9 29] and rs1800801[T] in *MGP* co-localised with both reduced *MGP* eQTL (online supplemental table S6) in several tissues and with reduced matrix Gla protein (encoded by the *MGP* gene) pQTL in plasma (online supplemental table S7), also consistent with previous results.[10 12 28 30] Since the *MGP* gene is expressed at a very low level in blood cells the protein in plasma primarily comes from other tissues. Furthermore, in our data, an increased plasma level of matrix Gla protein associated with lower odds of EHOA (OR=0.75 per SD, p=0.028, $N_{erosive}=55$, $N_{controls}=27\,083$, online supplemental figure S2).

Rs17013495[T] at the 4q22.1 locus co-localised with reduced mRNA expression of the *SPP1* gene in spleen (online supplemental table S6), and associated with decreased level of osteopontin (encoded by the *SPP1* gene) in plasma (online supplemental table S7), although not the strongest cis-pQTL for this protein in plasma. Increased levels of bone sialoprotein 2, encoded by the *IBSP* gene at the 4q22.1 locus, associated with reduced odds of EOHA (OR=0.74 per SD, p=0.023, online supplemental figure S2), although pQTL or eQTL for this gene did not co-localise with the EHOA variants. However, we note that expression of the *IBSP* gene is mostly restricted to bone and cartilage, tissues without public eQTL/pQTL datasets.

We did not detect eQTLs or pQTLs at the 6p24.3 locus. However, of the nine genes within 1.5 MB of rs11243284, *BMP6* is the most likely candidate gene because of the known role of the BMP signalling pathway in skeletal formation and homeostasis.[31–33] To uncover biological functions of BMP6 in vivo, we examined the consequences of complete loss of *bmp6* function in the zebrafish. We used CRISPR-Cas9 methods to generate F0 and germline deletions of *bmp6* (online supplemental figure S1). WT and *bmp6*$^{+/-}$ have a normally segmented vertebral column indicating that Bmp6 does not affect the overall development or patterning of the larval skeleton (figure 2 and online supplemental figure S3). In contrast to WT or control larvae, *bmp6*$^{+/-}$ have multiple defects reminiscent of EHOA, including bone erosions, structural defects in the vertebral precursors and ectopic cartilage formation. These data support that *BMP6* is a strong candidate gene in EHOA.

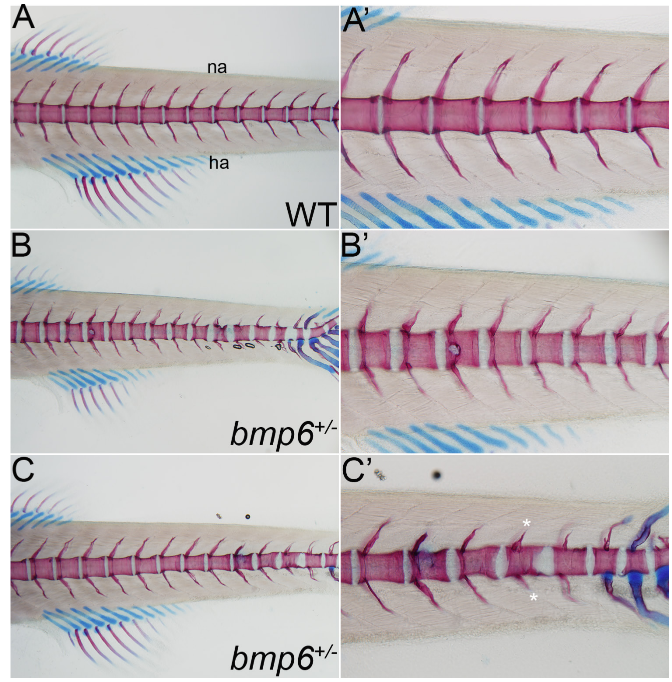


Figure 2 Loss of *bmp6* causes erosive-like phenotypes in the zebrafish vertebral precursors. (A–C′). Analysis of cartilage (blue) and bone (red) in the vertebral column of 14 days post fertilisation wild-type (WT) and *bmp6*$^{+/-}$ zebrafish larvae. (A, A′) WT larvae have a normally segmented and ossified centra (vertebral precursors) and neural (na) and hemal arches (ha), whereas (B, C′) *bmp6*$^{+/-}$ have multiple defects, including bone erosions (arrow in B and B′), structural defects in the centra (arrowhead in B, C and C′), ectopic cartilage formation (arrow in C′), and disruptions in the neural and hemal arches (asterisks in C′). No defects are observed in the cartilaginous structures of the fins. All images are lateral views with anterior to the left.

All the above genes (*ALDH1A2*, *MGP*, *BMP6*, *SPP1* and *IBSP*) are expressed in human cartilage,[34] with relative expression from the 0.01st percentile (*MGP*) to the 12th percentile (*BMP6*).

Association of EHOA variants with other OA subtypes and relevant diseases or traits

To address association of the four EHOA variants with other OA subtypes and other diseases or traits, we used data from the GO consortium[12] and public datasets (Open Targets Genetics and UK Biobank data). Furthermore, we generated EHOA PRS to run a non-hypothesis driven scan for genetic overlap with other diseases/traits in UK Biobank, and subsequently, assessed the genetic component shared by EHOA and other traits with LD score regression.

All four EHOA variants associated with finger OA in the GO consortium data ($P_{Bonferroni}<0.0025$) but with considerably lower OR estimate than for EHOA (table 3). All EHOA variants, except rs11243284 at 6p24.3, also associated nominally with thumb OA. Of special note is the opposite effect of rs1800801 (*MGP*) and rs11631127 (*ALDH1A2*) on knee OA compared with EHOA, that is, the EHOA risk alleles associated with reduced risk of knee OA, consistent with what was also observed in the GO consortium meta-analysis.[12] None of the EHOA variants associated with spine OA.

Three of the EHOA signals, rs17013495 (*SPP1*), rs1800801 (*MGP*) and rs11631127 (*ALDH1A2*), showed some multitrait associations, although mostly with musculoskeletal measures; hand grip strength and bone density (online supplemental table

Table 3 Association of the four EHOA variants with other osteoarthritis in the GO consortium meta-analysis

		Finger OA (N=10 804 cases/255 814 controls)		Thumb OA (N=10 536 cases/236 919 controls)		Hip OA (N=36 520 cases/317 590 controls)		Knee OA (N=63 498 cases/335 777 controls)		Spine OA (N=28 731 cases/307 798 controls)	
Variant (allele)	**Chr:position**	**OR**	**P value**	**OR**	**P value**	**OR**	**P value**	**OR**	**P value**	**OR**	**P value**
rs17013495(T)	chr4:87 885 460	1.08	2.3E-05*	1.05	7.9E-03*	0.99	0.16	1.00	0.58	1.01	0.32
rs11243284(C)	chr6:8 945 086	1.10	1.3E-06*	1.00	0.85	1.00	0.84	1.00	0.59	0.99	0.25
rs1800801(T)	chr12:14 885 854	1.16	8.6E-16*	1.06	2.5E-04*	0.97	5.5E-03	0.98	2.3E-03*	1.01	0.27
rs11631127(C)	chr15:57 977 811	1.09	3.7E-07*	1.10	1.3E-08*	1.02	0.079	0.97	1.3E-06*	1.00	0.64

Results are shown for OA subsets phenotypes in the Genetics of Osteoarthritis Consortium meta-analysis.[12] Chr is chromosome, Pos is the position in build GRCh38.
*Denotes significant associations after correction for multiple testing.
EHOA, erosive hand osteoarthritis; GO, genetics of osteoarthritis; OA, osteoarthritis.

S10). No disease or trait, except for OA, was shared by two or more of the EHOA loci. Of note is association of rs17013495[C] with increased levels of urate and risk of gout, another form of arthritis caused by uric acid crystal deposition, but severe gout can also result in bone erosions. Follow-up of these observations for all four EHOA variants in UK Biobank data and in our meta-analysis of bone density shows an association of all four EHOA risk alleles with reduced grip strength (online supplemental figure S4), but only rs17013495 (*SPP1/MEPE*) associated with urate (online supplemental table S11). All four EHOA variants also associated nominally with lumbar spine bone mineral density (LS-BMD), but the direction of effects was not consistent between the four variants. Only rs1800801 (*MGP*) associated with BMD estimated with heel ultrasound (eBMD).

Consistent with the above-described observations, the EHOA PRS scan was only significant ($p<1.0 \times 10^{-5}$, accounting for 5000 main phenotypes) for hand OA measures, other arthrosis diagnosis (ICD10:M19), polyarthrosis (ICD10:M15), pain due to OA and hand grip strength (online supplemental table S12).

We estimated the extent of shared genetics between EHOA and the other OA subtypes and the traits identified in the phenoscans through genetic correlation analysis. Although, not identified in the multitrait associations analysis nor in the PRS scan, we also included RA in this analysis because that it is another form of inflammatory arthritis that can result in bone erosions and is, as gout, a clinical differential diagnosis to EHOA.

We observed highest genetic correlation between EHOA and those types of OA of which EHOA is a subset, that is, finger OA and hand OA, followed by thumb OA, knee OA, hip OA and the weakest with spine OA (figure 3). Reduced grip strength, increased urate levels and gout were also nominally correlated genetically with EHOA, whereas measures of bone density and RA were not.

The extent of genetic correlation between EHOA and other OA types is also reflected by the associations of GO consortium variants with EHOA[12] (online supplemental table S13). Eight of the 10 GO independent associations with hand, finger or thumb OA, associated with EHOA under a false discovery rate of 5% in our data, whereas only 3 of the remaining 68 independent knee, hip, spine or any OA variants did so. The small sample size of our EHOA dataset may not be powered to detect associations with these variants, however, similar results were also observed for direct comparison of the ORs of EHOA and the other OA subsets, irrespective of the significance of the association (online supplemental figure S5). We note that as for the EHOA variants reported here, a majority of the finger/hand OA variants associated with EHOA with larger ORs than with finger/hand OA in the GO data, indicating that EHOA is a severe subset of finger OA.

DISCUSSION

Here, we describe the first GWAS of EHOA. Despite a modest sample size of 1484 cases, we found 4 significant EHOA loci, all of which confer relatively high effect on EHOA risk.

Two of the associated loci, rs1800801 (*MGP*) and rs11631127 (*ALDH1A2*), have previously been associated with hand OA overall.[9 10 12] Both of these loci also associated with knee OA with opposite effects to that of EHOA, that is, the EHOA risk alleles associate with protection of knee OA.[12] The EHOA risk alleles at these loci co-localise with lower mRNA expression of *ALDH1A2* and *MGP* in cartilage, other joint tissues as well as some other tissues,[9 10 12 29 30 35] and the rs1800801 (*MGP*) risk allele also co-localises with lower levels of matrix Gla protein levels in plasma, indicating that *ALDH1A2* and *MGP* genes are likely EHOA candidate causal genes at these loci.[9 10 12 29 30 35] We also show that the matrix Gla protein in plasma is lower in EHOA patients than in controls, further supporting a causal role

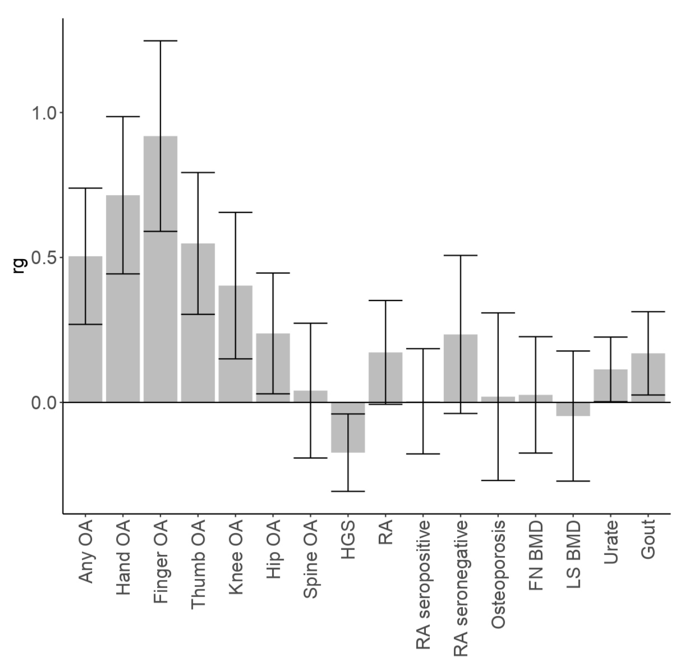


Figure 3 Genetic correlation between EHOA and other OA subtypes and diseases/traits The genetic correlation coefficient (r_g) and SE, of genetic correlation between EHOA and other OA subtypes, any OA (which includes all types of OA), and several diseases/traits are shown. HGS is hand grip strength, FN_BMD is femoral neck bone mineral density, LS_BMD is lumbar spine bone mineral density, and RA is rheumatoid arthritis. BMD, bone mineral density; EHOA, erosive hand osteoarthritis; LS-BMD, lumbar spine BMD; OA, osteoarthritis.

of the *MGP* gene in OA, with lower level of protein predisposing to the disease. rs1800801 (*MGP*) associated with EHOA under a recessive model, whereas the association with finger OA is consistent with an additive model. The matrix Gla protein is a vitamin K dependent inhibitor of ectopic tissue calcification, particularly of vascular and cartilage calcification.[36 37] The function of the protein depends on the post-translational Ca^{++} binding γ-carboxyglutamic acid residues (Gla), mediated by vitamin K, but fully carboxylated form of matrix Gla protein has been shown to be lower in OA cartilage than in normal cartilage.[38]

We found two association signals for EHOA that have not been associated with OA before, rs17013495 (*SPP1/MEPE*) and rs11243284 (*BMP6*). Both variants associated nominally with finger OA in our data, although with lower effect. The *SPP1/MEPE* locus is a well-known locus for BMD[39–41] and the EHOA risk variant also associated with increased LS-BMD in our data. We also observed association with increased levels of urate and risk of gout.

There are strong candidate genes at the *SPP1/MEPE* locus, that harbours a cluster of five genes that encode the SIBLNG (small integrin-binding ligand N-linked glycoprotein) family of extracellular matrix proteins, three of which are expressed in the relevant tissues of bone and/or cartilage: *IBSP* (bone sialoprotein 2), *MEPE* (matrix extracellular phosphoglycoprotein) and *SPP1* (osteopontin). *SPP1* is expressed in many tissues and cell types whereas expression of *IBSP* and *MEPE* is mostly restricted to bone, cartilage and teeth.[42–44] We found co-localisation of rs17013495 EHOA risk variant and lower expression of *SPP1* in spleen, and association with a secondary pQTL for plasma levels of osteopontin. We also observed lower levels of bone sialoprotein 2 in plasma of EHOA patients. The origin of this protein in plasma is most likely from bone as it constitutes approximately 12% of the non-collagenous proteins in human bone and is not expressed in other tissues than bone and/or cartilage. However, since no dataset is currently available to conduct well-powered eQTL or pQTL studies in joint tissues, the possible causal effect of these genes on EHOA cannot be differentiated at this stage. They all play key biological roles in the mineralisation of bone, form an integral part of the mineralised matrix and are involved in chondrocyte differentiation, bone formation and remodelling.[45]

The similarities in the bone phenotypes that we observed in the zebrafish *bmp6* mutants we created with the clinical hallmarks of EHOA suggests that *BMP6*, that has a role in maintaining bone and joint homeostasis, is the candidate causative gene at the 6p24.3 locus. Although several studies have examined the function of *BMP6* on bone formation, its precise role remains unclear possibly due to functional redundancy of other BMPs or genetic compensation.[32 46–48] Recently, a GWAS found that an intronic variant in *BMP6*, rs12190551[C], uncorrelated with the EHOA signal, associated with spine OA. The spine OA risk allele correlated with reduced expression of *BMP6* mRNA in the tibial nerve in the GTEx portal.[27] Previous transcriptomic analysis of musculoskeletal tissue from *bmp* mutants has demonstrated that loss of *bmp6* activated the NF-κB pathway, which inhibited development of osteoblasts and promoted osteoclast formation.[46] Further, gain-of-function and loss-of-function studies in animal models are needed to delineate the precise role and mechanism of *BMP6* function in OA.[49]

Our phenoscan of over 5000 different diseases and traits in UK Biobank using the EHOA PRS, as well as genetic correlation analysis using LD score regression, indicated that EHOA unsurprisingly shares genetics with different measures of OA, but also with decreased hand grip strength, increased urate concentrations and gout, but not RA. The genetic correlation with other OA subtypes shows, as expected, the most shared genetics between EHOA and finger and hand OA, of which EHOA is a subset. EHOA, gout and RA share the clinical features of joint inflammation, and erosions in the most severe cases of gout and RA. It should also be noted that it can be difficult to differentiate between EHOA and gout, both clinically and radiographically. In contrast to EHOA, there are several effective disease-modifying antirheumatic therapies available for RA that hinder progression to erosive disease, but those have not proven effective against EHOA, also indicating a different underlying pathogenesis.

Here, we describe the first robust loci for EHOA. All four loci conferred relatively high risk of the disease, with one locus, rs1800801 in *MGP*, associating with EHOA under recessive mode of inheritance with OR=2.0, compared with additive association with finger OA, thus differentiating EHOA from finger OA. All four risk variants associated with lowered hand grip strength. Two of the EHOA variants, rs17013495 (*SPP1/MEPE*) and rs11243284 (*BMP6*), only associated with EHOA and/or hand OA, and no other type of OA. Of special note is the opposite effect of rs1800801 in *MGP* and rs11631127 in *ALDH1A2* on knee OA compared with EHOA, that is, the EHOA risk allele of these variants confer protection of knee OA. The likely EHOA candidate genes at these loci implicate roles of cartilage calcification (*MGP*), vitamin A (*ALDH1A2*) and bone/cartilage mineralisation/remodelling (*BMP6*, SPP1/*IBSP/MEPE*) pathways in EHOA. Moreover, our results support the notion that EHOA is a severe form of hand OA as evident by higher risk of the EHOA and reported hand OA variants in EHOA than of fingers/thumbs OA, as well as high genetic correlation. Our results also indicate some genetic, and or functional or biological, distinction between EHOA and OA in the larger joints, since the EHOA risk alleles either do not confer risk, or confer protection, of OA in these joints, and the lower genetic correlation.

Author affiliations
[1]Population Genomics, deCODE genetics / Amgen Inc, Reykjavik, Iceland
[2]Statistics, deCODE genetics / Amgen Inc, Reykjavik, Iceland
[3]Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland
[4]Department of Human Genetics, University of Utah, Salt Lake City, Utah, USA
[5]Department of Orthopaedics, University of Utah, Salt Lake City, Utah, USA
[6]Rheumatology Division, A Coruna University Hospital, A Coruna, Spain
[7]Enterprise Analytics, Intermountain Healthcare, Salt Lake City, Utah, USA
[8]Radboud University Medical Center, Radboud Institute for Health Science, Radboud University, Nijmegen, The Netherlands
[9]Department of Phisiotherapy, Medicine and Biomedical Sciences, A Coruna University Hospital, A Coruna, Spain
[10]Precision Genomics, Intermountain Healthcare, Salt Lake City, Utah, USA
[11]Nutrition and Integrative Physiology, University of Utah, Salt Lake City, Utah, USA
[12]Department of Rheumatology, Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
[13]School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland
[14]Department of Medicine, Landspitali University Hospital, Reykjavik, Iceland
[15]deCODE genetics / Amgen Inc, Reykjavik, Capital, Iceland

Twitter Ignacio Rego-Perez @nacho_rego

Acknowledgements We thank the study subjects for their valuable participation. This research has been conducted using the UK Biobank Recourse under application numbers 23359 and 56270.

Contributors US, UT, GT, SS, HJ, SS, MJJ and KS designed the study and interpreted the results. DGF, LS, GT, SHL and US coordinated or performed statistical and bioinformatics analyses. OAS performed analyses of regulatory regions. MJJ, NK and KH coordinated, analysed or carried out zebrafish experiments. HJ, US, TR, SS, LAK, MK, TB, CH, FJB, NO, IR-P, MJJ, NK and KN carried out subject recruitment, ascertainment or managed phenotype data. US, UT, GT, SS and MJJ drafted the manuscript. MK, TB and FJB drafted the manuscript for intellectual content. All authors contributed to the final version of the manuscript. US and KS are responsible for the overall content as guarantors.

Funding The study is sponsored by deCODE genetics/Amgen. The Hostas study is financially supported by the Dutch Arthritis Society, who did not have influence on the design, performance and analysis of the study.

Competing interests The authors US, LS, GT, OAS, SS, SHL, DFG, UT and KS are affiliated with deCODE genetics/Amgen and declare competing interests as employees. The author MK reports consultancy/lecture fees outside the submitted work from Pfizer, Novartis, UCB, Galapagos, Flexion, Kiniksa, Jansen, Abbvie, CHDR, all paid to institution. Royalties from Wolters Kluwer and Springer Verlag, paid to institution.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants. All Icelandic participants who donated samples gave informed consent and the National Bioethics Committee of Iceland approved the study (VSN_14–148, VSN_14–015 v8), which was conducted in agreement with conditions issued by the Data Protection Authority of Iceland. The Dutch subjects gave written informed consent and the study was approved by the Leiden University Medical Center medical ethical committee (CCMO reference NL26201.058.08) and the Institutional Review Board of the Radboud University Medical CenterCentre. All UK participants gave informed consent and UK Biobank's scientific protocol and operational procedures were reviewed and approved by the North West Research Ethics Committee. US subjects were informed of the study protocol and procedures prior to providing consent and the study was approved by the Institutional Review Board at the University of Utah (IRB#: 79442) and Intermountain Healthcare (IRB#: 1051071). PROCOAC samples belong to the Sample Collection for Research on Rheumatic Diseases authorised by the Galician Research Ethics Committee (CAEIG) with registry code 2013/107 and inscribed in the National Registry of Biobanks - Collections Section (Code C.0000424). The patients have signed an informed consent agreement form prior to collection

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The summary results from the EHOA meta-analysis are available at https://www.decode.com/summarydata. Other data are available on reasonable request.

ORCID iDs
Unnur Styrkarsdottir http://orcid.org/0000-0001-8146-8278
Saedis Saevarsdottir http://orcid.org/0000-0001-9392-6184
Ignacio Rego-Perez http://orcid.org/0000-0003-1754-1164
Francisco J Blanco http://orcid.org/0000-0001-9821-7635
Margreet Kloppenburg http://orcid.org/0000-0002-9294-2307
Michael J Jurynec http://orcid.org/0000-0003-2702-9164

REFERENCES

1 Kloppenburg M, Kwok W-Y. Hand osteoarthritis -- a heterogeneous disorder. *Nat Rev Rheumatol* 2011;8:22–31.
2 Marshall M, Watt FE, Vincent TL, *et al*. Hand osteoarthritis: clinical phenotypes, molecular mechanisms and disease management. *Nat Rev Rheumatol* 2018;14:641–56.
3 Favero M, Belluzzi E, Ortolan A, *et al*. Erosive hand osteoarthritis: latest findings and outlook. *Nat Rev Rheumatol* 2022;18:171–83.
4 Bijsterbosch J, van Bemmel JM, Watt I, *et al*. Systemic and local factors are involved in the evolution of erosions in hand osteoarthritis. *Ann Rheum Dis* 2011;70:326–30.
5 Oreiro-Villar N, Raga AC, Rego-Pérez I. PROCOAC (prospective cohort of A coruñA) description: spanish prospective cohort to study osteoarthritis. *Reumatol Clin (Engl Ed)* 2020;18:100–4.
6 Pattrick M, Manhire A, Ward AM, *et al*. Hla-A, B antigens and alpha 1-antitrypsin phenotypes in nodal generalised osteoarthritis and erosive osteoarthritis. *Ann Rheum Dis* 1989;48:470–5.
7 Ramonda R, Musacchio E, Campana C, *et al*. Immunogenetic aspects of erosive osteoarthritis of the hand in patients from northern Italy. *Scand J Rheumatol* 2011;40:139–44.
8 Stern AG, de Carvalho MRC, Buck GA, *et al*. Association of erosive hand osteoarthritis with a single nucleotide polymorphism on the gene encoding interleukin-1 beta. *Osteoarthritis Cartilage* 2003;11:394–402.
9 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.
10 den Hollander W, Boer CG, Hart DJ, *et al*. Genome-Wide association and functional studies identify a role for matrix Gla protein in osteoarthritis of the hand. *Ann Rheum Dis* 2017;76:2046–53.
11 Boer CG, Yau MS, Rice SJ, *et al*. Genome-wide association of phenotypes based on clustering patterns of hand osteoarthritis identify *wnt9a* as novel osteoarthritis gene. *Ann Rheum Dis* 2021;80:367–75.
12 Boer CG, Hatzikotoulas K, Southam L, *et al*. Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell* 2021;184:6003–5.
13 Verbruggen G, Veys EM. Numerical scoring systems for the anatomic evolution of osteoarthritis of the finger joints. *Arthritis Rheum* 1996;39:308–20.
14 Damman W, Liu R, Kroon FPB, *et al*. Do comorbidities play a role in hand osteoarthritis disease burden? data from the hand osteoarthritis in secondary care cohort. *J Rheumatol* 2017;44:1659–66.
15 Wetzels JFM, Kiemeney LALM, Swinkels DW, *et al*. Age- and gender-specific reference values of estimated GFR in Caucasians: the Nijmegen biomedical study. *Kidney Int* 2007;72:632–7.
16 Kazmers NH, Meeks HD, Novak KA, *et al*. Familial clustering of erosive hand osteoarthritis in a large statewide cohort. *Arthritis Rheumatol* 2021;73:440–7.
17 Oreiro-Villar N, Raga AC, Rego-Pérez I, *et al*. PROCOAC (prospective cohort of a coruñA) description: Spanish prospective cohort to study osteoarthritis. *Reumatol Clin (Engl Ed)* 2020:S1699-258X(20)30231-X.
18 MANTEL N, HAENSZEL W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959;22:719–48.
19 Sveinbjornsson G, Albrechtsen A, Zink F, *et al*. Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nat Genet* 2016;48:314–7.
20 Kong A, Frigge ML, Thorleifsson G, *et al*. Selection against variants in the genome associated with educational attainment. *Proc Natl Acad Sci U S A* 2017;114:E727–32.
21 Bulik-Sullivan BK, Loh P-R, Finucane HK, *et al*. Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015;47:291–5.
22 ENCODE Project Consortium, Moore JE, Purcaro MJ, *et al*. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* 2020;583:699–710.
23 Meuleman W, Muratov A, Rynes E, *et al*. Index and biological spectrum of human DNase I hypersensitive sites. *Nature* 2020;584:244–51.
24 Giambartolomei C, Vukcevic D, Schadt EE, *et al*. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* 2014;10:e1004383.
25 Styrkarsdottir U, Lund SH, Thorleifsson G, *et al*. Meta-Analysis of Icelandic and UK data sets identifies missense variants in SMO, IL11, COL11A1 and 13 more new loci associated with osteoarthritis. *Nat Genet* 2018;50:1681–7.
26 Ferkingstad E, Sulem P, Atlason BA, *et al*. Large-Scale integration of the plasma proteome with genetics and disease. *Nat Genet* 2021;53:1712–21.
27 Zhang Y, Grant RA, Shivakumar MK, *et al*. Genome-Wide association analysis across 16,956 patients identifies a novel genetic association between BMP6, NIPAL1, CNGA1 and spondylosis. *Spine (Phila Pa 1976)* 2021;46:E625–31.
28 Boix CA, James BT, Park YP, *et al*. Regulatory genomic circuitry of human disease loci by integrative epigenomics. *Nature* 2021;590:300–7.
29 Shepherd C, Zhu D, Skelton AJ, *et al*. Functional characterization of the osteoarthritis genetic risk residing at aldh1a2 identifies rs12915901 as a key target variant. *Arthritis Rheumatol* 2018;70:1577–87.
30 Shepherd C, Reese AE, Reynard LN, *et al*. Expression analysis of the osteoarthritis genetic susceptibility mapping to the matrix Gla protein gene MGP. *Arthritis Res Ther* 2019;21:149.
31 Canalis E, Economides AN, Gazzerro E. Bone morphogenetic proteins, their antagonists, and the skeleton. *Endocr Rev* 2003;24:218–35.
32 Gitelman SE, Kobrin MS, Ye JQ, *et al*. Recombinant vgr-1/BMP-6-expressing tumors induce fibrosis and endochondral bone formation in vivo. *J Cell Biol* 1994;126:1595–609.
33 Wu M, Chen G, Li YP. Tgf-B and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. *Bone Res* 2016;4:16009.
34 Ramos YFM, den Hollander W, Bovée JVMG, *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.
35 Houtman E, Coutinho de Almeida R, Tuerlings M, *et al*. Characterization of dynamic changes in matrix gla protein (MGP) gene expression as function of genetic risk alleles, osteoarthritis relevant stimuli, and the vitamin K inhibitor warfarin. *Osteoarthritis Cartilage* 2021;29:1193–202.

36 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.
37 Newman B, Gigout LI, Sudre L, *et al*. Coordinated expression of matrix Gla protein is required during endochondral ossification for chondrocyte survival. *J Cell Biol* 2001;154:659–66.
38 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.
39 Estrada K, Styrkarsdottir U, Evangelou E, *et al*. Genome-Wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet* 2012;44:491–501.
40 Medina-Gomez C, Kemp JP, Trajanoska K, *et al*. Life-course genome-wide association study meta-analysis of total body BMD and assessment of age-specific effects. *Am J Hum Genet* 2018;102:88–102.
41 Morris JA, Kemp JP, Youlten SE, *et al*. An atlas of genetic influences on osteoporosis in humans and mice. *Nat Genet* 2019;51:258–66.
42 Petersen DN, Tkalcevic GT, Mansolf AL, *et al*. Identification of osteoblast/osteocyte factor 45 (OF45), a bone-specific cDNA encoding an RGD-containing protein that is highly expressed in osteoblasts and osteocytes. *J Biol Chem* 2000;275:36172–80.
43 Fisher LW, Fedarko NS. Six genes expressed in bones and teeth encode the current members of the sibling family of proteins. *Connect Tissue Res* 2003;44 Suppl 1:33–40.
44 Gullard A, Gluhak-Heinrich J, Papagerakis S, *et al*. MEPE localization in the craniofacial complex and function in tooth dentin formation. *J Histochem Cytochem* 2016;64:224–36.
45 Malaval L, Aubin JE, Vico L. Role of the small integrin-binding ligand N-linked glycoprotein (sibling), bone sialoprotein (BSP) in bone development and remodeling. *Osteoporos Int* 2009;20:1077–80.
46 Xu H, Tong G, Yan T, *et al*. Transcriptomic analysis provides insights to reveal the *bmp6* function related to the development of intermuscular bones in zebrafish. *Front Cell Dev Biol* 2022;10:821471.
47 Beederman M, Lamplot JD, Nan G, *et al*. Bmp signaling in mesenchymal stem cell differentiation and bone formation. *J Biomed Sci Eng* 2013;6:32–52.
48 Solloway MJ, Dudley AT, Bikoff EK, *et al*. Mice lacking BMP6 function. *Dev Genet* 1998;22:321–39.
49 Jurynec MJ, Gavile CM, Honeggar M, *et al*. The NOD/RIPK2 signaling pathway contributes to osteoarthritis susceptibility. *Genetics* [Preprint] 2022.

Supplementary Table 1. Association of EHOA GWS variants in the individual study populations

											Iceland (N = 918 cases / 109,249 controls)				UK Biobank (N = 63 cases / 430,875 controls)				The Netherlands (N = 139 cases / 5,102 controls)				Spain (N = 218 cases / 164 controls)				USA (N = 145 cases / 5,308 controls)			
rsName	Chr	Pos_build38	Effect Allele	Other Allele	Freq (%) EA	P_value	OR	P_value_GWS_threshold	P_bonferroni	P_{het}	I2	P value	OR	Freq (%) EA	Info	P value	OR	Freq (%) EA	Info	P value	OR	Freq (%) EA	Info	P value	OR	Freq (%) EA	Info	P value	OR	Freq (%) EA	Info
rs17013495	chr4	87885460	C	T	42.7	8.4E-14	0.72	1.2E-09	3.5E-06	0.18	36.9	2.5E-11	0.70	38.7	1.000	0.71	1.07	43.5	0.999	0.042	0.76	43.0	0.995	5.2E-03	0.60	45.9	0.987	0.020	0.71	42.3	0.990
rs11243284	chr6	8945086	C	T	29.8	4.2E-11	1.35	4.0E-10	5.2E-03	0.67	0	2.9E-08	1.35	29.5	1.000	0.33	1.20	30.3	1.000	0.15	1.22	29.7	0.987	0.011	1.71	29.7	0.817	0.018	1.47	30.0	0.867
rs1800801	chr12	14885854	T	C	39.0	3.6E-13	1.37	2.4E-09	7.5E-06	0.17	38.0	5.0E-07	1.30	37.2	1.000	8.3E-04	1.84	37.8	0.999	7.1E-05	1.70	37.4	1.000	0.14	1.29	45.1	0.989	0.026	1.38	37.3	0.988
rs11631127	chr15	57977811	G	C	47.8	7.1E-18	0.69	4.0E-10	8.9E-10	0.59	0	2.3E-11	0.70	42.4	1.000	0.25	0.81	46.5	1.000	8.4E-05	0.60	47.3	0.998	0.058	0.70	55.8	0.995	5.6E-04	0.61	47.1	0.994

Supplementary Table 2. Association of EHOA variants with EHOA under additive, recessive and full genotype models

		Additive model			Recessive model			Genotype specific model: Heterozygotes			Genotype specific model: Homozygotes			
Variant[allele]	**Chr**	**OR (95% CI)**	**P value**	**P_{het}**	**OR (95% CI)**	**P value**	**P_{het}**	**OR (95% CI)**	**P value**	**P_{het}**	**OR (95% CI)**	**P value**	**P_{het}**	**P model**
rs17013495[T]	chr4	1.395 (1.279-1.522)	8.77E-14	0.17	1.600 (1.346-1.901)	9.64E-08	0.16	1.342 (1.098-1.641)	0.0041	0.48	2.011 (1.630-2.481)	7.07E-11	0.63	0.289
rs11243284[C]	chr6	1.354 (1.237-1.482)	4.22E-11	0.67	1.674 (1.387-2.022)	8.47E-08	0.41	1.307 (1.147-1.491)	6.27E-05	0.77	1.773 (1.457-2.157)	1.07E-08	0.66	0.446
rs1800801[T]	chr12	1.368 (1.257-1.488)	3.55E-13	0.17	1.848 (1.594-2.143)	3.86E-16	0.14	1.151 (1.002-1.323)	0.047	0.87	2.012 (1.705-2.373)	1.09E-16	0.34	0.0011
rs11631127[C]	chr15	1.456 (1.337-1.587)	7.15E-18	0.59	1.608 (1.376-1.880)	2.68E-09	0.45	1.320 (1.095-1.591)	0.0036	0.42	2.089 (1.726-2.528)	3.69E-14	0.48	0.241

Association of the four EHOA variants with EHOA is shown for the additive model, the recessive model, and for the full model evaluating risk at the heterozygous genotypes and homozygous genotypes. The effect allele of each variant is shown within square brackets, with the odds ratio (OR) with 95% confidence interval (CI), the P value, and the heterogeneity P value (P_{het}) for each model, and the P value (P model) for deviation from the additive model.

Supplementary Table 3. Association of rs1800801 in 5'UTR of MGP with hand osteoarthritis subtypes under additive, recessive and full genotype model

		Additive model			Recessive model			Genotype specific model: Heterozygotes			Genotype specific model: Homozygotes			
Phenotype	N cases / N controls	OR (95% CI)	P value	Phet	OR (95% CI)	P value	Phet	OR (95% CI)	P value	Phet	OR (95% CI)	P value	Phet	P model
Erosive hand OA	1,484 / 550,680	1.368 (1.257-1.488)	3.6E-13	0.17	1.848 (1.594-2.143)	3.9E-16	0.14	1.151 (1.002-1.323)	0.047	0.87	2.012 (1.705-2.373)	1.1E-16	0.34	0.0011
Finger OA	7,871 / 608,869	1.143 (1.099-1.188)	1.5E-11	0.026	1.258 (1.173-1.349)	1.1E-10	0.035	1.103 (1.037-1.173)	0.0017	0.60	1.349 (1.242-1.464)	1.0E-12	0.031	0.12
Thumb OA	9,865 / 623,814	1.066 (1.031-1.103)	1.6E-04	0.28	1.064 (0.999-1.133)	0.055	0.22	1.108 (1.052-1.166)	1.0E-04	0.44	1.129 (1.050-1.214)	0.0010	0.19	0.056
Hand OA	14,841 / 626,618	1.080 (1.050-1.111)	8.4E-08	0.081	1.132 (1.074-1.193)	3.31E-06	0.059	1.073 (1.027-1.120)	0.0016	0.34	1.181 (1.112-1.254)	6.8E-08	0.08	0.19

Association of the rs180081[T] with hand osteoarthritis subtypes is shown for the additive model, the recessive model, and for the full model evaluating risk at the heterozygous genotypes and homozygous genotypes. The odds ratio (OR) with 95% confidence interval (CI), the P value, and the heterogeneity P value (Phet) is shown for each model, and the P value (P model) for deviation from the additive model. The finger, thumb and hand OA analysis included data from Iceland, US, UK, and The Netherlands, whereas all datasets were included in the erosive hand OA analysis.

Supplementary Table 4. EHOA variants, or their correlated variants, are located in regions defined as candidate cis-regulatory elements by ENCODE project (screen.encodeproject.org).

	GWAS association, lead sequence variant for each signal			
cCRE annotation:	**rs17013495** (chr4:87885460) LD class = 68	**rs11243284** (chr6:8945086) LD class = 17	**rs1800801** (chr12:14885854) LD class = 107	**rs11631127** (chr15:57977811) LD class =155
DNase-H3K4me3				chr15:58008570:SG
Promoter-like sequence (PLS)				chr15:58065219:IG
Promoter-like sequence (PLS)-CTCF-bound			chr12:14885854:SG	
Enhancer-like sequence, distal (dELS)	chr4:87868563:SG, chr4:87868643:SG	chr6:8948008:SG, chr6:8948226:SG	chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14851053:SG, chr12:14851097:IG, chr12:14899824:SG, chr12:14899901:SG, chr12:14900018:SG, chr12:14910656:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:58040343:SG, chr15:58040385:SG
Enhancer-like sequence, distal (dELS)-CTCF-bound	chr4:87863666:SG, chr4:87885460:SG	chr6:8949691:SG	chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14901082:SG	chr15:57923529:SG
Enhancer-like sequence, proximal (pELS)			chr12:14839301:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14883768:SG	chr15:58063976:IG, chr15:58064657:SG
Enhancer-like sequence, proximal (pELS)-CTCF bound				chr15:58064164:SG

The variants are shown by their position in Build38, with SG ending for SNPs and IG for indels

Supplementary Table 5. Location of EHOA variants, or their correlated variants, in candidate cis-regulatory elements (cCRE) by tissue or cell type (encodeproject.org)

Tissue/cell type	UBERON/CL	rs17013495 (chr4:87885460)	rs11243284 (chr6:8945086)	rs1800801 (chr12:14885854)	rs11631127 (chr15:57977811)
activated CD4-positive, alpha-beta T cell	CL_0000896			chr12:14910656:SG	chr15:58065219:IG
activated CD8-positive, alpha-beta T cell	CL_0000906			chr12:14910656:SG	
adipose tissue	UBERON_0001013			chr12:14885854:SG	
adrenal gland	UBERON_0002369			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14910656:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
amnion	UBERON_0000305			chr12:14836364:SG, chr12:14885854:SG	chr15:58065219:IG
amniotic stem cell	CL_0002639			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14885854:SG	
aorta	UBERON_0000947			chr12:14836364:SG, chr12:14839301:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG
ascending aorta	UBERON_0001496			chr12:14885854:SG	
astrocyte	CL_0000127	chr4:87885460:SG		chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG, chr12:14910656:SG	
astrocyte of the cerebellum	CL_0002603			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG	
astrocyte of the hippocampus	CL_0002604			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG	
astrocyte of the spinal cord	CL_0002606			chr12:14836364:SG, chr12:14885854:SG	
B cell	CL_0000236				chr15:58065219:IG
bipolar neuron	CL_0000103		chr6:8948008:SG, chr6:8948226:SG, chr6:8949691:SG		chr15:58065219:IG
body of pancreas	UBERON_0001150			chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG	chr15:57923529:SG
brain	UBERON_0000955			chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
brain microvascular endothelial cell	CL_2000044	chr4:87868563:SG, chr4:87868643:SG		chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14885854:SG	
brain pericyte	CL_2000043			chr12:14836364:SG	chr15:57923529:SG
breast epithelium	UBERON_0008367			chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG, chr12:14910656:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
bronchial epithelial cell	CL_0002328			chr12:14836364:SG, chr12:14910656:SG	chr15:58040343:SG, chr15:58040385:SG
cardiac muscle cell	CL_0000746			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14910656:SG	chr15:57923529:SG, chr15:58065219:IG
cardiac myoblast	CL_0010021			chr12:14851053:SG, chr12:14851097:IG	chr15:58065219:IG
cardiovascular progenitor cell	CL_0002664				chr15:58064164:SG, chr15:58064657:SG
caudate nucleus	UBERON_0001873			chr12:14885854:SG, chr12:14910656:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	
CD14-positive monocyte	CL_0001054	chr4:87885460:SG		chr12:14836364:SG	chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
CD1c-positive myeloid dendritic cell	CL_0002399				chr15:58065219:IG
CD4-positive, alpha-beta memory T cell	CL_0000897				chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
CD4-positive, alpha-beta T cell	CL_0000624			chr12:14839301:SG, chr12:14910656:SG	chr15:58065219:IG
CD4-positive, CD25-positive, alpha-beta regulatory T cell	CL_0000792				chr15:58065219:IG
CD8-positive, alpha-beta memory T cell	CL_0000909				chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
CD8-positive, alpha-beta T cell	CL_0000625			chr12:14836364:SG	
cell of skeletal muscle	CL_0000188			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14885854:SG, chr12:14910656:SG	chr15:57923529:SG
cerebellar cortex	UBERON_0002129			chr12:14883768:SG, chr12:14885854:SG	
cerebellum	UBERON_0002037			chr12:14883768:SG	
chorion	UBERON_0003124			chr12:14836364:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG	chr15:58065219:IG
chorionic villus	UBERON_0007106			chr12:14836364:SG, chr12:14839301:SG, chr12:14840674:SG, chr12:14840920:SG	
choroid plexus epithelial cell	CL_0000706			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14910656:SG	chr15:57923529:SG
cingulate gyrus	UBERON_0002967			chr12:14885854:SG	
colonic mucosa	UBERON_0000317			chr12:14885854:SG	chr15:57923529:SG
common myeloid progenitor, CD34-positive	CL_0001059	chr4:87885460:SG		chr12:14885854:SG	chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
coronary artery	UBERON_0001621			chr12:14885854:SG	
dermis blood vessel endothelial cell	CL_2000010			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG, chr12:14910656:SG	chr15:57923529:SG, chr15:58065219:IG
dermis microvascular lymphatic vessel endothelial cell	CL_2000041			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14910656:SG	chr15:57923529:SG
duodenal mucosa	UBERON_0000320			chr12:14885854:SG	
ectodermal cell	CL_0000221				chr15:58065219:IG
effector memory CD4-positive, alpha-beta T cell	CL_0000905				chr15:58065219:IG
embryonic facial prominence	UBERON_0012314			chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58064657:SG
endocrine pancreas	UBERON_0000016			chr12:14839301:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14883768:SG, chr12:14885854:SG	chr15:58064657:SG, chr15:58065219:IG
endodermal cell	CL_0000223				chr15:58065219:IG
endothelial cell of umbilical vein	CL_0002618			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG, chr12:14910656:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
epithelial cell of amnion	CL_0002536			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14885854:SG	chr15:58040343:SG, chr15:58040385:SG
epithelial cell of esophagus	CL_0002252			chr12:14836364:SG, chr12:14910656:SG	chr15:58040343:SG, chr15:58040385:SG
epithelial cell of prostate	CL_0002231			chr12:14836364:SG, chr12:14910656:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:58040343:SG, chr15:58040385:SG
epithelial cell of proximal tubule	CL_0002306			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG	
esophagus	UBERON_0001043			chr12:14885854:SG	chr15:58065219:IG
esophagus muscularis mucosa	UBERON_0004648			chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
esophagus squamous epithelium	UBERON_0006920			chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
eye	UBERON_0000970				chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
fat cell	CL_0000136			chr12:14836364:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG
femur	UBERON_0000981			chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG
fibroblast of cardiac tissue	CL_0002548			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG, chr12:14910656:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
fibroblast of dermis	CL_0002551			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG
fibroblast of gingiva	CL_0002552			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14885854:SG	chr15:57923529:SG
fibroblast of lung	CL_0002553	chr4:87885460:SG		chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG, chr12:14910656:SG	chr15:57923529:SG, chr15:58064657:SG
fibroblast of mammary gland	CL_0002555			chr12:14836364:SG, chr12:14885854:SG	chr15:57923529:SG
fibroblast of peridontal ligament	CL_2000017			chr12:14836364:SG, chr12:14885854:SG	
fibroblast of pulmonary artery	CL_0002557			chr12:14836364:SG, chr12:14885854:SG, chr12:14910656:SG	chr15:57923529:SG, chr15:58065219:IG
fibroblast of skin of abdomen	CL_2000013			chr12:14836364:SG, chr12:14885854:SG	
fibroblast of skin of back	CL_0011022			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG	
fibroblast of the aortic adventitia	CL_0002547			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG	
fibroblast of the conjunctiva	CL_0002550			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG	chr15:57923529:SG
fibroblast of upper back skin	CL_0011021			chr12:14836364:SG, chr12:14885854:SG	
fibroblast of villous mesenchyme	CL_0002558			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14854918:IG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG
forelimb muscle	UBERON_0003662			chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG
foreskin fibroblast	CL_1001608			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG
foreskin keratinocyte	CL_1001606			chr12:14836364:SG, chr12:14910656:SG	chr15:58040343:SG, chr15:58040385:SG, chr15:58065219:IG
foreskin melanocyte	CL_2000045	chr4:87885460:SG		chr12:14836364:SG, chr12:14901082:SG	
gastrocnemius medialis	UBERON_0011907			chr12:14836364:SG, chr12:14885854:SG	chr15:57923529:SG

gastroesophageal sphincter	UBERON_0004550	chr4:87863666:SG		chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
germinal center	UBERON_0010754	chr4:87885460:SG		chr12:14854918:IG	
glomerular endothelial cell	CL_0002188			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG, chr12:14910656:SG	chr15:57923529:SG
glutamatergic neuron	CL_0000679		chr6:8949691:SG		
heart	UBERON_0000948			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
heart left ventricle	UBERON_0002084			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
heart right ventricle	UBERON_0002080			chr12:14834162:SG, chr12:14834298:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
hematopoietic multipotent progenitor cell	CL_0000837	chr4:87885460:SG		chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14839301:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
hepatocyte	CL_0000182			chr12:14836364:SG, chr12:14885854:SG	chr15:58065219:IG
hindlimb muscle	UBERON_0003663			chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58065219:IG
immature natural killer cell	CL_0000823				chr15:58065219:IG
inferior parietal cortex	UBERON_0006088			chr12:14883768:SG	
inflammatory macrophage	CL_0000863			chr12:14836364:SG	chr15:58065219:IG
iris pigment epithelial cell	CL_0002565			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14910656:SG	
keratinocyte	CL_0000312			chr12:14836364:SG	
kidney	UBERON_0002113	chr4:87863666:SG		chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
kidney capillary endothelial cell	CL_1000892			chr12:14836364:SG, chr12:14910656:SG	
kidney epithelial cell	CL_0002518			chr12:14836364:SG	
kidney glomerular epithelial cell	CL_1000510			chr12:14836364:SG	chr15:58065219:IG
kidney tubule cell	CL_1000507			chr12:14836364:SG	
large intestine	UBERON_0000059	chr4:87868563:SG, chr4:87868643:SG		chr12:14836364:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:57923529:SG, chr15:58040343:SG, chr15:58040385:SG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
layer of hippocampus	UBERON_0002305			chr12:14885854:SG	chr15:58065219:IG
left cardiac atrium	UBERON_0002079			chr12:14885854:SG	chr15:57923529:SG, chr15:58065219:IG
left colon	UBERON_0008971			chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	
left forelimb	UBERON_8300002			chr12:14885854:SG	chr15:58065219:IG
left hindlimb	UBERON_8300004			chr12:14883768:SG, chr12:14885854:SG	chr15:58065219:IG
left kidney	UBERON_0004538			chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
left lobe of liver	UBERON_0001115				chr15:57923529:SG
left lung	UBERON_0002168			chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
left renal cortex interstitium	UBERON_0018117			chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
left renal pelvis	UBERON_0018115			chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
limb	UBERON_0002101			chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
liver	UBERON_0002107			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14840674:SG, chr12:14840920:SG	chr15:57923529:SG, chr15:58065219:IG
lower leg skin	UBERON_0004264			chr12:14885854:SG	
lower lobe of left lung	UBERON_0008953	chr4:87863666:SG			chr15:58065219:IG
lung	UBERON_0002048			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58040343:SG, chr15:58040385:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
lung microvascular endothelial cell	CL_2000016			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14910656:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
mammary gland epithelial cell	CL_0002327			chr12:14836364:SG, chr12:14885854:SG, chr12:14910656:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:58040343:SG, chr15:58040385:SG
medulla oblongata	UBERON_0001896			chr12:14885854:SG	chr15:57923529:SG
mesenchymal stem cell	CL_0000134			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14854918:IG, chr12:14885854:SG	
mesodermal cell	CL_0000222			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
middle frontal area -8	UBERON_0006483			chr12:14883768:SG, chr12:14885854:SG	chr15:58065219:IG
middle frontal gyrus	UBERON_0002702			chr12:14883768:SG	
mucosa of descending colon	UBERON_0004992			chr12:14885854:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:57923529:SG
mucosa of gallbladder	UBERON_0005033			chr12:14885854:SG	chr15:57923529:SG, chr15:58065219:IG
mucosa of rectum	UBERON_0003346			chr12:14883768:SG, chr12:14885854:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:57923529:SG, chr15:58065219:IG
mucosa of stomach	UBERON_0001199			chr12:14885854:SG	chr15:57923529:SG
muscle cell	CL_0000187			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG	chr15:57923529:SG
muscle layer of colon	UBERON_0012489			chr12:14883768:SG, chr12:14885854:SG	chr15:58065219:IG
muscle layer of duodenum	UBERON_0012488			chr12:14883768:SG, chr12:14885854:SG	chr15:58064657:SG, chr15:58065219:IG
muscle of arm	UBERON_0001499			chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
muscle of back	UBERON_0002324			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
muscle of leg	UBERON_0001383			chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
myoepithelial cell of mammary gland	CL_0002324			chr12:14885854:SG	chr15:58065219:IG
myotube	CL_0002372			chr12:14836364:SG	chr15:57923529:SG
naive thymus-derived CD4-positive, alpha-beta T cell	CL_0000895			chr12:14910656:SG	chr15:58065219:IG
natural killer cell	CL_0000623			chr12:14885854:SG	chr15:58064657:SG
nephron	UBERON_0001285			chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
neural cell	CL_0002319				chr15:58065219:IG
neural crest cell	CL_0011012			chr12:14851053:SG, chr12:14851097:IG	chr15:58065219:IG
neural progenitor cell	CL_0011020				chr15:58064164:SG, chr15:58065219:IG
neuroepithelial stem cell	CL_0002259				chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
neuronal stem cell	CL_0000047				chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
neutrophil	CL_0000775				chr15:58065219:IG
non-pigmented ciliary epithelial cell	CL_0002304			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14910656:SG	
omental fat pad	UBERON_0010414				chr15:58065219:IG
ovary	UBERON_0000992			chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064657:SG, chr15:58065219:IG
pancreas	UBERON_0001264			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG	chr15:57923529:SG
peripheral blood mononuclear cell	CL_2000001				chr15:58065219:IG
Peyers patch	UBERON_0001211				chr15:58040343:SG, chr15:58040385:SG
placenta	UBERON_0001987			chr12:14836364:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG, chr12:14885854:SG	chr15:58064657:SG, chr15:58065219:IG
podocyte	CL_0000653			chr12:14836364:SG	
posterior cingulate gyrus	UBERON_0002740	chr4:87863666:SG			
prostate gland	UBERON_0002367			chr12:14883768:SG, chr12:14885854:SG	chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
psoas muscle	UBERON_0008450			chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14839301:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
pulmonary artery endothelial cell	CL_1001568			chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG	chr15:57923529:SG
radial glial cell	CL_0000681				chr15:58065219:IG
rectal smooth muscle tissue	UBERON_0018112			chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	

renal cortex interstitium	UBERON_0005270		chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
renal cortical epithelial cell	CL_0002584		chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG	
renal pelvis	UBERON_0001224		chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
retina	UBERON_0000966			chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
retinal pigment epithelial cell	CL_0002586		chr12:14836364:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14885854:SG	
right atrium auricular region	UBERON_0006631		chr12:14885854:SG	
right cardiac atrium	UBERON_0002078		chr12:14836364:SG, chr12:14839301:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
right forelimb	UBERON_8300001		chr12:14883768:SG, chr12:14885854:SG	chr15:58065219:IG
right hindlimb	UBERON_8300003		chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58065219:IG
right kidney	UBERON_0004539		chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
right lobe of liver	UBERON_0001114			chr15:57923529:SG
right lung	UBERON_0002167		chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
right renal cortex interstitium	UBERON_0018118		chr12:14854918:IG, chr12:14883768:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
right renal pelvis	UBERON_0018116		chr12:14854918:IG, chr12:14883768:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
sciatic nerve	UBERON_0001322		chr12:14885854:SG	
sigmoid colon	UBERON_0001159		chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58065219:IG
skeletal muscle myoblast	CL_0000515		chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14901082:SG	chr15:57923529:SG
skeletal muscle of trunk	UBERON_0001774		chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
skeletal muscle tissue	UBERON_0001134		chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
skin epidermis	UBERON_0001003		chr12:14851053:SG, chr12:14851097:IG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG, chr12:14910656:SG	chr15:58065219:IG
skin of body	UBERON_0002097		chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG	chr15:57923529:SG, chr15:58063976:IG
small intestine	UBERON_0002108	chr4:87868563:SG, chr4:87868643:SG	chr12:14836364:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:58040343:SG, chr15:58040385:SG, chr15:58064657:SG, chr15:58065219:IG
smooth muscle cell	CL_0000192		chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG, chr12:14899824:SG, chr12:14899901:SG, chr12:14900018:SG, chr12:14910656:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	chr15:58064657:SG, chr15:58065219:IG
smooth muscle cell of the brain vasculature	CL_0002590		chr12:14836364:SG	chr15:57923529:SG
spinal cord	UBERON_0002240		chr12:14834162:SG, chr12:14834298:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
spleen	UBERON_0002106		chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG, chr12:14847226:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
stomach	UBERON_0000945	chr4:87885460:SG	chr12:14836364:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58040343:SG, chr15:58040385:SG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
stomach smooth muscle	UBERON_0004222		chr12:14883768:SG, chr12:14885854:SG	chr15:58065219:IG
stromal cell of bone marrow	CL_0010001		chr12:14836364:SG, chr12:14854918:IG, chr12:14885854:SG	chr15:57923529:SG
subcutaneous abdominal adipose tissue	UBERON_0014455		chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
substantia nigra	UBERON_0002038		chr12:14885854:SG	
superior temporal gyrus	UBERON_0002769		chr12:14883768:SG	
suppressor macrophage	CL_0000862		chr12:14836364:SG	chr15:58065219:IG
suprapubic skin	UBERON_0036149		chr12:14847029:SG, chr12:14847226:SG, chr12:14851053:SG, chr12:14851097:IG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG, chr12:14910656:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
testis	UBERON_0000473		chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
T-helper 1 cell	CL_0000545		chr12:14910656:SG	
T-helper 17 cell	CL_0000899			chr15:58065219:IG
T-helper 2 cell	CL_0000546		chr12:14910656:SG	
thoracic aorta	UBERON_0001515		chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58065219:IG
thymus	UBERON_0002370			chr15:58064657:SG, chr15:58065219:IG
thyroid gland	UBERON_0002046		chr12:14847029:SG, chr12:14847226:SG, chr12:14851053:SG, chr12:14851097:IG, chr12:14854918:IG, chr12:14885854:SG	chr15:57923529:SG
tibial artery	UBERON_0007610		chr12:14885854:SG	
tibial nerve	UBERON_0001323		chr12:14836364:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG
tongue	UBERON_0001723		chr12:14883768:SG, chr12:14885854:SG	chr15:57923529:SG, chr15:58064657:SG, chr15:58065219:IG
transverse colon	UBERON_0001157		chr12:14885854:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	
trophoblast	UBERON_0000088		chr12:14836364:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG, chr12:14910656:SG, chr12:14911149:IG, chr12:14911328:SG, chr12:14911429:SG	
trophoblast cell	CL_0000351	chr4:87868563:SG, chr4:87868643:SG, chr4:87885460:SG	chr12:14836364:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG	
umbilical cord	UBERON_0002331		chr12:14834162:SG, chr12:14834298:SG, chr12:14836364:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14899824:SG, chr12:14899901:SG, chr12:14900018:SG	chr15:57923529:SG
upper lobe of left lung	UBERON_0008952		chr12:14885854:SG	chr15:58065219:IG
ureter	UBERON_0000056		chr12:14885854:SG	
urinary bladder	UBERON_0001255	chr4:87885460:SG	chr12:14847029:SG, chr12:14847226:SG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
uterus	UBERON_0000995		chr12:14847029:SG, chr12:14847226:SG, chr12:14854918:IG, chr12:14883768:SG, chr12:14885854:SG, chr12:14901082:SG	chr15:57923529:SG, chr15:58063976:IG, chr15:58064164:SG, chr15:58064657:SG, chr15:58065219:IG
vagina	UBERON_0000996		chr12:14885854:SG	

The variants are shown by their position in Build38, with SG ending for SNPs and IG for indels

Supplementary Table 6. Co-localisation of EHOA variants and expression of genes at the EHOA loci (eQTL)

EHOA variants	chr:pos(hg38)	EA / OA	Freq% EA	OR	Gene	Tissue	eQTL variant	r^2	EA / OA	Freq% EA	Effect	*P value*	Source	# individuals/tissue	COLOC PP3	COLOC PP4
rs17013495	chr4:87885460	T/C	59.6	1.4	SPP1	Spleen	rs4693198	0.91	C/T	59.6	-0.48	1.1E-09	GTEx v8	227	0.12	0.88
					"	Esophagus – Mucosa	rs4693897	0.91	G/T	59.5	-0.32	1.4E-08	GTEx v8	497	1.00	0.00
					"	Whole blood	rs12644436	0.91	G/A	59.4	-0.18	9.0E-09	GTEx v8	670	1.00	0.00
rs1800801	chr12:14885854	T/C	37.2	1.37	MGP	Lung	rs11614330	0.98	T/C	36.9	-0.19	5.9E-13	GTEx v8	515	0.09	0.91
					"	Thyroid	rs4581512	0.95	T/G	37.4	-0.17	3.1E-08	GTEx v8	574	0.05	0.95
					"	Adipose	rs9668569	0.91	T/C	37.2	-0.53	2.8E-22	deCODE	770	0.08	0.92
					"	Blood*	rs11056199	0.89	C/A	39.9	0.40	6.9E-226	deCODE	17,940	1.00	0.00
rs11631127	chr15:57977811	C/G	57.6	1.46	ALDH1A2	Cultured fibroblasts	rs3742961	0.93	C/T	60.2	-0.30	9.0E-11	GTEx v8	483	0.14	0.86

Data is shown for datasets in GTEx and deCODE genetics. For each variant the gene whose expression is correlated with the erosive variants is shown (Gene), the tissue (Tissue), the top expression variant (eQTL variant), the correlation between the top expression variant and the erosive variant (r2), the effect allele (EA) and the other allele (OA) of the variants, the frequency of their effect allele (Freq% EA), the effect on transcription in standard deviation (Effect), the P value of the expression correlation, the source of data (Source), and the number of individuals in each analysis (# individuals/tissue). The position of the erosive variants are shown in build 38, and the OR af the association with erosive osteoarthritis. PP3 is the posterior probability for two independent signals, and PP4 is the posterior probability for one shared signals using COLOC (Giambartolomei et al, PLoS genetics. 2014;10(5):e1004383)

* The expression of MGP in blood is very low but the direction of effect is consistent with that reported by den Hollander, W. et al, 2017.

Supplementary Table 7. Co-localisation of the EHOA associated variants and levels of proteins in plasma (cis-pQTL)

Erosive variants	chr:pos(hg38)	EA / OA	Freq% EA	OR	Gene	Protein	pQTL variant	r^2	EA / OA	Freq% EA	Effect	P value	COLOC PP3	COLOC PP4	Comment
rs17013495	chr4:87885460	T/C	59.6	1.396	SPP1	Osteopontin	rs990862	0.80	T/C	65.9	-0.063	1.8E-13	1.00	0.00	Five independent cis-signals for SPP1, and 2 independent trans signals. Rs990862 explains 5% of the variance explained by the pQTLs
rs1800801	chr12:14885854	T/C	37.2	1.37	MGP	Matrix Gla Protein	rs7294636	0.99	A/G	37.4	-0.250	8.3E-111	0.12	0.88	Two independent cis-signals for the MGP protein (in opposite directions), and 6 trans signals. Rs12307494 explains 72% of the variance explained by the pQTLs.

Data is based on proteins measured in plasma from 35,339 in Iceland (deCODE genetics) using the Somalogic platform. The top variant that correlates with the levels of the protein (Protein) and its encoding gene (Gene) in plasma (pQTL variant) is shown, and the correlation between the top pQTL variant and the erosive variant (r2), the effect allele (EA) and the other allele (OA) of the variants, the frequency of their effect allele (Freq% EA), the effect on protein levels in standard deviation (Effect), and the P value of the protein level-variant correlation. The position of the erosive variants are shown in build 38 (chr:pos(hg38)), and the odds ratio (OR) af the association with erosive osteoarthritis. PP3 is the posterior probability for two independent signals, and PP4 is the posterior probability for one shared signals using COLOC (Giambartolomei et al, PLoS genetics. 2014;10(5):e1004383)

Supplementary Table 8. Lead sequence variants for two of the four EHOA signals, or their correlated variants, reside within enhancer elements that are predicted to affect nearby genes in different tissue/cell types based on EpiMap (http://compbio.mit.edu/epimap/).

Tissue / cell type	rs1800801 (chr12:14885854)	rs11631127 (chr15:57977811)
ACUTE_LYMPHOBLASTIC_LEUKEMIA	MGP (chr12:14854918:IG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14901082:SG)	
ACUTE_PROMYELOCYTIC_LEUKEMIA	MGP (chr12:14854918:IG)	
ADENOID_CYSTIC_CARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)	
ADIPOCYTE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)	
ADIPOCYTE_FROM_MSC	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG)	
ADIPOSE_TISSUE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14894016:SG, chr12:14901082:SG)	
ADRENAL_GLAND	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG), WBP11 (chr12:14847029:SG, chr12:14894016:SG)	ALDH1A2 (chr15:58061348:SG)
AMMONS_HORN	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)	
AMNION	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG), WBP11 (chr12:14844512:SG)	
AMNION_EPITHELIAL_CELL	ART4 (chr12:14835521:SG)	
AMNION_STEM_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
AMNIOTIC_FLUID_FROM_MSC	MGP (chr12:14854918:IG)	
ANGULAR_GYRUS	ART4 (chr12:14835521:SG), MGP (chr12:14894016:SG)	
AORTA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14854918:IG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14894016:SG, chr12:14899824:SG, chr12:14899901:SG, chr12:14900018:SG, chr12:14901082:SG)	
AORTA_FIBROBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG)	
ARM_MUSCLE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG), WBP11 (chr12:14894016:SG)	
ASCENDING_AORTA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG), WBP11 (chr12:14844512:SG, chr12:14890950:SG, chr12:14890963:SG), MGP (chr12:14844512:SG, chr12:14847029:SG, chr12:14854918:IG, chr12:14878220:SG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG)	
ASTROCYTE	ART4 (chr12:14835521:SG, chr12:14847029:SG), MGP (chr12:14847029:SG)	
ASTROCYTE_HIPPOCAMPUS	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG)	
ASTROCYTE_SPINAL_CORD	MGP (chr12:14854918:IG)	
B_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG)	ALDH1A2 (chr15:58061348:SG)
B_CELL_LYMPHOMA	ART4 (chr12:14835521:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14854918:IG, chr12:14878220:SG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG)	
BACK_MUSCLE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG), WBP11 (chr12:14894016:SG)	
BODY_OF_PANCREAS	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), WBP11 (chr12:14847029:SG, chr12:14894016:SG), MGP (chr12:14847029:SG, chr12:14894016:SG)	
BONE_ARM	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)	
BONE_FEMUR	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)	
BONE_LEG	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)	
BONE_MARROW_EPITHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
BONE_MARROW_STROMA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
BONE_MARROW_STROMAL_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG), WBP11 (chr12:14894016:SG)	
BRAIN	MGP (chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG)	
BRAIN_MICROVASCULAR_ENDOTHELIAL_CELL	MGP (chr12:14854918:IG)	
BRAIN_VASCULATURE_SMOOTH_MUSCLE_CELL	MGP (chr12:14894016:SG)	
BREAST_EPITHELIAL_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), GUCY2C (chr12:14911328:SG)	
BREAST_EPITHELIUM	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14854918:IG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14901082:SG), H2AJ (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), WBP11 (chr12:14894016:SG)	
BREAST_FIBROBLAST	ART4 (chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG)	
BRONCHIAL_EPITHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
BURKITT_LYMPHOMA	MGP (chr12:14854918:IG), WBP11 (chr12:14854918:IG)	
CARDIAC_FIBROBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG)	
CARDIAC_MUSCLE_DERIV	ART4 (chr12:14835521:SG)	
CARDIAC_MYOCYTE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG)	
CAUDATE_NUCLEUS	ART4 (chr12:14835521:SG), MGP (chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14894016:SG), WBP11 (chr12:14894016:SG)	
CD34_CMP	ART4 (chr12:14835521:SG), MGP (chr12:14854918:IG), GUCY2C (chr12:14911328:SG)	
CD4_T_CELL	ERP27 (chr12:14911328:SG)	ALDH1A2 (chr15:58061348:SG)
CD8_T_CELL		CGNL1 (chr15:58061348:SG), ALDH1A2 (chr15:58061348:SG), GCOM1 (chr15:58061348:SG), LIPC (chr15:58061348:SG)
CEREBELLAR_CORTEX	MGP (chr12:14897475:SG, chr12:14897803:SG)	
CEREBELLUM	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)	
CERVIX_ADENOCARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	

CHORION	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG, chr12:14854918:IG, chr12:14901082:SG), ERP27 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), H2AJ (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), WBP11 (chr12:14847029:SG)
CHORIONIC_VILLUS	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), ERP27 (chr12:14840920:SG)
CHOROID_PLEXUS_EPITHELIAL_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)
CINGULATE_GYRUS	ART4 (chr12:14835521:SG), MGP (chr12:14894016:SG), WBP11 (chr12:14894016:SG)
COLON_CARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
COLON_EPITHELIAL_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)
COLON_MUCOSA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), WBP11 (chr12:14894016:SG), MGP (chr12:14894016:SG)
COLON_MUSCLE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG), WBP11 (chr12:14894016:SG)
COLORECTAL_ADENOCARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
CONJUNCTIVA_FIBROBLAST	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG)
CORONARY_ARTERY	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG), WBP11 (chr12:14844512:SG), MGP (chr12:14844512:SG, chr12:14847029:SG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14901082:SG)
DERMIS_BLOOD_VESSEL_ENDOTHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG, chr12:14894016:SG), WBP11 (chr12:14894016:SG), GUCY2C (chr12:14911328:SG)
DERMIS_FIBROBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG, chr12:14901082:SG)
DERMIS_LYMPHATIC_VESSEL_ENDOTHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG), WBP11 (chr12:14894016:SG)
DESMOPLASTIC_MEDULLOBLASTOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
DUODENUM_MUCOSA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG)
DUODENUM_MUSCLE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG)
EMBRYONIC_FACIAL_PROMINENCE	MGP (chr12:14854918:IG, chr12:14894016:SG)
ENDOCRINE_PANCREAS	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)
ENDODERMAL_DERIV	MGP (chr12:14894016:SG)
ENDOMETRIAL_ADENOCARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG), WBP11 (chr12:14844512:SG), MGP (chr12:14847029:SG, chr12:14894016:SG)
EPIDERMAL_MELANOCYTE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
ESC	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14854918:IG)
ESOPHAGUS	WBP11 (chr12:14894016:SG), MGP (chr12:14894016:SG)
ESOPHAGUS_MUSCULARIS_MUCOSA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG, chr12:14901082:SG)
ESOPHAGUS_SQUAMOUS_EPITHELIUM	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14901082:SG)
EYE_RETINOBLASTOMA	ART4 (chr12:14835521:SG)
FIBROSARCOMA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
FORESKIN_FIBROBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)
FORESKIN_KERATINOCYTE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
FORESKIN_MELANOCYTE	ART4 (chr12:14835521:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14854918:IG, chr12:14894016:SG)
FRONTAL_CORTEX	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
GASTROCNEMIUS_MEDIALIS	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG), WBP11 (chr12:14894016:SG)
GASTROESOPHAGEAL_SPHINCTER	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG, chr12:14901082:SG)
GERMINAL_CENTER	MGP (chr12:14854918:IG)
GERMINAL_MATRIX	ART4 (chr12:14835521:SG)
GINGIVAL_FIBROBLAST	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)
GLIOBLASTOMA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
GLOBUS_PALLIDUS	MGP (chr12:14894016:SG), WBP11 (chr12:14894016:SG)
GLOMERULUS_ENDOTHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG)
GLOMERULUS_EPITHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
GLOMERULUS_VISCERAL_EPITHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
HEART	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG, chr12:14878220:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG), WBP11 (chr12:14894016:SG)
HEART_LEFT_ATRIUM	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), WBP11 (chr12:14844512:SG), MGP (chr12:14878220:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG)
HEART_LEFT_VENTRICLE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG, chr12:14899824:SG, chr12:14899901:SG, chr12:14900018:SG, chr12:14901082:SG), WBP11 (chr12:14894016:SG, chr12:14901082:SG)
HEART_RIGHT_ATRIUM	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG), WBP11 (chr12:14844512:SG), MGP (chr12:14847029:SG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14899824:SG, chr12:14899901:SG, chr12:14900018:SG, chr12:14901082:SG)

HEART_RIGHT_VENTRICLE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14894016:SG, chr12:14901082:SG)	
HEPATIC_STELLATE_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
HEPATOCELLULAR_CARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14847029:SG)	
HEPATOCYTE_DERIV	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG)	
HIPPOCAMPUS	ART4 (chr12:14835521:SG), MGP (chr12:14894016:SG)	
INFERIOR_PARIETAL_CORTEX	MGP (chr12:14894016:SG)	
iPSC	MGP (chr12:14854918:IG)	
IRIS_PIGMENT_EPITHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)	
ISLET_PRECURSOR_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
KERATINOCYTE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14897475:SG, chr12:14897803:SG)	
KIDNEY	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG, chr12:14901082:SG), WBP11 (chr12:14894016:SG)	
KIDNEY_CAPILLARY_ENDOTHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
KIDNEY_CELL	WBP11 (chr12:14835521:SG), ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
KIDNEY_CLEAR_CELL_CARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
KIDNEY_EPITHELIAL_CELL	ART4 (chr12:14835521:SG)	
KIDNEY_RHABOID_TUMOR	MGP (chr12:14897475:SG, chr12:14897803:SG)	
LARGE_CELL_LUNG_CANCER	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
LARGE_INTESTINE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG), H2AJ (chr12:14836364:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG, chr12:14894016:SG), WBP11 (chr12:14844512:SG, chr12:14894016:SG)	
LEG_MUSCLE	ART4 (chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG), WBP11 (chr12:14894016:SG)	
LIMB_EMBRYO	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG), WBP11 (chr12:14894016:SG)	
LIVER	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG), H2AJ (chr12:14836364:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG, chr12:14894016:SG, chr12:14901082:SG), GUCY2C (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), WBP11 (chr12:14844512:SG)	
LUNG	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG), WBP11 (chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG)	
LUNG_ADENOCARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
LUNG_EPITHELIAL_CARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), ERP27 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), WBP11 (chr12:14844512:SG, chr12:14847029:SG)	CGNL1 (chr15:57923529:SG), ALDH1A2 (chr15:57923529:SG), RNF111 (chr15:57923529:SG)
LUNG_FIBROBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG, chr12:14901082:SG), WBP11 (chr12:14894016:SG)	
LUNG_MICROVASCULAR_ENDOTHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)	
LYMPHOBLASTOID_CELL_LINE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14854918:IG)	
LYMPHOCYTE	MGP (chr12:14854918:IG)	
MAMMARY_EPITHELIAL_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
MAMMARY_FIBROBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG, chr12:14901082:SG)	
MAMMARY_GLAND_ADENOCARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14840920:SG, chr12:14854918:IG, chr12:14901082:SG)	
MAMMARY_GLAND_DUCTAL_CARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14840920:SG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14901082:SG)	
MAMMARY_LUMINAL_EPITHELIAL_CELL	ART4 (chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG)	
MAMMARY_MYOEPITHELIAL_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG), GUCY2C (chr12:14911328:SG)	
MAMMARY_STEM_CELL	ART4 (chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG)	
MEDULLA_OBLONGATA	MGP (chr12:14894016:SG)	
MEDULLOBLASTOMA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), ERP27 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
MELANOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14894016:SG), WBP11 (chr12:14894016:SG), GUCY2C (chr12:14911328:SG)	
MESENCHYMAL_STEM_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), H2AJ (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG)	
MESODERMAL_DERIV	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG, chr12:14899824:SG, chr12:14899901:SG, chr12:14900018:SG)	
MIDBRAIN	MGP (chr12:14894016:SG)	
MIDDLE_FRONTAL_AREA	MGP (chr12:14894016:SG)	
MIDDLE_FRONTAL_GYRUS	MGP (chr12:14894016:SG)	
MPP	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG), ATF7IP (chr12:14836364:SG), WBP11 (chr12:14844512:SG, chr12:14901082:SG), MGP (chr12:14844512:SG, chr12:14847029:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG)	
MUSCLE_EWING_SARCOMA	MGP (chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG)	
MYELOGENOUS_LEUKEMIA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
MYELOMA	ART4 (chr12:14851053:SG, chr12:14851097:IG)	

MYOCYTE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
NEURAL_PROGENITOR_DERIV	ART4 (chr12:14835521:SG), MGP (chr12:14854918:IG, chr12:14894016:SG)
NEUROBLASTOMA	ART4 (chr12:14835521:SG)
NEUROEPITHELIOMA	MGP (chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG)
NEUROGLIOMA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
NEURON_DERIV	ART4 (chr12:14835521:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14901082:SG), WBP11 (chr12:14901082:SG)
NEUROSPHERE	ART4 (chr12:14835521:SG)
NON-PIGMENTED_CILIARY_EPITHELIAL_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)
OCCIPITAL_LOBE	MGP (chr12:14894016:SG)
OLFACTORY_NEUROSPHERE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG), WBP11 (chr12:14894016:SG)
OMENTAL_FAT_PAD	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG), WBP11 (chr12:14894016:SG)
OSTEOBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)
OSTEOSARCOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14897475:SG, chr12:14897803:SG)
OVARY	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14901082:SG), STRAP (chr12:14890950:SG, chr12:14890963:SG), WBP11 (chr12:14894016:SG)
PANCREAS	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14901082:SG)
PANCREATIC_DUCT_EPITHELIAL_CELL	WBP11 (chr12:14835521:SG), ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
PARATHYROID_ADENOMA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)
PERICYTE	MGP (chr12:14854918:IG)
PERIDONTAL_LIGAMENT_FIBROBLAST	MGP (chr12:14854918:IG)
PEYERS_PATCH	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG, chr12:14901082:SG)
PLACENTA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14844512:SG, chr12:14894016:SG, chr12:14901082:SG), WBP11 (chr12:14901082:SG)
PLASMA_CELL_MYELOMA	ART4 (chr12:14851053:SG, chr12:14851097:IG)
PONS	MGP (chr12:14894016:SG)
PROSTATE_ADENOCARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
PROSTATE_EPITHELIAL_CARCINOMA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), SMCO3 (chr12:14840920:SG)
PROSTATE_EPITHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
PROSTATE_GLAND	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)
PROXIMAL_TUBULE_EPITHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
PSOAS_MUSCLE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG, chr12:14901082:SG)
PULMONARY_ARTERY_ENDOTHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
PULMONARY_ARTERY_FIBROBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG), WBP11 (chr12:14894016:SG)
PUTAMEN	MGP (chr12:14894016:SG)
RECTUM_MUCOSA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG), GUCY2C (chr12:14911328:SG)
RECTUM_MUSCLE	MGP (chr12:14894016:SG, chr12:14901082:SG)
RENAL_CELL_ADENOCARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
RENAL_CELL_CARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
RENAL_CORTEX_INTERSTITIUM	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG), WBP11 (chr12:14894016:SG)
RENAL_CORTICAL_EPITHELIAL_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), GUCY2C (chr12:14911328:SG)
RENAL_PELVIS	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG), WBP11 (chr12:14894016:SG)
SIGMOID_COLON	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG), GUCY2C (chr12:14911328:SG)
SKELETAL_MUSCLE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG)
SKELETAL_MUSCLE_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
SKELETAL_MUSCLE_MYOBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
SKELETAL_MUSCLE_SATELLITE_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG)
SKIN_FIBROBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), WBP11 (chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14897475:SG, chr12:14897803:SG)
SKIN_LEG	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)
SKIN_OF_BODY	MGP (chr12:14847029:SG, chr12:14894016:SG)
SMALL_INTESTINE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)
SMOOTH_MUSCLE_DERIV	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14899824:SG, chr12:14899901:SG, chr12:14900018:SG)
SPINAL_CORD	ART4 (chr12:14835521:SG), MGP (chr12:14854918:IG, chr12:14894016:SG), WBP11 (chr12:14894016:SG)
SPLEEN	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG, chr12:14894016:SG), ERP27 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), H2AJ (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), WBP11 (chr12:14844512:SG)
SQUAMOUS_CELL_CARCINOMA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)
STOMACH	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG), WBP11 (chr12:14894016:SG)

STOMACH_MUSCLE	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG), WBP11 (chr12:14894016:SG)	
SUBSTANTIA_NIGRA	ART4 (chr12:14835521:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG)	
SUPERIOR_TEMPORAL_GYRUS	MGP (chr12:14894016:SG)	
T17_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
TEMPORAL_LOBE	ART4 (chr12:14835521:SG), MGP (chr12:14894016:SG)	
TESTIS	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14847029:SG, chr12:14894016:SG), WBP11 (chr12:14894016:SG)	
THORACIC_AORTA	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14844512:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG, chr12:14901082:SG)	
THYROID_GLAND	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), H2AJ (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14844512:SG, chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG)	
TIBIAL_ARTERY	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG)	
TIBIAL_NERVE	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14901082:SG), WBP11 (chr12:14901082:SG)	
TONGUE	MGP (chr12:14854918:IG, chr12:14894016:SG)	
TRANSVERSE_COLON	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG), WBP11 (chr12:14894016:SG)	
TROPHOBLAST	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14854918:IG, chr12:14901082:SG)	
TROPHOBLAST_DERIV	ART4 (chr12:14835521:SG, chr12:14851053:SG, chr12:14851097:IG)	
TRUNK_MUSCLE	MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG), ART4 (chr12:14847029:SG)	
TUBULE_CELL	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
UMBILICAL_CORD	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG), MGP (chr12:14894016:SG, chr12:14897475:SG, chr12:14897803:SG)	
UMBILICAL_VEIN_ENDOTHELIAL_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14854918:IG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14894016:SG), WBP11 (chr12:14847029:SG, chr12:14894016:SG)	
URINARY_BLADDER	ART4 (chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG, chr12:14901082:SG), WBP11 (chr12:14894016:SG)	
UROTHELIUM_CELL	ART4 (chr12:14835521:SG, chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG)	
UTERUS	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14847029:SG), MGP (chr12:14847029:SG, chr12:14894016:SG, chr12:14901082:SG)	ALDH1A2 (chr15:58061348:SG)
VAGINA	ART4 (chr12:14840136:SG, chr12:14840214:SG, chr12:14840505:SG, chr12:14840674:SG, chr12:14840920:SG, chr12:14851053:SG, chr12:14851097:IG), MGP (chr12:14847029:SG, chr12:14879684:SG, chr12:14879827:SG, chr12:14879925:IG, chr12:14879926:SG, chr12:14890950:SG, chr12:14890963:SG, chr12:14894016:SG, chr12:14901082:SG)	
VILLOUS_MESENCHYME_FIBROBLAST	MGP (chr12:14854918:IG)	

The variants are shown by their position in Build38, with SG ending for SNPs and IG for indels

Supplementary Table 9. Enrichment-Analysis: EHOA association signals are nominally enriched within regulatory regions specific for vascular/endothelial cell types.

Annotation	Number of overlapping GWAS loci	*P* -value	Expected intersection (95%CI)	Observed intersection (95%CI)	Enrichment (95%CI)
Vascular-endothelial	4	0.011	0.35 (0-0.75)	1 (1.00-1.00)	2.84 (2.84-2.84)
Pulmonary development	3	0.06	0.31 (0-0.75)	0.75 (0.25-1)	2.44 (0.813-3.25)
Musculoskeletal	3	0.27	0.50 (0-1)	0.75 (0.25-1)	1.52 (0.505-2.02)
Digestive	3	0.29	0.51 (0.25-1)	0.75 (0.369-1)	1.47 (0.723-1.96)
Myeloid-erythroid	3	0.41	0.57 (0-1)	0.75 (0.369-1)	1.32 (0.65-1.76)
StromalA	1	0.59	0.21 (0-0.5)	0.25 (0-0.75)	1.21 (0-3.61)
Renal-cancer	2	0.69	0.47 (0-0.75)	0.5 (0-1)	1.06 (0-2.12)
Organ development-renal	2	0.72	0.50 (0-1)	0.5 (0-1)	0.99 (0-1.99)
StromalB	3	0.76	0.73 (0.25-1)	0.75 (0.25-1)	1.03 (0.342-1.37)
Lymphoid	2	0.52	0.60 (0.25-1)	0.5 (0-1)	0.83 (0-1.66)
Primitive-embryonic	3	0.55	0.82 (0.5-1)	0.75 (0.25-1)	0.91 (0.304-1.22)
Placental-trophoblast	2	0.48	0.63 (0.25-1)	0.5 (0-1)	0.80 (0-1.6)
Cardiac	1	0.51	0.38 (0-0.75)	0.25 (0-0.75)	0.66 (0-1.99)
Cancer-epithelial	1	0.47	0.43 (0-1)	0.25 (0-0.75)	0.58 (0-1.75)
Neural	2	0.30	0.71 (0.25-1)	0.5 (0-1)	0.70 (0-1.41)
Tissue-invariant	0	0.034	0.52 (0-1)	0 (0-0)	0 (0-0)

Supplementary Table 10. Association of EHOA variants and correlated GWS variants in public datasets

Locus	Variant	EHOA / Correlated GWS	EA	NEA	LD (r2)	P-value	Beta	Odds Ratio	Trait	PMID	N Cases	N Overall	Study ID
4q22.1-MEPE	rs17013495	EHOA	T	C		8.5E-21	0.018		Urate			411,640	UKBio_deCODE
4q22.1-MEPE	rs17013495	EHOA	T	C		1.7E-07		1.06	Gout		16,353	431,047	UKBio_deCODE
4q22.1-MEPE	rs17013495	EHOA	T	C		1.8E-07		0.93	Plantar_fascial_fibromatosis		12,959	431,047	UKBio_deCODE
4q22.1-MEPE	rs17013495	EHOA	T	C		4.8E-07	0.038		Pelvis_DXA_BMD			35,596	UKBio_deCODE
4q22.1-MEPE	rs17013495	EHOA	T	C		5.2E-07	-0.013		Alkaline phosphatase			412,141	UKBio_deCODE
4q22.1-MEPE	rs17013495	EHOA	T	C		5.3E-07	-0.090		Hand grip strength (left)			359,704	NEALE2_46_raw
4q22.1-MEPE	rs17013495	EHOA	T	C		8.1E-07	0.037		Trunk_DXA_BMD			35,596	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		5.0E-35	-0.029		Mean_grip_strength			427,745	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		8.6E-33	-0.219		Hand grip strength (left)			359,704	NEALE2_46_raw
12p12.3-MGP	rs1800801	EHOA	T	C		8.2E-27	-0.197		Hand grip strength (right)			359,729	NEALE2_47_raw
12p12.3-MGP	rs1800801	EHOA	T	C		1.1E-14	-0.015		Heel bone mineral density	PMID:30598549		426,824	GCST006979
12p12.3-MGP	rs1800801	EHOA	T	C		6.4E-11		1.05	Any fracture over 40 years		55,982	431,047	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		2.2E-10	-0.016		Speed_of_sound_through_heel_SOS			399,133	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		1.4E-09	-0.015		Heel bone mineral density			398,823	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		2.9E-09		1.03	Any fracture		103,590	431,047	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		4.5E-09	-0.015		Heel_bone_ultrasound_T_score			401,039	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		1.5E-08		1.04	Low hand grip strength (60 years and older) (EWGSOP)	PMID:33510174	48,596	256,523	GCST90007526
12p12.3-MGP	rs1800801	EHOA	T	C		7.2E-08	-0.041		DXA_Arms_BMD			35,597	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		1.7E-07	-0.013		Heel_broadband_ultrasound_attenuation_BUA			398,131	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		1.9E-07	-0.010		Appendicular lean mass	PMID:33097823		450,243	GCST90000025
12p12.3-MGP	rs1800801	EHOA	T	C		2.3E-07		1.04	Arthrosis_unspecified		46,615	431,047	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		4.8E-07		1.04	Fractured/broken bones in last 5 years		34,780	359,241	NEALE2_2463
12p12.3-MGP	rs1800801	EHOA	T	C		8.9E-07	-0.012		Creatinine			411,927	UKBio_deCODE
12p12.3-MGP	rs1800801	EHOA	T	C		9.3E-07		1.04	Fracture low trauma		35,439	431,047	UKBio_deCODE
12p12.3-MGP	rs3887182*	Correlated GWS	A	G	0.92	3.0E-213	0.721		Blood protein levels [ART4, 6576_1_3]	PMID:30072576		3,200	GCST006585_131
12p12.3-MGP	rs67482087*	Correlated GWS	G	T	0.98	8.0E-182	0.690		Blood protein levels [ART4, 6576_1_3]	PMID:30072576		3,200	GCST006585_131
12p12.3-MGP	rs2287226	Correlated GWS	G	A	0.91	5.3E-33	-0.220		Hand grip strength (left)			359,704	NEALE2_46_raw
12p12.3-MGP	rs11056198	Correlated GWS	A	G	0.91	7.7E-28	-0.201		Hand grip strength (right)			359,729	NEALE2_47_raw
12p12.3-MGP	rs2430689	Correlated GWS	G	C	0.87	3.0E-16	-0.016		Heel bone mineral density	PMID:30598549		426,824	GCST006979
12p12.3-MGP	rs67482087	Correlated GWS	G	T	0.98	2.0E-15	-0.199		Blood protein levels [MGP, 6520_87_3]	PMID:30072576		3,200	GCST006585_1144
12p12.3-MGP	rs4764133	Correlated GWS	T	C	0.97	2.0E-15	0.830		Osteoarthritis of the hand	PMID:28855172		12,754	GCST009596
12p12.3-MGP	rs2430690	Correlated GWS	C	T	0.84	3.0E-15			Heel bone mineral density	PMID:30595370		446,000	GCST007066
12p12.3-MGP	rs2430689	Correlated GWS	G	C	0.87	5.0E-14			Heel bone mineral density	PMID:30048462		394,929	GCST006433
12p12.3-MGP	rs4764133	Correlated GWS	T	C	0.97	5.0E-14	0.650		Finger osteoarthritis severity (hand Klsum)	PMID:33055079		2,994	GCST90010717
12p12.3-MGP	rs2287226	Correlated GWS	G	A	0.91	1.0E-12	-0.002		Hand grip strength	PMID:29691431		334,825	GCST005830
12p12.3-MGP	rs3887182	Correlated GWS	A	G	0.92	2.0E-12	-0.172		Blood protein levels [MGP, 6520_87_3]	PMID:30072576		3,200	GCST006585_1144
12p12.3-MGP	rs11614333	Correlated GWS	T	C	0.91	2.0E-12	-0.160		Hand grip strength	PMID:29313844		195,180	GCST005235
12p12.3-MGP	rs4764133	Correlated GWS	T	C	0.97	3.0E-12	0.810		Hand osteoarthritis severity (hand Klsum)	PMID:33055079		6,032	GCST90010716
12p12.3-MGP	rs10846071	Correlated GWS	T	C	0.92	4.0E-12	-0.002		Hand grip strength	PMID:29691431		334,825	GCST005830
12p12.3-MGP	rs10630224	Correlated GWS	TGC	T	0.76	1.4E-10		1.06	Low hand grip strength (60 years and older) (EWGSOP)	PMID:33510174		256,523	GCST90007526
12p12.3-MGP	rs10630224	Correlated GWS	TGC	T	0.76	3.2E-09		1.07	Low hand grip strength (60 years and older) (EWGSOP)	PMID:33510174		135,468	GCST90007527
12p12.3-MGP	rs11419786	Correlated GWS	TG	T	0.63	1.3E-08		1.05	Fractured/broken bones in last 5 years			359,241	NEALE2_2463
12p12.3-MGP	rs11056244	Correlated GWS	A	T	0.73	3.7E-08	0.524		Impedance of arm (left)			354,807	NEALE2_23110_raw
15q21.3-ALDH1A2	rs11631127	EHOA	C	G		1.2E-07		1.17	Polyarthrosis		2,610	149,831	FINNGEN_R5_M13_ARTHROSIS_POLY
15q21.3-ALDH1A2	rs11631127	EHOA	C	G		1.5E-07	-0.094		Hand grip strength (right)			359,729	NEALE2_47_raw
15q21.3-ALDH1A2	rs11631127	EHOA	C	G		1.9E-07		1.05	Low hand grip strength (60 years and older) (EWGSOP)	PMID:33510174	34,589	135,468	GCST90007527
15q21.3-ALDH1A2	rs11631127	EHOA	C	G		2.1E-07		0.97	Knee pain | pain type(s) experienced in last month		76,628	360,391	NEALE2_6159_7
15q21.3-ALDH1A2	rs11631127	EHOA	C	G		2.2E-07		0.96	Knee_joint_operation		28,317	431,047	UKBio_deCODE
15q21.3-ALDH1A2	rs11631127	EHOA	C	G		2.3E-07		1.04	Low hand grip strength (60 years and older) (EWGSOP)	PMID:33510174	48,596	256,523	GCST90007526
15q21.3-ALDH1A2	rs11631127	EHOA	C	G		3.8E-07		0.96	Knee osteoarthritis		37,270	430,938	UKBio_deCODE
15q21.3-ALDH1A2	rs11631127	EHOA	C	G		3.3E-07	-0.091		Hand grip strength (left)			359,704	NEALE2_46_raw

15q21.3-ALDH1A2	rs11631127	EHOA	C	G		4.7E-07	0.95	Knee osteoarthritis	PMID:30664745	24,955	403,124	GCST007090
15q21.3-ALDH1A2	rs4775006	Correlated GWS	A	C	0.74	3.0E-22		Brain region volumes [X4th ventricle]	PMID:31676860		19,629	GCST009518_4
15q21.3-ALDH1A2	rs4775006	Correlated GWS	A	C	0.74	1.0E-18		Subcortical volume (min-P)	PMID:32665545		26,502	GCST010698
15q21.3-ALDH1A2	rs4775006	Correlated GWS	A	C	0.74	2.0E-18		Brain morphology (min-P)	PMID:32665545		26,502	GCST010699
15q21.3-ALDH1A2	rs3204689	Correlated GWS	C	G	0.65	1.0E-11		Osteoarthritis (hand, severe)	PMID:24728293		78,162	GCST002410
15q21.3-ALDH1A2	rs66725070	Correlated GWS	G	GACAT	0.73	3.0E-10	0.87	Barrett's esophagus	PMID:27527254		23,326	GCST003738
15q21.3-ALDH1A2	rs4775006	Correlated GWS	A	C	0.74	8.4E-10	1.06	Knee osteoarthritis	PMID:30664745		403,124	GCST007090
15q21.3-ALDH1A2	rs8033270	Correlated GWS	C	G	0.64	9.0E-10	0.84	Polyarhtrosis			149,831	FINNGEN_R5_M13_ARTHROSIS_POLY
15q21.3-ALDH1A2	rs4775006	Correlated GWS	A	C	0.74	7.1E-09	1.03	Knee pain | pain type(s) experienced in last month			360,391	NEALE2_6159_7

Association results assessed by UKBiobank associations at deCODE genetics, and by Open Targets Genetics (https://genetics.opentargets.org/) which summarizes association data for the variants in public datasets (UK Biobank, FinnGen, and GWAS Catalog). The site was accessed on Febuary, 23rd, 2022. The look-up results for the EHOA variants are shown directly, and for correlated variants ($r^2 > 0.60$) that have been reported to associate with a given trait at a GWS level in the Open Targets Genetics database. The effect allele, the other allele, r^2 with EHOA variant at the locus, P value and beta or odds ratio (OR) are shown for each trait, along with publication ID, number of cases and/or overal study sample, and the Study ID. Associations with P < 1e-6 are shown.

*We note that this association is most likely due to a missense variant in ART4 which changes the binding of the Somalogic probe to the plasma protein

Supplementary Table 11. Association of EHOA variants with bone density, grip strength and urate levels

				FN_BMD (N = 107,310)		LS_BMD (N = 106,228)		eBMD (N = 398,823)		Grip strength (N = 427,745)		Urate (N = 411,640)	
Variant	**Chr**	**EA**	**NEA**	**P value**	**Effect**	**P value**	**Effect**	**P value**	**Effect**	**P value**	**Effect**	**P value**	**Effect**
rs17013495	chr4	T	C	4.9E-04	0.015	1.6E-09	0.028	0.89	0	1.9E-05	-0.010	8.5E-21	0.018
rs11243284	chr6	C	T	0.082	-0.008	0.0035	-0.015	0.78	-0.001	2.7E-03	-0.007	0.76	0.001
rs1800801	chr12	T	C	1.4E-05	-0.019	8.9E-08	-0.025	1.4E-09	-0.015	5.0E-35	-0.029	0.036	-0.004
rs11631127	chr15	C	G	0.59	-0.002	8.3E-07	0.023	0.21	-0.003	9.8E-10	-0.014	0.95	0

Results for eBMD, grip strength and urate levels are from the UK Biobank resource, run at deCODE genetics. Results for FN (femoral neck) and LS (lumbar spine) BMD are derived from our unpublished meta-analysis of BMD in Iceland, UK Biobank, and the publicly available GEFOS consortium (Zheng et al, Nature, 2015).

Supplementary Table 12. Significant association of EHOA polygenic risk score with phenotypes in UK biobank

Phenotype	P value	Effect /OR	N cases	N controls	N overall	nR2
Grip strength (mean, age, sex, height adj.)	6.0E-41	-0.022			427,745	0.00048
Other arthrosis (ICD10:M19)	4.8E-29	1.05	73,440	357,607		0.00048
Any OA	2.1E-19	1.04	103,173	327,874		0.00029
Polyarthrosis (ICD10:M15)	4.9E-17	1.08	12,326	418,612		0.00072
Hand OA	4.7E-16	1.15	3,416	427,631		0.00175
Pain due to OA	7.5E-14	1.05	44,262	98,258		0.00053
Pain in hands in last three months	2.8E-13	1.11	5,766	64,039		0.00174
Heberden nodes with arthropathy	1.1E-12	1.30	758	428,428		0.00464
Finger OA	1.3E-13	1.29	834	428,352		0.00407
Other arthritis (ICD10:M13)	6.3E-10	1.04	33,303	397,744		0.00021
Operation of joint of finger (OPCS:Z83)	1.3E-09	1.10	4,289	426,758		0.00082

A PRS for EHOA was generated from the Icelandic, the Dutch, the US and Spanish EHOA datasets. The MHC region was excluded from the EHOA PRS. The results are shown from a scan of diverse phenotypes derived from the UK Biobank. Significance was set as $P < 1.0 \times 10^{-5}$, accounting for 5,000 main phenotypes. nR2 is the Nagelkerke's correlation coefficient.

Supplementary Table 13. Association of finger, hand, thumb, knee, hip, spine, and all OA GWS variants from the GO consortium in EHOA meta-analysis

							GO consortium results		EHOA_meta	
OA phenotype	Variant	EA	NEA	EA_freq%	GO_locus_number	Associated GWS OA phenotypes	OR	P	OR	P
FingerOA	rs7294636	A	G	37.3	21	finger	1.16	3.0E-16	1.36	6.8E-13
FingerOA	rs9396861	A	C	61.0	77	hand,finger	1.13	9.3E-11	1.23	3.2E-06
FingerOA	rs11588154	T	G	16.6	7	finger	0.83	6.1E-10	0.86	0.01
FingerOA	rs8031133	T	G	54.5	30	hand, finger, thumb, knee	1.11	1.1E-09	1.40	2.4E-14
FingerOA	rs11550348	A	G	11.0	45	hand,finger	0.84	6.2E-09	0.76	1.9E-04
HandOA	rs11071366	A	T	61.4	30	hand, finger, thumb, knee	0.90	4.9E-17	0.71	7.1E-15
HandOA	rs3993110	A	C	60.8	10	hand	1.09	3.8E-11	1.14	0.0028
HandOA	rs3771498	T	C	52.1	52	hand, thumb, hip, all	0.92	6.8E-11	0.86	2.8E-04
HandOA	rs8112559	C	G	88.6	45	hand,finger	1.13	7.3E-11	1.32	1.8E-04
HandOA	rs10062749	T	G	26.9	73	hand,thumb	1.08	2.0E-09	1.15	0.0029
HandOA	rs7748189	A	G	73.2	77	hand,finger	1.08	6.1E-09	1.24	1.2E-05
HandOA	rs1560080	A	G	82.5	71	hand	0.91	9.6E-09	0.84	0.0011
ThumbOA	rs4238326	T	C	60.7	30	hand, finger, thumb, knee	0.89	7.3E-12	0.71	7.4E-15
ThumbOA	rs2862851	T	C	46.5	52	hand, thumb, hip, all	1.11	3.2E-10	1.16	4.4E-04
ThumbOA	rs11588850	A	G	82.0	6	thumb	0.87	3.5E-10	1.06	0.34
ThumbOA	rs10062749	T	G	26.9	73	hand,thumb	1.11	1.3E-08	1.15	0.0029
SpineOA	rs201194999	T	C	30.1	69	all,spine	0.85	1.2E-08	0.57	0.27
KneeOA	rs143384	A	G	59.1	53	knee,all	1.07	1.0E-23	1.08	0.09
KneeOA	rs9940278	T	C	43.5	35	knee,hip	1.06	3.2E-16	0.97	0.56
KneeOA	rs34195470	A	G	44.5	36	knee	0.95	3.1E-13	0.99	0.85
KneeOA	rs4548913	A	G	62.8	37	knee, all	0.95	3.2E-12	0.92	0.04
KneeOA	rs72760655	A	C	33.1	92	knee,all	1.05	7.3E-11	1.05	0.29
KneeOA	rs7581446	T	C	48.3	50	knee	0.95	1.7E-10	0.98	0.62
KneeOA	rs753350451	D	I	20.2	19	knee	0.93	3.4E-10	0.98	0.77
KneeOA	rs58973023	A	T	48.9	27	knee	1.06	4.7E-10	1.04	0.41
KneeOA	rs4775006	A	C	41.6	30	hand, finger, thumb, knee	1.05	8.5E-10	0.69	3.0E-16
KneeOA	rs4380013	A	G	18.8	29	knee	1.06	8.7E-10	1.05	0.36
KneeOA	rs1426371	A	G	27.1	18	knee	0.95	8.9E-10	1.03	0.51
KneeOA	rs66906321	T	C	17.5	51	knee	0.95	1.7E-09	0.97	0.64
KneeOA	rs7967762	T	C	15.6	24	knee	1.06	2.1E-09	1.08	0.19
KneeOA	rs72979233	A	G	75.3	15	knee	0.95	2.5E-09	0.92	0.07
KneeOA	rs2163832	T	C	32.1	43	knee,all	1.05	2.7E-09	1.23	6.2E-06
KneeOA	rs11705555	A	C	76.4	56	knee	1.05	3.0E-09	1.00	0.97
KneeOA	rs2791549	A	C	29.6	5	knee, hip	1.05	3.1E-09	1.05	0.26
KneeOA	rs10842226	A	G	42.0	22	knee	1.05	3.6E-09	0.91	0.03
KneeOA	rs10974438	A	C	64.6	99	knee	1.04	4.9E-09	1.07	0.14
KneeOA	rs6500609	C	G	11.0	34	knee	0.94	5.2E-09	0.97	0.67
KneeOA	rs10038860	A	G	27.4	73	knee	1.05	5.6E-09	1.16	0.0013
KneeOA	rs12914479	C	G	66.0	33	knee	1.04	7.1E-09	1.07	0.13
KneeOA	rs2066928	A	G	48.3	75	knee	0.96	1.2E-08	0.97	0.47
KneeOA	rs7680647	T	C	63.1	68	knee,hip,all	0.96	1.2E-08	0.85	3.2E-04
HipOA	rs10843013	A	C	78.4	23	hip	0.90	2.9E-24	0.96	0.46
HipOA	rs12209223	A	C	11.1	83	hip	1.15	1.9E-22	1.16	0.03
HipOA	rs11164653	T	C	41.3	1	hip,all	0.92	2.8E-18	1.01	0.84
HipOA	rs12908498	C	G	53.8	32	hip	1.08	1.9E-16	1.06	0.18
HipOA	rs2416564	T	C	59.8	95	hip	0.93	1.0E-15	1.06	0.18
HipOA	rs765002298	D	I	19.7	90	hip	0.90	1.8E-15	1.03	0.58
HipOA	rs4252548	T	C	2.4	46	hip	1.25	2.2E-15	1.36	0.03
HipOA	rs1046934	A	C	64.9	4	hip	1.07	3.8E-14	0.99	0.80
HipOA	rs79895530	T	C	13.0	91	hip	0.90	7.0E-14	1.05	0.42
HipOA	rs2268023	A	T	41.1	63	hip	1.07	1.6E-13	0.95	0.21
HipOA	rs1913707	A	G	60.5	66	hip,all	1.07	1.8E-13	1.06	0.16
HipOA	rs2862851	T	C	46.5	52	hand, thumb, hip, all	1.07	3.9E-13	1.16	4.4E-04
HipOA	rs9475400	T	C	9.8	82	hip	1.11	8.0E-13	1.04	0.56
HipOA	rs111844273	A	G	2.1	89	hip	1.26	1.0E-12	1.10	0.49
HipOA	rs6908606	A	G	71.1	81	hip	0.93	3.9E-12	0.91	0.06
HipOA	rs2605098	A	G	33.2	5	hip	1.07	6.8E-12	1.09	0.05
HipOA	rs1330349	C	G	58.9	93	hip	1.06	6.9E-12	1.02	0.59
HipOA	rs4411121	T	C	31.4	2	hip	1.07	2.2E-11	1.06	0.22
HipOA	rs12377624	C	G	36.2	97	hip	0.94	4.6E-11	1.02	0.66
HipOA	rs143083812	T	C	0.11	86	hip	2.90	8.2E-11	3.48	0.0087
HipOA	rs1401796	A	C	51.3	39	hip,all	0.94	1.4E-10	0.97	0.52
HipOA	rs746239049	D	I	20.5	31	hip	0.92	3.3E-10	0.89	0.05
HipOA	rs12160491	A	G	71.1	57	hip	0.94	4.4E-10	1.01	0.81
HipOA	rs67924081	A	G	73.9	13	hip	1.07	7.8E-10	1.02	0.71
HipOA	rs10831477	T	G	81.1	17	hip,all	1.07	1.2E-09	1.01	0.81
HipOA	rs9835230	A	G	24.3	61	hip	1.07	1.3E-09	1.11	0.04
HipOA	rs2521348	T	C	38.7	41	hip	1.06	1.6E-09	0.94	0.14
HipOA	rs34560402	T	C	6.5	14	hip	0.89	1.6E-09	0.96	0.63
HipOA	rs9940278	T	C	43.5	35	knee,hip	1.06	1.8E-09	0.97	0.56
HipOA	rs3740129	A	G	45.9	8	hip	1.06	1.8E-09	1.05	0.25
HipOA	rs79056043	A	G	93.7	25	hip	0.89	2.0E-09	0.84	0.05
HipOA	rs79220007	T	C	92.7	78	hip	0.90	2.2E-09	1.01	0.91

HipOA	rs798756	T	C	19.4	68	knee,hip,all	0.93	2.2E-09	0.85	0.0044
HipOA	rs4073717	T	G	20.1	74	hip	0.94	2.5E-09	0.90	0.05
HipOA	rs17677724	T	C	16.1	72	hip,all	1.07	3.5E-09	1.12	0.05
HipOA	rs1809889	T	C	28.0	20	hip	1.06	3.6E-09	1.07	0.17
HipOA	rs10983775	T	C	54.2	96	hip	0.95	4.7E-09	0.99	0.80
HipOA	rs66989638	A	G	12.7	48	hip	1.08	4.8E-09	1.00	0.97
HipOA	rs7862601	A	G	62.4	94	hip	0.94	6.2E-09	0.94	0.15
HipOA	rs7222178	A	T	19.5	40	hip	1.07	7.4E-09	1.01	0.82
HipOA	rs10940168	A	G	39.4	76	hip	0.95	7.7E-09	1.01	0.76
HipOA	rs6855246	A	G	92.8	64	hip,all	0.90	7.9E-09	1.08	0.53
HipOA	rs10465114	A	G	22.0	98	hip	1.06	9.0E-09	1.01	0.89
AllOA	rs13107325	T	C	7.1	64	hip,all	1.08	3.2E-17	0.94	0.65
AllOA	rs3771501	A	G	46.8	52	hand, thumb, hip, all	1.04	4.0E-15	1.16	6.4E-04
AllOA	rs1913707	A	G	60.5	66	hip,all	1.03	1.4E-12	1.06	0.16
AllOA	rs2425061	A	G	62.8	53	knee,all	1.03	2.1E-12	0.94	0.14
AllOA	rs216175	A	C	82.8	37	all	1.04	2.7E-12	1.09	0.11
AllOA	rs2622873	T	C	88.0	1	hip,all	1.05	4.2E-11	1.03	0.66
AllOA	rs10405617	A	G	31.9	43	knee,all	1.03	9.3E-11	1.22	1.5E-05
AllOA	rs12901372	C	G	52.7	32	all	1.03	1.0E-10	1.07	0.10
AllOA	rs11731421	A	G	34.6	68	knee,hip,all	1.03	1.9E-10	1.19	1.0E-04
AllOA	rs75621460	A	G	2.6	44	all	1.10	1.1E-09	1.06	0.65
AllOA	rs4979341	T	C	27.5	92	knee,all	1.03	1.4E-09	1.04	0.41
AllOA	rs12667224	A	G	52.0	85	all	0.97	1.7E-09	0.96	0.30
AllOA	rs62242105	A	G	33.1	62	hip	0.97	2.9E-09	1.01	0.79
AllOA	rs201194999	T	C	30.1	69	all,spine	0.88	3.1E-09	0.57	0.27
AllOA	rs62182810	A	G	54.4	49	all	1.03	3.8E-09	1.07	0.15
AllOA	rs11729628	T	G	23.9	65	all	0.97	4.7E-09	0.91	0.05
AllOA	rs1401795	A	G	50.0	39	all	1.03	6.2E-09	1.03	0.47
AllOA	rs10831476	A	C	81.1	17	hip,all	1.03	7.8E-09	1.01	0.82
AllOA	rs17677555	C	G	25.6	72	hip,all	1.03	1.1E-08	1.08	0.13

The Genetics of Osteoarthritis (GO) consortium data is from Boer et al, Cell, 2021. OR (odds ratio) and P values and ORs are shown for the respective osteoarthritis (OA) phenotypes in the GO consortium data. The OA phenotypes that are significantly associated with the respective signal in GO are listed under the column "Associated GWS OA phenotypes" (often represented by a different, but highly correlated, variant).

SUPPLEMENTARY MATERIAL: Meta-analysis of erosive hand osteoarthritis

SUPPLEMENTARY METHODS

Study populations

Iceland**:** EHOA (918 cases) was diagnosed from conventional dorsopalmar radiographs taken of individuals who had been diagnosed with hand OA and compared to 109,249 controls. The proximal and distal interphalangeal joints were scored according to Verbruggen-Veys (VV) (1) and patients with at least 1 joint in the E phase (erosive) or R phase (remodelled) were classified as having EHOA. The number of erosive joints per individual was recorded. All radiographs were scored by the same clinician, a co-author of this paper (HJ). Individuals diagnosed with rheumatoid arthritis (RA) were excluded.

Any type of OA was excluded from the controls (ICD10 codes: M15, M16, M17, M18, M19, or M47, ICD9 code 715, and subcodes). The information was derived from Landspitali University Hospital electronic health records, from The Directorate of Health electronic health records, clinicians, and from a national Icelandic hip or knee arthroplasty registry.

All participants who donated samples gave informed consent and the National Bioethics Committee of Iceland approved the study (VSN_14-148, VSN_14-015v8) which was conducted in agreement with conditions issued by the Data Protection Authority of Iceland.

The Netherlands**:** The Dutch samples were derived from two studies: the patients from the Hand OSTeoArthritis in Secondary care (HOSTAS) study (2) , and the controls from the Nijmegen Biomedical Study study (NBS) (3). EHOA cases were scored according to Verbruggen-Veys (VV) (1) and defined as EHOA cases, same as in Iceland. Hostas is an observational cohort with consecutive patients with hand OA diagnosed at a rheumatology outpatient clinic by their treating rheumatologist. Patients with secondary OA or inflammatory joint diseases, such as rheumatoid arthritis, or other conditions that could explain their hand symptoms were excluded. Dorsovolar hand radiographs were scored by one reader, with good reliability (for details see ref Damman *et al* (2). Both cases (N=139) and controls (N=5,102) were genotyped on the same Illumina chip type. Individuals from the NBS were invited to participate in a study on gene-environment interactions in multifactorial diseases. The details of this study were reported previously (3). The study protocol of the Nijmegen Biomedical Study was approved by the Institutional Review Board of the Radboud University Medical Center and all study subjects gave written informed consent. All individuals included in this study were genetically determined to be of European descent.

United Kingdom: The UK Biobank resource (http://www.ukbiobank.ac.uk) includes data from 500,000 volunteer participants who were recruited between the age of 40-69 years in 2006-2010 across the United Kingdom. All individuals in the current study (63 EHOA cases/430,875 controls) were of White British descent. The EHOA included those with the ICD10 code M15.4. All participants gave informed consent and UK Biobank’s scientific protocol and operational procedures were reviewed and approved by the North West Research Ethics Committee. This research has been conducted using the UK Biobank Resource under Application Number 23359.

United States: The Utah EHOA cases (N=145) have been previously described in Kazmers et al (4). Individuals with the ICD-10 code M15.4 in the Utah Population Database between October 1, 2015 and December 31, 2019 were included, excluding those with rheumatoid arthritis (ICD-9 714.0, ICD-10 M05), other rheumatoid arthritis subtypes (ICD-9 714.2, ICD-10 M06), or juvenile rheumatoid arthritis (ICD-9 714.3, ICD-10 M08). Manual chart review was performed to confirm the EHOA diagnosis. Additional individuals were identified by querying those enrolled in the Intermountain Healthcare HerediGene: Population Study using the ICD-10 code M15.4 and excluding individuals with rheumatoid arthritis. Subjects (male and female, ≥ 18 years of age, and a United States resident) visiting an Intermountain Healthcare facility or event were recruited for study participation (Utah, USA). Subjects were informed of the study protocol and procedures prior to providing consent. A consent waiver was granted for the use of residual blood that would otherwise be discarded following a standard of care blood draw performed before a subject expired. Control subjects (N=5,308) were from the Intermountain Healthcare study, excluding those with any OA. All individuals included in this study were genetically determined to be of European descent. Study procedures were in accordance with the ethical standards of the responsible institution and approved by the Institutional Review Board at the University of Utah (IRB#: 79442, Salt Lake City, UT USA) and Intermountain Healthcare (IRB#: 1051071, Salt Lake City, UT USA).

Spain: The Spanish samples are all from A Coruña. The cases (N=218) were derived from the PROCOAC (PROspective COhort of A Coruña) cohort (5), and the controls (N=164) were from other projects at A Coruña University Hospital who had not been diagnosed with hand OA on radiographs. EHOA cases were scored according to Verbruggen-Veys (VV) (1). All individuals included in this study were genetically determined to be of European descent.

We applied ancestry analysis to the UK, US, Spanish and Dutch cohorts and excluded samples that were identified as ethnic outliers (see below). For the remaining samples we constructed genetic principal components that were used as covariates in the association analysis to adjust for remaining population substructure. Related individuals are included in the analysis and any inflation this leads to in the test statistics is adjusted for using a genomic control adjustment.

Genotyping and imputation: The Icelandic samples, the Dutch, the US, and the Spanish samples, were genotyped by deCODE genetics, using Illumina HumanHap and HumanOmni genotyping chips for the Icelandic samples, HumanOmni-1 Quad chip for the Dutch samples, and Illumina GSA chip for the Spanish and US samples. For each sample set, variants were excluded if they (i) had<98% yield, (ii) had<1% MAF, (iii) failed Hardy-Weinberg test ($P<1 \times 10^{-6}$) or (iv) showed significant ($P<1 \times 10^{-6}$) difference between genotype batches. Samples with<96% yield were excluded. The UK Biobank genotyping was performed using a custom-made Affimetrix chip, UK BiLEVE Axiom (6), in the first 50,000 participants, and with Affimetrix UK Biobank Axiom array in the remaining participants (7).

In the Icelandic samples, variants were derived from whole genome sequencing (WGS) 49,962 Icelanders using GAIIx, HiSeq, HiSeqX, and NovaSeq Illumina technology (8), (9), the genotypes of SNPs and indels called jointly by Graphtyper (10), haplotyped long range phased (11) and high-quality sequence variants imputed into all samples. All variants tested had imputation information over 0.8.

The samples from the Netherlands and Spain, and the erosive samples from the US, phased using SHAPEIT (12) and used to impute un-genotyped variants using IMPUTE2 (13). The samples were imputed using the 1000 Genomes Phase 3 reference data (October 2014 release) that includes phased genotypes for about 80 million variants and for 2,504 individuals of various ethnicities (14).

The variants in the US hand, finger, and thumb samples were derived from sequencing 9,268 individuals of non-Icelandic northern European descent, 245 million variants in total, long range phased using SHAPEIT4 (15) and imputed into the US chip data.

The variants imputed into the UK Biobank samples were derived from WGS of 131,958 UK individuals, performed jointly by deCODE genetics and the Welcome Trust Sanger Institute (16) where over 245 million high-quality sequence variants and indels were identified using Graphtyper (10). Quality-controlled chip genotype data were phased using SHAPEIT 4 (15). A phased haplotype reference panel was prepared from the sequence variants using the long-range phased chip-genotyped samples using inhouse tools and methods described previously (8, 9) and imputed into the phases genotype data.

Ancestry analysis: For UK Biobank, we used a British-Irish ancestry subset defined previously (16). It was defined by applying uniform manifold approximation and projection (UMAP) dimension reduction of 40 genetic principal components provided by the UK Biobank and ADMIXTURE analysis supervised on five reference populations and self-reported ethnicity information and defined three cohorts in the UK Biobank data; British-Irish, South-Asian and African ancestry. For the current study we used only data from the British-Irish ancestry group (N = 431,805). For this group 20 principal components were calculated as and included in the association analysis to adjust for remaining population structure.

To study the population structure and the ancestry of samples in the Dutch, Spanish and US cohorts we used the ADMIXTURE (v 1.2) (17) and EIGENSOFT (v 6.0.1) (18) software. Samples were excluded if they were identified as ethnic outliers in the respective cohort, and to adjust for remaining population substructure ten principal components were included as covariates in the subsequent association analysis.

Association analysis: Logistic regression was used to test for association between variants and disease, assuming a multiplicative model, treating disease status as the response and expected genotype counts from imputation as covariates. Testing was performed using the likelihood ratio statistic. For the Icelandic and UK cohorts this was done using software developed at deCODE genetics (8). For Iceland we included county of birth, age, age squared, sex and an indicator function for the overlap of the lifetime of the individual with the time span of phenotype collection as covariates to account for differences between cases and controls. We used county of birth as a proxy covariate for the first principal components (PCs) in our analysis because county of birth has been shown to be in concordance with the first PC in Iceland (19).

The UK association was adjusted for sex, age and the 20 PCs.

The US, Dutch and Spanish associations were analysed using the SNPTEST (v.2.5) software (20), including age, sex and 20 PC‘s as covariates.

We used LD score regression (21) to account for distribution inflation due to cryptic relatedness and population stratification in each of the cohorts respectively.

For genome-wide significance thresholds we used the weighted Holm-Bonferroni method to allocate familywise error rate of 0.05 equally between five annotation-based classes of sequence variants (22); $P \leq 2.4 \times 10^{-7}$ for high-impact variants (including stop-gained and loss, frameshift, splice acceptor or donor and initiator codon variants), $P \leq 4.9 \times 10^{-8}$ for missense, splice-region variants and in-frame-indels, $P \leq 4.4 \times 10^{-9}$ for low-impact variants (including synonymous, 3’ and 5’ UTR, and upstream and downstream variants), $P \leq 2.2 \times 10^{-9}$ for deep intronic and intergenic variants in DNase I hypersensitivity sites (DHS), and $P \leq 7.4 \times 10^{-10}$ for other non-DHS deep intronic and intergenic variants.

Polygenic risk score (PRS) and phenotype correlation analysis: We used PRS analysis based on a EHOA meta-analysis of Icelandic, Dutch, Spanish and US GWASs to investigate its correlation with about 5,000 quantitative and case/control traits in the UK Biobank dataset. The PRSs was calculated using genotypes for about 600,000 autosomal markers included on the Illumina SNP chips to avoid uncertainty due to imputation quality (23). We estimated linkage disequilibrium (LD) between markers using 14,938 phased Icelandic samples and used this LD information to calculate adjusted effect estimates using LDpred (24). The adjusted effects were used as weights to generate the weighted PRS for testing in the UK. We created several PRSs assuming different fractions of causal markers (the P parameter in LDpred). Subsequently, we selected the PRS that was the most predictive of erosive hand OA in UK Biobank data to test for correlation with other traits. The model selected corresponds to assuming that 0.3% of the markers are causal, and this explains 0.4% (P = 0.02) of the variance in the correlation with erosive hand OA based on an Nagelkerke pseudo R^2 estimate. The correlation between the outcome phenotypes and the PRS was done in the same way as for the correlation with genetic variants and using the same software developed at deCODE genetics. For case/control outcome we used logistic regression to test for association between variants and disease treating disease status as the response and the PRS as covariate. Testing was performed using the likelihood ratio statistic and the analysis was adjusted for sex, age and 20 PC's. For quantitative outcome traits we used logistic regression with the PRS as covariate. Prior to association analysis of quantitative traits, measurements were adjusted for sex, age, year of birth, measurement site and population structure. Average of multiple measurements for an individual was used, and the measurements were normalized to a standard normal distribution using quantile normalization. In both cases likelihood ratio test was used to calculate the P-values, and the P values were adjusted for distribution inflation due to cryptic relatedness and population stratification using LD score regression and association results for about 1.2 million unlinked genetic variants. We have now added this description to the methods section. Accounting for 5,000 main phenotypes in the PRS scan, which included all main disease-categories and measured quantitative traits, we set the significance threshold at $P < 1.0 \times 10^{-5}$.

Additional phenotypes: The quantitative phenotypes in UK Biobank were adjusted for covariates for each sex separately and only included individuals of a British-Irish ancestry. For grip strength we used the mean of right and left measures ($N_{grip\_strength}$ = 427,745), adjusted for age and height, and the urate (N urate = 411,640) and BMD measures were adjusted for age and BMI. We downloaded summary statistics from a meta-analysis of lumbar spine (LS) BMD and femoral-neck (FN) BMD from the GEFOS consortium that did not include Icelandic data (25), and meta-analysed with the summary statistics from Iceland and UK Biobank ($N_{LS\text{-}BMD}$ = 106,228, $N_{FN\text{-}BMD}$ = 107,310). eBMD was estimated from heel ultrasound measures as described in Morris et al (26) (N_{eBMD} = 398,823). Osteoporosis was defined by ICD10 codes M80 and M81 ($N_{osteoporosis}$ = 6,626).

For the genetic correlation analysis, we used meta-analyses of rheumatoid arthritis (RA) overall ($N_{RA\_overall}$ = 27,700), sero-positive RA ($N_{RA\ sero\text{-}positive}$ = 16,273), sero-negative RA ($N_{RA\ sero\text{-}negative}$ = 7,446) in North-western European populations (27), excluding the Icelandic data, since Iceland had the largest EHOA sample-set, and gout from UK Biobank, captured both by ICD10 codes M10.0 and M10.9 and by gout-specific drugs (allopurinol, febuxostat, or probenecid) (N_{gout}=15,806).

Functional annotation of sequence variants: We downloaded the cell type agnostic definition of candidate cis-regulatory elements (cCRE) from the ENCODE project (28) (screen.encodeproject.org) and tissue specific regulatory elements from Meuleman et al (zenodo.org/record/3838751#.YYUyjhrP2Ul) (29). We then determined whether the lead sequence variant or any of their correlated variants (r^2 > 0.80) are located within cCRE or tissue specific regulatory regions. We looked for association signals in enhancer elements defined in EpiMap (compbio.mit.edu/epimap) to then see if those same enhancers are predicted to influence nearby genes based on per-sample analysis datasets: personal.broadinstitute.org/cboix/epimap/links/links_corr_only.

Enrichment of association signals in functional annotations: We determined how many of the four association signals identified for EHOA intersect with one of sixteen tissue specific regulatory regions defined in Meuleman et al. (29). Here, we define an association signal as a lead sequence variant along with other sequence variants found in strong correlation (linkage disequilibrium; LD) to the lead variant; r^2>0.80. We refer to this intersection as the „observed intersection“. To find the „expected intersection“, we made use of association signals from the GWAS catalogue (see details in next paraphrase). We binned the signals according to LD class, i.e., the number of correlated variants for each lead association signal in the GWAS catalogue. We then selected, at random, one „lead variant“ from the GWAS catalogue for each of the four EHOA association loci, but ensure that they are selected from the same LD class bins as the observed association signals are found in. LD class bins: 1-10, 11-20, 21-50, 51-100, 101-200, 201-Inf. We then obtain the fraction of overlap to the tissue specific regulatory regions for these four randomly selected and LD class matched loci. This is the „expected intersection“, and, we record whether or not the expected intersection is larger or equal to the observed intersection. We then repeat this process 5,000 times to obtain the mean and confidence intervals for the expected intersection and, importantly, the number of times we see the expected intersections to be higher than or equal to the observed intersection gives the P-value. The enrichment estimates are obtained by computing: observed intersection / mean of expected intersections.

We compiled a robust set of association signals from the NHGRI-EBI catalogue of GWAS association signals; downloaded on 4-AUG-2021 (GWAS catalogue v1.00; www.ebi.ac.uk/gwas). For each disease (or other traits) we selected associations where P-value < 1e-9 and, for each chromosome, we ordered the associations according to P-value to then select the strongest association on each chromosome. We then select the „second strongest“ association on the same chromosome only if it is located more than 1Mb away from the strongest association. This same process was then continued down the list of remaining associations; only those located more than 1Mb away from the stronger associations were selected. Further, as our enrichment algorithm takes LD into account, which we compute in 28,075 whole genome sequenced individuals from the Icelandic population, we selected GWAS´s carried out in individuals of European descent. Finally, we deleted 240 trait association signals as the lead variant of these signals was somewhat correlated (r^2>0.20) to a stronger lead variant on that same chromosome for the same disease/trait. This resulted in 42.669 association signals in 1.875 diseases or other human traits. It is this large set of trait associations that enables us to estimate the expected fraction of association signals intersecting with a given genome annotation.

Co-localisation: To test for co-localization of the EHOA signals with signals in other traits we used the COLOC software package implemented in R (30). Using summary statistics for traits A and B, i.e., effects and P-values, we calculated Bayes factors for each of the variants in the associated region tor the two traits and used COLOC to calculate posterior probability for two hypotheses: (1) that the association with trait A and trait B are independent signals (PP3) and (2) that the association with trait A and trait B are due to a shared signal (PP4).

Zebrafish experiments:

Zebrafish: Danio rerio were maintained in accordance with approved institutional protocols at the University of Utah. Adult zebrafish were maintained under standard conditions and kept on a light-dark cycle of 14 hours in light and 10 hours in dark at 27°C. The Tu strain was used in all experiments.

Bmp6 Mutant Zebrafish Generation: Mutations were induced with CRISPR/Cas9 reagents as described in Hoshijima *et al* (31). gRNA target sequences are as follows: bmp6_gRNA1 (in exon 5) – TTTCAGAGAATTGAGCTGGC(AGG) and bmp6_gRNA2 (in exon 7) – AGTAGAGCACGGAGATTGCG(TGG) (Figure S1a). The PAM sequence is indicated in parentheses. Target-specific Alt-R® crRNA and common Alt-R® tracrRNA were synthesized by IDT and dissolved in duplex buffer (IDT) as a 100µM stock solution. Equal volumes of the Alt-R® crRNA and Alt-R® tracrRNA stock solutions were mixed together and annealed in a PCR machine using the following settings: 95°C, 5 min; cool at 0.1°C/sec to 25°C; 25°C, 5 min; 4°C. Cas9 protein (Alt-R® S.p. Cas9 nuclease, V3, IDT, dissolved in 20mM HEPES-NaOH (pH 7.5), 350mM KCl, 20% glycerol) and crRNA:tracrRNA duplex mixed to generate a 5µM gRNA:Cas9 RNP complex (referred to as RNPs). Prior to microinjection, the RNP complex solution was incubated at 37°C, 5 min and then placed at room temperature. Approximately one nanoliter of 5µM RNP complex was injected into the cytoplasm of one-cell stage zebrafish embryos. To remove *bmp6* gene function in F0 embryos, a mixture of gRNA:Cas9 RNPs targeting exon 5 and exon 7 were injected into the cytoplasm of one-cell stage embryos. To generate zebrafish lacking *bmp6* gene function in the germline, RNP injected embryos were raised to adulthood and individual F1 embryos carrying deletions at the *bmp6* locus were identified using the primers below. We identified one allele, *z52* - a 1,749 bp deletion, which stably transmitted through the germline (Figure S1c).

Genomic DNA extraction, High Resolution Melt Analysis (HRMA), and PCR genotyping: For HRMA analysis and embryos genotyping, genomic DNA was extracted from individual embryos at 24 hours post fertilization (hpf). Dechorionated embryos were incubated in 30 ul 50 mM NaOH at 95°C, 20 min. 1/10 volume of 1 M Tris-HCl (pH 8.0) was added to neutralize. Genome sequences containing CRISPR/Cas9 target sites were amplified with pairs of primers: *bmp* exon 5

HRMA F3 – ACAGCCTGCAGAAAGCATGA and *bmp* exon 5 HRMA R3 – GCCAGCATTTGTTTACAGTACAGAG; *bmp6* exon 7 HRMA F4 – AGAACGTCCCAAAGCCATGT and *bmp6* exon 7 HRMA R4 – AACGCACCACCATGTTCCT. To determine if individual gRNA:Cas9 RNPs produced mutations at the desired target sites, HRMA was performed on DNA isolated from 8 individual 24 hpf gRNA:Cas9 RNP-injected embryos using LightScanner PCR Master Mix (BioFire) (32). To detect deletion events, PCR was performed on DNA isolated from 8 individual 24 hpf F0 gRNA:Cas9 RNP injected embryos using KAPA HiFi HotStart Ready Mix with the following primer pairs: *bmp6* F1 – CATGTGCTGGATAAGATGGTGA and *bmp6* R2 – TCCATAGATTCAGCGACGTTC (Figure S1b). These same primer pairs were used to detect deletion events in F1 embryos and adults. The following primer pairs were used to detect the WT *bmp6* locus: *bmp6* F1 – CATGTGCTGGATAAGATGGTGA and *bmp6* R1 – GTTCGATCCGCCTACATTTG.

<u>Cartilage and Bone Staining</u>: Fourteen days post fertilization (dpf) zebrafish larvae were anesthetized with Tricaine (3-amino benzoic acidethylester) and processed as previously described (33, 34) with the following modifications. Larvae were fixed in 2% paraformaldehyde for 1 hour, washed for 10 minutes in 50% EtOH, and then transferred to a solution containing 0.01% Alizarin Red and 0.04% Alcian Blue for 24 hours. Larvae were washed in 80 EtOH/10mM MgCl2 for 60 minutes, 50% EtOH for 30 minutes, 25 % EtOH for 30 minutes, bleached in 3% H_2O_2/0.5% KOH for 15 minutes, washed in 2X 25% glycerol/0.1% KOH and then transferred to 50% glycerol/0.1% KOH for imaging.

References:

1. Verbruggen G, Veys EM. Numerical scoring systems for the anatomic evolution of osteoarthritis of the finger joints. Arthritis Rheum. 1996;39(2):308-20.
2. Damman W, Liu R, Kroon FPB, Reijnierse M, Huizinga TWJ, Rosendaal FR, et al. Do Comorbidities Play a Role in Hand Osteoarthritis Disease Burden? Data from the Hand Osteoarthritis in Secondary Care Cohort. J Rheumatol. 2017;44(11):1659-66.
3. Wetzels JFM, Kiemeney LALM, Swinkels DW, Willems HL, Heijer Md. Age- and gender-specific reference values of estimated GFR in Caucasians: The Nijmegen Biomedical Study. Kidney Int. 2007;72(5):632-7.

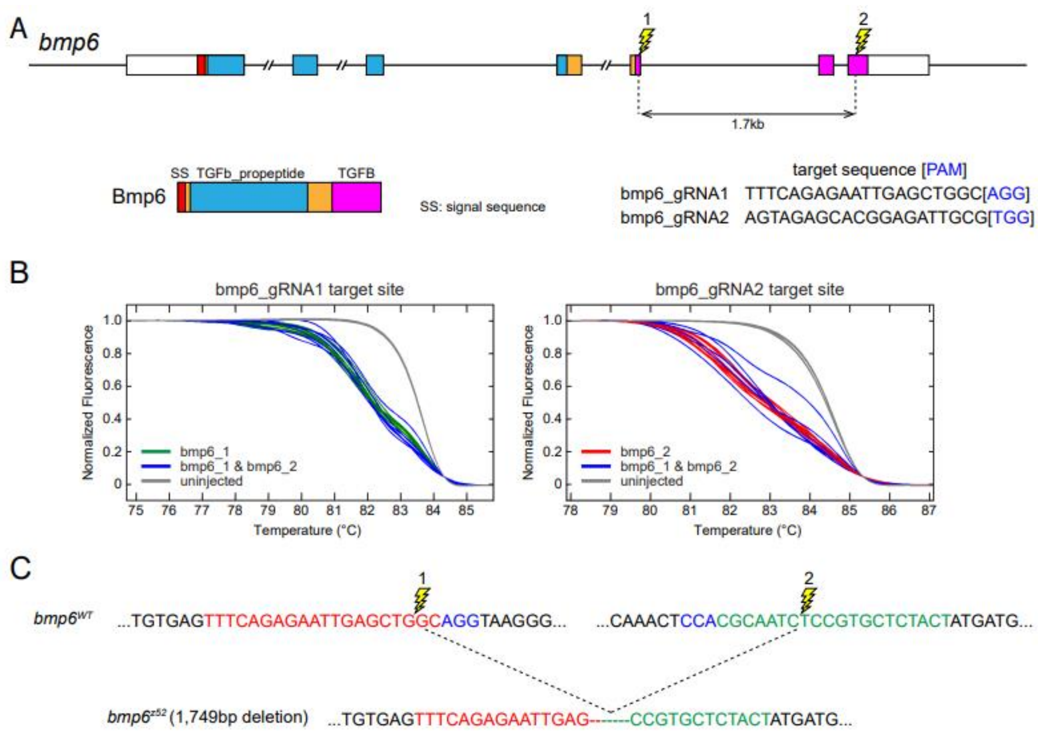
4. Kazmers NH, Meeks HD, Novak KA, Yu Z, Fulde GL, Thomas JL, et al. Familial Clustering of Erosive Hand Osteoarthritis in a Large Statewide Cohort. Arthritis Rheumatol. 2021;73(3):440-7.
5. Oreiro-Villar N, Raga AC, Rego-Pérez I, Pértega S, Silva-Diaz M, Freire M, et al. PROCOAC (PROspective COhort of A Coruña) description: Spanish prospective cohort to study osteoarthritis. Reumatologia clinica. 2020.
6. Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, Artigas MS, et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. The Lancet Respiratory Medicine. 2015;3(10):769-81.
7. Welsh S, Peakman T, Sheard S, Almond R. Comparison of DNA quantification methodology used in the DNA extraction protocol for the UK Biobank cohort. BMC Genomics. 2017;18(1):26.
8. Gudbjartsson DF, Helgason H, Gudjonsson SA, Zink F, Oddson A, Gylfason A, et al. Large-scale whole-genome sequencing of the Icelandic population. Nat Genet. 2015;47(5):435-44.
9. Jónsson H, Sulem P, Kehr B, Kristmundsdottir S, Zink F, Hjartarson E, et al. Whole genome characterization of sequence diversity of 15,220 Icelanders. Scientific Data. 2017;4:170115.
10. Eggertsson HP, Jonsson H, Kristmundsdottir S, Hjartarson E, Kehr B, Masson G, et al. Graphtyper enables population-scale genotyping using pangenome graphs. Nat Genet. 2017;49(11):1654-60.
11. Kong A, Masson G, Frigge ML, Gylfason A, Zusmanovich P, Thorleifsson G, et al. Detection of sharing by descent, long-range phasing and haplotype imputation. Nat Genet. 2008;40(9):1068.
12. Delaneau O, Howie B, Cox AJ, Zagury JF, Marchini J. Haplotype estimation using sequencing reads. Am J Hum Genet. 2013;93(4):687-96.
13. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS genetics. 2009;5(6):e1000529.
14. The Genomes Project C. A global reference for human genetic variation. Nature. 2015;526:68.
15. Delaneau O, Zagury JF, Robinson MR, Marchini JL, Dermitzakis ET. Accurate, scalable and integrative haplotype estimation. Nat Commun. 2019;10(1):5436.
16. Halldorsson BV, Eggertsson HP, Moore KHS, Hauswedell H, Eiriksson O, Ulfarsson MO, et al. The sequences of 150,119 genomes in the UK Biobank. Nature. 2022;607(7920):732-40.
17. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009;19(9):1655-64.
18. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006;38(8):904-9.
19. Price AL, Helgason A, Palsson S, Stefansson H, St Clair D, Andreassen OA, et al. The impact of divergence time on the nature of population structure: an example from Iceland. PLoS genetics. 2009;5(6):e1000505.
20. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007;39(7):906-13.
21. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47(3):291-5.
22. Sveinbjornsson G, Albrechtsen A, Zink F, Gudjonsson SA, Oddson A, Masson G, et al. Weighting sequence variants based on their annotation increases power of whole-genome association studies. Nat Genet. 2016;48(3):314-7.
23. Kong A, Frigge ML, Thorleifsson G, Stefansson H, Young AI, Zink F, et al. Selection against variants in the genome associated with educational attainment. Proc Natl Acad Sci U S A. 2017;114(5):E727-E32.
24. Vilhjálmsson BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S, et al. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. American journal of human genetics. 2015;97(4):576-92.

25. Zheng HF, Forgetta V, Hsu YH, Estrada K, Rosello-Diez A, Leo PJ, et al. Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. Nature. 2015;526(7571):112-7.
26. Morris JA, Kemp JP, Youlten SE, Laurent L, Logan JG, Chai RC, et al. An atlas of genetic influences on osteoporosis in humans and mice. Nature genetics. 2019;51(2):258-66.
27. Saevarsdottir S, Stefansdottir L, Sulem P, Thorleifsson G, Ferkingstad E, Rutsdottir G, et al. Multiomics analysis of rheumatoid arthritis yields sequence variants that have large effects on risk of the seropositive subset. Ann Rheum Dis. 2022;81(8):1085-95.
28. Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, Adrian J, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature. 2020;583(7818):699-710.
29. Meuleman W, Muratov A, Rynes E, Halow J, Lee K, Bates D, et al. Index and biological spectrum of human DNase I hypersensitive sites. Nature. 2020;584(7820):244-51.
30. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS genetics. 2014;10(5):e1004383.
31. Hoshijima K, Jurynec MJ, Klatt Shaw D, Jacobi AM, Behlke MA, Grunwald DJ. Highly Efficient CRISPR-Cas9-Based Methods for Generating Deletion Mutations and F0 Embryos that Lack Gene Function in Zebrafish. Developmental cell. 2019;51(5):645-57.e4.
32. Dahlem TJ, Hoshijima K, Jurynec MJ, Gunther D, Starker CG, Locke AS, et al. Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. PLoS genetics. 2012;8(8):e1002861.
33. Teerlink CC, Jurynec MJ, Hernandez R, Stevens J, Hughes DC, Brunker CP, et al. A role for the MEGF6 gene in predisposition to osteoporosis. Annals of Human Genetics. 2021;85(2):58-72.
34. Walker MB, Kimmel CB. A two-color acid-free cartilage and bone stain for zebrafish larvae. Biotechnic & histochemistry : official publication of the Biological Stain Commission. 2007;82(1):23-8.
35. Styrkarsdottir U, Lund SH, Saevarsdottir S, Magnusson MI, Gunnarsdottir K, Norddahl GL, et al. The CRTAC1 Protein in Plasma Is Associated With Osteoarthritis and Predicts Progression to Joint Replacement: A Large-Scale Proteomics Scan in Iceland. Arthritis & Rheumatology. 2021;73(11):2025-34.

SUPPLEMENTARY FIGURES

Supplementary Figure 1. Generation of zebrafish lacking bmp6 gene function

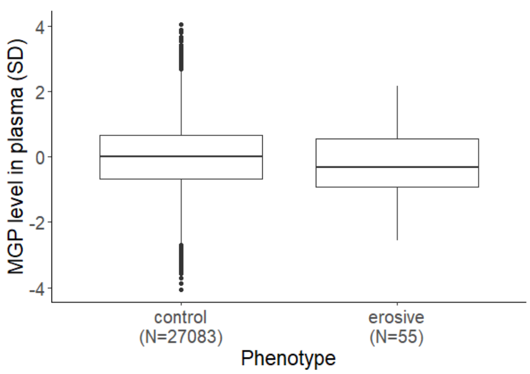
(A) Schematic illustration of the zebrafish *bmp6* locus indicating conserved protein domains (coloured regions) and the guide RNAs (lightning bolts) used to generate a deletion in the *bmp6* gene. (B) High resolution melt analysis (HRMA) detects indels generated in the genomes of 24 hpf WT or *bmp6* RNP injected embryos. HRMA analysis of WT embryos is represented as grey curves, bmp6_gRNA1 or bmp6_gRNA2 RNP as green and red curves, respectively, and embryos injected with both bmp6_gRNA1 and bmp6_gRNA2 RNPs as blue curves. (C) Schematic representation of WT and *bmp6*^{z52} loci. The *z52* is a 1,749 bp deletion that is stably transmitting through the germline.



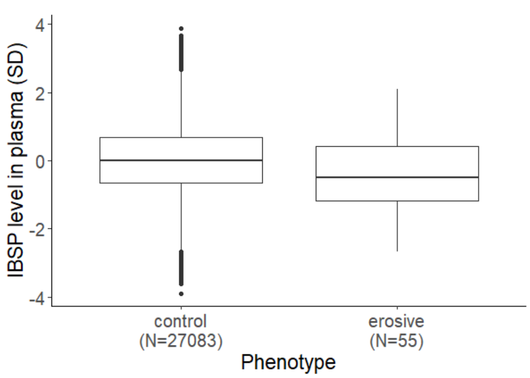
Supplementary Figure 2. Protein levels in plasma according to EHOA disease status

Standardized protein levels, adjusted for the age of the individual at the time of plasma collection, sex, collection site, and the storage age of the sample. After adjustment, the plasma protein levels were rank transformed onto the standard normal distribution with mean 0 and standard deviation 1 (35). Association of standardized protein levels with EHOA disease status (EHOA vs. controls) was estimated with logistic regression, adjusting for age at the time of plasma collection, sex, and BMI. Both proteins associate strongly with BMI (MGP: effect= 0.13, $P = 0$, IBSP: effect= −0.03, $P = 7.1 \times 10^{-38}$) and with age (MGP: effect=0.009, $P = 3.5 \times 10^{-4}$, IBSP: effect= 0.05, $P=5.6 \times 10^{-86}$).

a) Matrix gla protein (MGP)



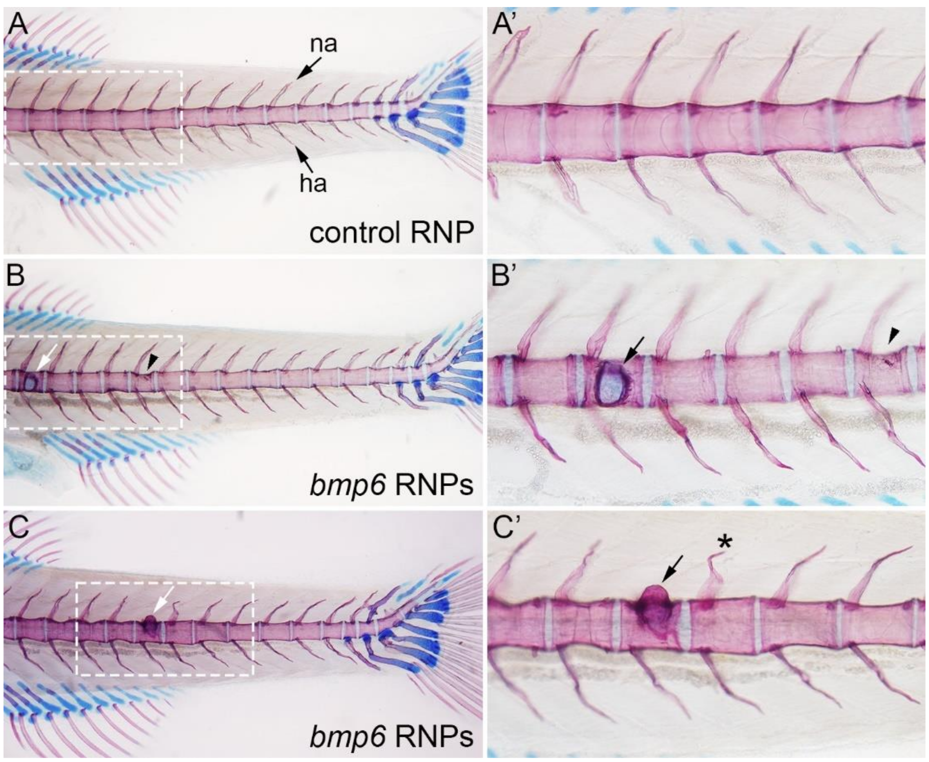
b) Bone sialoprotein 2 (IBSP)



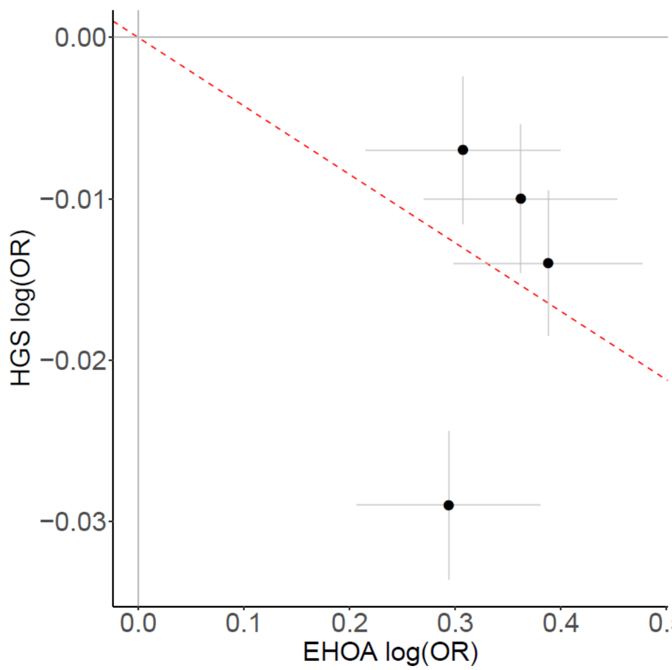
Styrkarsdottir U, *et al. Ann Rheum Dis* 2023;0:1–8. doi: 10.1136/ard-2022-223468

Supplementary Figure 3. Loss of bmp6 in F0 zebrafish larvae causes erosive-like phenotypes in the vertebral precursors similar to germline mutants.

(A-C'). Analysis of cartilage (blue) and bone (red) in the vertebral column of 14 days post fertilization control RNP and *bmp6* RNP injected zebrafish larvae. (A and A') Control RNP larvae have normally segmented and ossified centra (vertebral precursors) and neural (na) and hemal arches (ha). *bmp6* F0 mutant animals were generated by co-injection of bmp6_rRNA1 and bmp6_gRNA2 RNPs (see Figure S1) at the one-cell stage. In contrast to control RNP injected larvae (A and A'), *bmp6*$^{+/-}$ F0 mutant larvae have multiple defects, including large bone erosions (arrow in B and B'), ectopic bone formation in the centra (arrow in C and C'), structural defects in the centra (arrowhead in B and B'), and disruption of the neural arches (asterisk in C'). These are defects are also seen the in the germline allele (Figure 2). No defects are observed in the cartilaginous structures of the fins. All images are lateral views with anterior to the left.



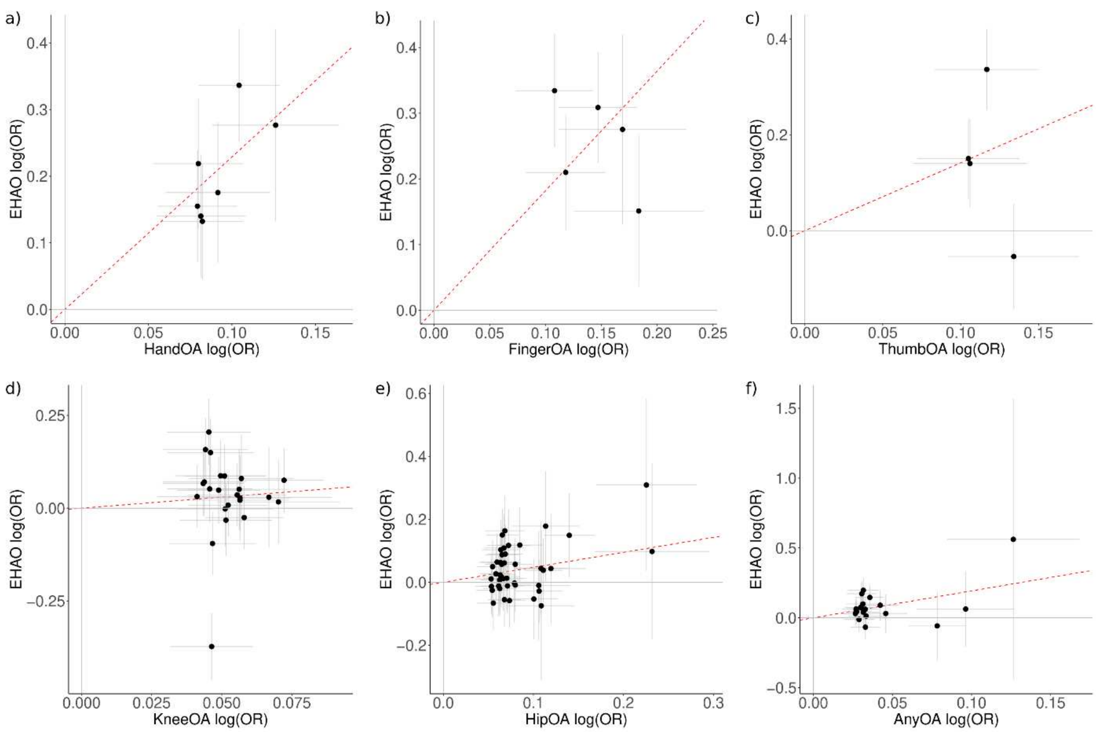
Supplementary Figure 4. Correlation between effects of EHOA variants on EHOA and hand grip strength



Styrkarsdottir U, *et al. Ann Rheum Dis* 2023;0:1–8. doi: 10.1136/ard-2022-223468

Supplementary Figure 5. Correlation of OR's between EHOA and other OA

The variants were identified by the GO consortium for a) hand OA, b) finger OA, c) thumb OA, d) knee OA, d) hip OA, and d) any type of OA (Boer et al, Cell, 2021). The logOR of these OA phenotypes in the GO data were plotted against the logOR of association of these variants with EHOA. Each variant is indicated by a dot and plotter with standard errors.



18