
Genetic underpinning of the comorbidity between type 2 diabetes and osteoarthritis

Authors

Ana Luiza Arruda, April Hartley, Georgia Katsoula,
George Davey Smith, Andrew P. Morris,
Eleftheria Zeggini

Correspondence

eleftheria.zeggini@helmholtz-munich.de

Multimorbidity is a rising public health challenge. Here, we study the genetic underpinning of the comorbidity between type 2 diabetes and osteoarthritis and provide insights into its biological basis. The approach can be applied to any pair of co-occurring diseases.

Arruda et al., 2023, *The American Journal of Human Genetics* 110, 1304–1318

August 3, 2023 © 2023 The Author(s).

<https://doi.org/10.1016/j.ajhg.2023.06.010>



Genetic underpinning of the comorbidity between type 2 diabetes and osteoarthritis

Ana Luiza Arruda,^{1,2,4} April Hartley,⁵ Georgia Katsoula,^{1,4} George Davey Smith,⁵ Andrew P. Morris,^{1,6} and Eleftheria Zeggini^{1,3,*}

Summary

Multimorbidity is a rising public health challenge with important implications for health management and policy. The most common multimorbidity pattern is the combination of cardiometabolic and osteoarticular diseases. Here, we study the genetic underpinning of the comorbidity between type 2 diabetes and osteoarthritis. We find genome-wide genetic correlation between the two diseases and robust evidence for association-signal colocalization at 18 genomic regions. We integrate multi-omics and functional information to resolve the colocalizing signals and identify high-confidence effector genes, including *FTO* and *IRX3*, which provide proof-of-concept insights into the epidemiologic link between obesity and both diseases. We find enrichment for lipid metabolism and skeletal formation pathways for signals underpinning the knee and hip osteoarthritis comorbidities with type 2 diabetes, respectively. Causal inference analysis identifies complex effects of tissue-specific gene expression on comorbidity outcomes. Our findings provide insights into the biological basis for the type 2 diabetes-osteoarthritis disease co-occurrence.

Introduction

Multimorbidity is defined as the coexistence of multiple chronic diseases in a single individual.¹ Worldwide, over 50% of the population older than 65 years is affected by more than one long-term medical condition simultaneously.² Commensurate with the rise in life expectancy and average population age, multimorbidity is an increasing global health challenge. However, the majority of health and drug development research is focused on treating and/or preventing individual diseases, leading to interventions that are currently not optimally designed to assist individuals suffering from multiple health conditions.

The most prevalent multimorbidity pattern among women and men is the combination of cardiometabolic and osteoarticular diseases,³ exemplified by the highly prevalent co-occurrence of type 2 diabetes and osteoarthritis.⁴ Between 2009 and 2016, approximately one in three adults with prediabetes in the US suffered from arthritis.⁵ Osteoarthritis is the most common whole-joint chronic disorder, affecting over 520 million people worldwide.⁶ It is a degenerative disorder characterized by a local and systemic low-grade inflammation state, irreversible loss of cartilage, and additional bone formation that results in pain, its most prevalent symptom.⁷ Across the globe, type 2 diabetes affects over 430 million people and is characterized by elevated blood glucose levels and insulin resistance.⁶ Both osteoarthritis and type 2 diabetes are complex diseases influenced by genetic, demographic, and lifestyle factors, such as older age and obesity.⁸

The majority of observational studies have reported a positive epidemiological association between type 2 diabetes and osteoarthritis of the hip or knee.⁴ In a meta-analysis including 1,040,175 individuals, the unadjusted odds ratio (OR) for type 2 diabetes in individuals with osteoarthritis versus non-osteoarthritis was 1.41 (95% confidence interval [CI] = [1.21, 1.65]).⁹ For individuals with type 2 diabetes, the overall risk of osteoarthritis was also higher than for individuals without type 2 diabetes (unadjusted OR = 1.46, 95% CI = [1.08, 1.96], $n = 32,137$).⁹ Articular joint-specific analyses have shown a stronger link between type 2 diabetes and knee osteoarthritis than hip osteoarthritis.⁹

Mendelian randomization (MR) analyses¹⁰ suggest no causal relation between liability to type 2 diabetes and knee osteoarthritis,¹¹ whereas body-mass index (BMI) has been shown to be causal for both diseases.^{12,13} When adjusting for BMI, studies linking type 2 diabetes and osteoarthritis have yielded conflicting results.^{4,9,14} Considering that obesity is a major risk factor for both diseases studied here, genetic variants associated with different physiological characteristics of increased adiposity are expected to be shared risk variants for the comorbidity. However, those variants could exert their effects on the comorbidity through alternative biological pathways to obesity through horizontal pleiotropy.¹⁰

Given the increase of the world's elderly population and the chronic nature of this highly prevalent pair of diseases, understanding their shared genetic background is important. Here, we focus on disentangling shared genetic risk loci between type 2 diabetes and osteoarthritis, including integration

¹Institute of Translational Genomics, Helmholtz Zentrum München – German Research Center for Environmental Health, 85764 Neuherberg, Germany; ²Munich School of Data Science, Helmholtz Zentrum München – German Research Center for Environmental Health, 85764 Neuherberg, Germany; ³TUM School of Medicine, Technical University Munich and Klinikum Rechts der Isar, 81675 Munich, Germany; ⁴Technical University of Munich (TUM), School of Medicine, Graduate School of Experimental Medicine, 81675 Munich, Germany; ⁵MRC Integrative Epidemiology Unit, University of Bristol, BS8 2BN Bristol, UK; ⁶Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, The University of Manchester, M13 9PT Manchester, UK

*Correspondence: eleftheria.zeggini@helmholtz-munich.de
<https://doi.org/10.1016/j.ajhg.2023.06.010>

© 2023 The Author(s). This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



Table 1. Sample sizes of GWASs used in this study

OA phenotype	Affected individuals	Control individuals	Total
All	177,517	649,173	826,690
Knee	62,497	333,557	396,054
Knee and/or hip	89,741	400,604	490,345
Hip	36,445	316,943	353,388
TKR	18,200	233,841	252,041
TJR	40,887	327,689	368,576
THR	23,021	296,016	319,037

An overview of the osteoarthritis (OA) phenotypes used in this work and the number of affected individuals, control individuals, and total number of individuals included in the corresponding GWAS (total knee replacement [TKR], total hip replacement [THR], total joint replacement [TJR], and osteoarthritis at any site [all]).

with functional genomics data in relevant cell types, in order to identify effector genes and provide insights into common underpinning mechanisms of disease development.

Material and methods

Datasets

For osteoarthritis, we used a recent large genome-wide association study (GWAS) meta-analysis from the Genetics of Osteoarthritis (GO) consortium.¹⁵ In total, it comprises data from 826,690 individuals (177,517 affected individuals) from mostly white European ancestry for 11 different osteoarthritis phenotypes. In this study, we used the following osteoarthritis phenotypes: knee, hip, knee and/or hip, total knee replacement (TKR), total hip replacement (THR), total joint replacement (TJR), and osteoarthritis at any site (all). An overview of the number of affected individuals, control individuals, and the total individuals for each study can be found in Table 1. For type 2 diabetes, the GWAS meta-analysis unadjusted for BMI from the DIAM ANTE consortium was used.¹⁶ It includes data from 898,130 individuals (74,124 affected individuals) of European ancestry.

We also employed molecular quantitative trait locus (QTL) data from disease-specific tissues. For osteoarthritis, we used expression quantitative trait locus (eQTL) data from intact cartilage ($n = 95$), degenerated cartilage ($n = 87$), and synovium ($n = 77$), as well as protein abundance quantitative trait locus (pQTL) data from intact and degenerated cartilage ($n = 99$).¹⁷ All samples were collected from individuals with osteoarthritis. For type 2 diabetes, we used eQTL data from pancreatic islets from the Inspire consortium.¹⁸ In the pancreatic islets dataset, 37 individuals out of 420 were diabetic.

We aligned the effect alleles of all datasets used in this paper by inverting the sign of the effect sizes when a mismatch was detected. Chromosome X was not included in any analysis. All datasets used the Genome Reference Consortium Human Build 37 (GRCh37) assembly.

Measures of adiposity

We used four measures of adiposity: BMI, waist-to-hip ratio (WHR) unadjusted for BMI, whole-body fat mass, and body fat percentage. For BMI ($n = 806,834$) and WHR ($n = 697,734$), we used the recent meta-analysis combining data from the GIANT consortium and the UK biobank.¹⁹ The inverse rank-normalized GWAS summary statistics for whole-body fat mass ($n = 330,762$) and body fat percentage ($n = 331,117$) were taken from the Neale's Lab website (<http://www.nealelab.is/uk-biobank/>). For each adiposity phenotype, we looked up the effect direction and signifi-

cance of all variants in the 95% credible set of the colocalized regions between type 2 diabetes and osteoarthritis (Table S11).

Quantification and statistical analysis

Genetic overlap of type 2 diabetes and osteoarthritis phenotypes

We conducted a linkage disequilibrium (LD) score regression analysis using the LDSC software (v.1.0.1) with `-rg` flag to estimate the genetic correlation between each osteoarthritis phenotype and type 2 diabetes (Table S1).²⁰ Because the majority of the GWASs used here comprise data of European ancestry individuals only, pre-computed LD scores from the 1000 Genomes European ancestry haplotypes were used.²¹ To assess the potential for chance findings when performing multiple statistical analyses, we performed a permutation-based analysis. We randomly permuted the effects (Z scores) of the variants for the osteoarthritis phenotypes 10,000 times while fixing the effects for type 2 diabetes. Running LD score regression on each permuted dataset yielded an empirical p value for the genetic correlation of type 2 diabetes and each analyzed osteoarthritis phenotype.

Statistical colocalization analysis

We defined regions of 2 Mb (± 1 Mb) around established independent association signals from each disease. For type 2 diabetes, we selected all primary and secondary independent signals from the BMI-unadjusted GWAS (p value threshold = 5×10^{-8}). For osteoarthritis, we selected the risk signals for the respective phenotype at the adjusted genome-wide significance of 1.3×10^{-8} . For each osteoarthritis phenotype, we performed regional pairwise statistical colocalization analysis with type 2 diabetes using the *coloc.abf* function from the *coloc* R package (version 3.2.1).²² Colocalization analyses were conducted using estimated regression coefficients (effect sizes) and standard errors (Table S2). In short, this function calculates posterior probabilities for five association configurations under the assumption of a single causal variant per trait. These configurations are summarized in the hypotheses below:

- H0: no trait has a genetic association in the region.
- H1: trait 1 has a genetic association in the region.
- H2: trait 2 has a genetic association in the region.
- H3: both traits have a genetic association in the region but with different causal variants.
- H4: both traits share a genetic association (single causal variant) in the region.

For all osteoarthritis phenotypes, we used the default prior probabilities of the *coloc* R package. We considered evidence for

colocalization if the posterior probability of H4 (PP4) > 0.8. For each genomic locus of colocalization, we calculated a 95% credible set for the causal variant by taking the cumulative sum of the variants' posterior probabilities to be causal conditional on H4 being true. LD between the single-nucleotide polymorphisms (SNPs) was calculated using plink (v.2.0 alpha)²³ based on the UK biobank²⁴ and was used for visualizing the results in regional association plots.

Knockout mouse phenotypes

We performed a schematic search for each gene in the vicinity of the genomic loci that colocalize between type 2 diabetes and osteoarthritis to screen for knockout mice showing phenotypes related to type 2 diabetes or osteoarthritis. The databases used in this scope were the International Mouse Phenotyping Consortium (IMPC) (<https://www.mousephenotype.org/>), Mouse Genome Informatics (MGI) (<http://www.informatics.jax.org/>), and Rat Genome Database (RGD) (<https://rgd.mcw.edu/>). For IMPC and RGD, we extracted the knockout mice phenotypes for each potential effector gene using the programmatic data access via their application programming interface (API). For MGI, we used the MGI batch query.

For type 2 diabetes, we looked for insulin- and diabetes-related phenotypes that included the following terms: insulin, glucose, diabetes, hyperglycemia, pancreas, pancreatic, obesity, BMI, body weight, body mass, body fat, beta cell, and glucosuria. For osteoarthritis, we looked for musculoskeletal phenotypes including the terms skeletal, muscle, bone, osteo, arthritis, muscular, joint, body size, growth, stature, and height.

Rare and syndromic human diseases

To investigate whether any analyzed genes are associated with a monogenic disorder, we extracted data from the Online Mendelian Inheritance in Man (OMIM) (<https://omim.org/>) by using their API. The terms we looked up for osteoarthritis-related phenotypes were bone, muscle, skeleton, osteo, arthritis, muscular, joint, body size, growth, skeletal, stature, height, hand-foot-uterus, synostosis, Martsolf, Warburg, leukodystrophy, squalene, and FINCA (as an abbreviation of fibrosis, neurodegeneration, and cerebral angiomas). For type 2 diabetes, we searched for insulin, glycemia, glucose, diabetes, pancreas, pancreatic, obesity, BMI, body weight, body mass, body fat, beta cell, glucosuria, Martsolf, aciduria, Aicardi-Goutières, and FINCA.

Differential gene expression

We explored whether the analyzed genes show differential expression for type 2 diabetes and osteoarthritis using published summary statistics from RNA sequencing (RNA-seq) datasets. For osteoarthritis, we assessed differential expression by comparing paired intact and degraded osteoarthritis cartilage from 124 individuals.²⁵ Because the samples were collected within person, the data is automatically robust against cofactors such as age and population structure. For type 2 diabetes, we used RNA-seq data from surgical pancreatic tissue samples from metabolically phenotyped pancreatectomized individuals. Samples were collected from 18 non-diabetic individuals and 39 individuals who were previously diagnosed with type 2 diabetes.²⁶ The differential expression analysis was based on a linear model with age, sex, and BMI as covariates. We considered genes that changed more than 1.5-fold in either direction and had an adjusted p value < 0.05 to be differentially expressed between degraded (high-grade) and intact (low-grade) osteoarthritis cartilage and diabetic versus healthy pancreatic islets for osteoarthritis and type 2 diabetes, respectively.

Multi-trait statistical colocalization analysis with eQTL and pQTL data

First, we superimposed molecular QTL information from disease-specific tissues by performing multi-trait molecular QTL-GWAS co-

localization analyses. The analyses were performed only on the variants in the 95% credible set. The input consisted of three summary statistics: one from the type 2 diabetes GWAS, one from the osteoarthritis phenotype GWAS, and one from the disease-relevant tissue molecular QTL dataset. Because one variant is tested for multiple genes in an eQTL dataset or multiple proteins for the pQTL datasets, we performed the colocalization gene-wise or protein-wise, respectively, such that for each analysis a single molecular QTL summary statistic is available for each variant. If the 95% credible set consisted only of a single variant for each gene or protein, we included all variants in a 1-Mb window in the analysis.

For the multi-trait statistical colocalization analyses, we used the R package *HyPrColoc* (v.1.0.0).²⁷ We conducted regional gene-wise analysis to assess whether all traits colocalize by switching off the Bayesian divisive clustering algorithm (bb.alg = FALSE). In a similar manner to the *coloc* package, we used *HyPrColoc* to estimate the posterior probabilities and identify candidate effector genes using multiple traits as input. For consistency, evidence for colocalization was considered at a threshold of 0.8 for PP4. The type 2 diabetes and osteoarthritis GWAS meta-analyses share five cohorts. Although the samples overlap, we assumed independence between the datasets, as instructed by the developers of the *HyPrColoc* package.

Using the prior knowledge that type 2 diabetes and osteoarthritis colocalize in the analyzed genomic loci, we adapted the prior parameters of the *HyPrColoc* algorithm accordingly. The first parameter, *prior.1*, which denotes the probability that a SNP is associated with one trait only, was set to 1×10^{-10} , six times smaller than the default. We set the second parameter *prior.2* to 0.7 instead of the default of 0.98. *1-prior.2* denotes the prior probability of a SNP being associated with an additional trait and *1-(prior.2)*² with the SNP being associated with the two other traits. LD between SNPs was again calculated using plink (v.2.0 alpha)²³ based on the UK biobank.²⁴

Scoring of potential effector genes

In genomic loci that colocalized between type 2 diabetes and at least one osteoarthritis phenotype with a PP4 > 0.8, we analyzed all genes in a 1-Mb window on either side of the lead variant of the 95% credible set. We incorporated orthogonal multi-omics and functional information to derive a list of high-confidence effector genes for the type 2 diabetes-osteoarthritis comorbidity.

Except for the pQTL analysis, all four above-mentioned biological lines of evidence were tested for both osteoarthritis and type 2 diabetes, yielding one separate score for each disease. Additionally, we incorporated information about previously established high-confidence effector genes for the individual diseases. For type 2 diabetes, we defined genes as high confidence if their top score in the type 2 diabetes knowledge portal was at least 4 (<https://t2d.hugeamp.org/>). For osteoarthritis, we selected genes scored as high confidence by the GO consortium.¹⁵ Because our analysis overlaps with criteria used to define a gene as high confidence for the individual diseases, we followed an approach to incorporate this information orthogonally: if a gene is high confidence for a disease but scored zero in our analysis, we updated the respective disease score to one.

We also looked up all variants in the 95% credible sets and searched for any missense variants for the genes located in the colocalized genomic loci. The results of this lookup were consolidated into an additional score for each gene, defined as the missense variant score. The total score was defined as the sum of the osteoarthritis score, the type 2 diabetes score, and the missense

variant lookup. However, if for a gene only the missense variant score is non-zero, the total score was set to zero because it is not relevant for the type 2 diabetes-osteoarthritis comorbidity.

Based on the scoring of the six orthogonal biological lines of evidence, we defined genes as potential effector genes if they showed at least one line of evidence for either one of the diseases. Genes that scored at least one line of evidence for osteoarthritis and one for type 2 diabetes were defined as likely effector genes for comorbidity. High-confidence effector genes were a subset of the likely effector genes that scored at least 3 in the total score (Table S3).

To further analyze our set of effector genes, we grouped them according to the osteoarthritis localization. If a gene is located in a genomic locus that colocalizes only between type 2 diabetes and the following three osteoarthritis phenotypes (osteoarthritis at any site, knee and/or hip osteoarthritis, and/or TJR), then it was considered to be associated equally with knee and hip osteoarthritis. If, in addition, the genomic locus colocalizes between type 2 diabetes and knee and/or TKR, then we considered the gene to be mostly associated with knee osteoarthritis. Similarly, if it also colocalizes between type 2 diabetes and hip and/or THR, then the gene was classified as mostly related to hip osteoarthritis.

Multi-trait statistical colocalization analysis with adiposity measures

In the genomic regions that colocalized between type 2 diabetes and osteoarthritis, we performed multi-trait colocalization analyses between type 2 diabetes, osteoarthritis, and the above-mentioned measures of adiposity. The analyses were performed only on the variants in the 95% credible set. If the 95% credible set consisted only of a single variant, for each gene or protein, we included all variants in a 1-Mb window on either side of the single variant in the analysis. As for the molecular QTL colocalization, we used the same functions of the R package *HyPrColoc* (v.1.0.0) and adjusted the prior parameters accordingly ($prior.1 = 1 \times 10^{-10}$, $prior.2 = 0.7$).²⁷ For consistency, evidence for colocalization was considered at a threshold of 0.8 for PP4.

Pathway analysis

We performed gene set enrichment analyses on the likely and high-confidence effector genes stratified by knee or hip osteoarthritis association (Table S5). The number of genes in each set is summarized in Table 2. We used the human resources and the enrichment software from the ConsensusPathDB (<http://cpdb.molgen.mpg.de/>) to examine the functional annotation of each gene set by testing their enrichment among curated networks in humans.²⁸ We used the networks from Reactome,²⁹ KEGG,³⁰ WikiPathways³¹ and Gene Ontology.³² For the latter, we included the subcategories molecular function, biological processes, and cellular component up to level 4. We required a minimum overlap of two genes for enrichment. The significance threshold was set at false discovery rate (FDR) < 0.05.

Classification of high-confidence genes on the basis of association with obesity

We have classified the high-confidence genes based on the level of association with obesity. For level 1, we searched OMIM for diseases or susceptibility to diseases associated with obesity. For level 2, we conducted a search on PhenoScanner (v.2, <http://www.phenoscanner.medschl.cam.ac.uk/>)³³ by querying their database using their R package *phenoscanner* and Ensembl (GRCh37) by using their API in R. If variants in the genes were associated with phenotypes associated with obesity, the gene was defined as level 2. We tried to capture different aspects of obesity by looking for following adiposity phenotypes: BMI, WHR, weight, fat percentage, and fat mass. We combined the results of PhenoScanner

and Ensembl to maximize the genes associated with adiposity measures. The remaining genes were included in level 3 and were defined as having no association with obesity.

Druggable genome

To outline drug repurposing targets, we queried the druggability status of the 72 likely effector genes for the comorbidity. We used the Druggable Genome database, which consists of 4,479 genes that are classified into three tiers depending on their progress in the drug development pipeline.³⁴ Tier 1 included 1,427 genes that are clinical-phase drug candidates or targets of already-approved small molecules and biotherapeutic drugs. Tier 2 consisted of 682 genes that encode targets with known bioactive drug-like small-molecule binding partners and genes with $\geq 50\%$ identity (over 75% of the sequence) with approved drug targets. Tier 3 comprised 2,370 genes encoding secreted or extracellular proteins, proteins with more distant similarity to approved drug targets, and members of key druggable gene families that were not included in tier 1 or 2. Tier 3 was further subdivided to prioritize genes in proximity (± 50 kbp) to a GWAS SNP from the GWAS catalog³⁵ and had an extracellular location (Tier 3A). Tier 3B is composed of the remaining genes.

For the likely effector genes included in tier 1, we further examined the approved or in-clinical-trial drugs by using the DrugBank online database (<https://www.drugbank.com>, accessed on August 1, 2022).

Causal inference analysis

Causal inference was strengthened through use of bidirectional two-sample MR between type 2 diabetes and all analyzed osteoarthritis phenotypes.¹⁰ We used the *TwoSampleMR* R package (v.0.5.6), which is curated by MR-Base.³⁶ We performed causal inference analyses on the full summary statistics (Table S6). For all analyses, instrumental variables (IVs) were selected as the genome-wide significant (p value $\leq 5 \times 10^{-8}$) and independent SNPs from the full data. Independence was defined as LD-based clumped SNPs with a strict LD threshold of $R^2 = 0.001$ over a 10-Mb window on either side of the index variant. To assure that the IVs are more strongly related to the exposure than to the outcome, we applied Steiger filtering.¹⁰ We applied the inverse-variance weighted (IVW) method, which performs a random-effects meta-analysis of the Wald ratios for each SNP, and the weighted median (WM) method. Finally, we performed sensitivity analyses by testing for heterogeneity based on the Q statistic using the *mr_heterogeneity* function from the *TwoSampleMR* R package. Horizontal pleiotropy was assessed through the intercept of the MR-Egger regression.³⁷ To account for multiple testing, p values were adjusted via the FDR approach.¹⁰

Two-step MR

We performed a two-step MR analysis between different adiposity measures and type 2 diabetes or osteoarthritis using *cis* eQTLs of each high-confidence effector gene in disease-relevant tissues as mediators (Table S7).³⁸ In the first step, we assessed whether adiposity was causal for the expression of our genes in the respective analyzed tissues. To assure independence of IVs between the two steps, we excluded independent eQTLs from the risk variants of each adiposity measure. Independence was defined by local LD-based clumping with $R^2 = 0.001$ over a 10-Mb window on either side of the index variant.

For the second step, we used independent genetics variants associated with the expression of each high-confidence gene as IVs and conducted a two-sample MR analysis between each of our genes and type 2 diabetes or osteoarthritis. For each analyzed tissue and each gene-disease pair, we conducted one MR analysis using the

Table 2. Number of genes in each gene set

Gene set	Number of genes
Likely effector genes	72
Likely effector genes related to knee	67
Likely effector genes related to hip	43
HC effector genes	19
HC effector genes related to knee	18
HC effector genes related to hip	10

HC = high confidence

TwoSampleMR R package (v.0.5.6).³⁶ If only one SNP was available after clumping and harmonizing the data, we employed the Wald ratio method. If more than one SNP remained after the pre-processing steps, we applied the IVW method and tested for heterogeneity with the *mr_heterogeneity* function. Moreover, if more than three SNPs were used for the causal inference analysis, we also tested for horizontal pleiotropy through MR-Egger regression.³⁷ Additionally, we estimated the F statistics from summary-level data as $mean(beta^2 / se^2)$ to assess the strength of the IVs.¹⁰ Finally, we adjusted the p values for multiple testing by using the FDR approach.

Tissue-specific effects

We determined the tissue-specific role of BMI in both osteoarthritis and type 2 diabetes using MR restricted to BMI instruments colocalizing with eQTLs in brain and adipose tissue, respectively, as described in Leyden et al.³⁹ (Table S12). In brief, summary-level MR was performed restricted to the 86 adipose tissue-colocalizing SNPs and the 140 brain tissue-colocalizing SNPs, where the numerator of the Wald ratio is the SNP effect on osteoarthritis or type 2 diabetes and the denominator is the effect estimate for the SNP on BMI from a GWAS meta-analysis of UK Biobank and the GIANT consortium⁴⁰ available at the MR-Base platform.³⁶ Osteoarthritis summary statistics were extracted from the recent GWAS of hip and knee osteoarthritis from the GO consortium.¹⁵ Type 2 diabetes summary statistics were extracted from the recent European DIAMANTE consortium GWAS.¹⁶

We used IVW meta-analysis of the individual SNP Wald ratios to estimate the causal effects of adipose-tissue-instrumented BMI and brain tissue-instrumented BMI on each outcome. As sensitivity analyses, we performed MR-Egger³⁷ to determine the potential role of pleiotropic effects (i.e., mediated via BMI-independent pathways), which gives a pleiotropy robust estimate of the causal effect assuming that there is no correlation between instrument strength (i.e., the association of the SNP with BMI) and the pleiotropic effect.¹⁰ We also performed WM analysis, which gives an unbiased estimate of the causal effect as long as less than 50% of the SNPs are invalid instruments.⁴¹ We performed a Z test to assess the effect difference between adipose tissue- and brain tissue-instrumented BMI MR analyses.

Results

Insights into disease biology and treatment targets

We first assessed the genetic correlation between type 2 diabetes ($N_{cases} = 74,124$, $N_{controls} = 824,006$) and osteoarthritis (knee: $N_{cases} = 62,497$, $N_{controls} = 333,557$; hip:

$N_{cases} = 36,445$, $N_{controls} = 316,943$) on a genome-wide scale by using data from the recent large GWAS meta-analyses (Table S1 and Figure 1). In line with epidemiological evidence, we found a greater magnitude of genetic correlation between type 2 diabetes and knee osteoarthritis ($r_g = 0.241$, $SE = 0.028$, $p = 2.65 \times 10^{-18}$) than osteoarthritis of the hip ($r_g = 0.078$, $SE = 0.029$, $p = 0.008$) (Figure 1A). To assess the potential for bias due to overlapping samples and different sample sizes, we also performed a permutation-based analysis (empirical p value for knee = 0.005, empirical p value for hip = 0.142) (Figure 1B and Table S2). Causal inference analyses using MR showed evidence for a non-causal relationship between the two diseases (Table S6), consistent with smaller-scale studies in the literature.¹²

Using pairwise Bayesian colocalization analyses on genome-wide significant regions for osteoarthritis ($p = 1.3 \times 10^{-8}$) or type 2 diabetes ($p = 5 \times 10^{-8}$), we found robust evidence of colocalization (posterior probability of a shared causal variant ≥ 0.8) for 51 signals. Some of those signals colocalize between type 2 diabetes, and more than one osteoarthritis phenotype related to the knee and/or hip, resulting in 18 unique colocalizing genomic loci (Table S2 and Figures S5–S22). Ten of those loci colocalize with type 2 diabetes for both hip and knee osteoarthritis, two colocalize for hip osteoarthritis only, and six colocalize only for knee osteoarthritis. In three genomic loci, the 95% credible set for the causal variant from the colocalization analysis consisted of a single variant (Figure 2).

To resolve the colocalizing signals between type 2 diabetes and osteoarthritis, we have incorporated multi-omics data and functional information. On the basis of six complementary lines of evidence, we analyzed and scored all 906 genes located in the colocalized genomic loci and identified shared high-confidence candidate effector genes for the type 2 diabetes-osteoarthritis comorbidity (Figure 3 and Table S3). Our analysis included further statistical colocalization of the shared signals with gene eQTLs and pQTLs from disease-relevant tissues (cartilage chondrocytes, synoviocytes, and/or pancreatic beta cells). Twelve of the 18 colocalizing regions between type 2 diabetes and osteoarthritis showed statistical evidence for colocalization with a molecular QTL. We searched the variants in the 95% credible set for any missense variant associated

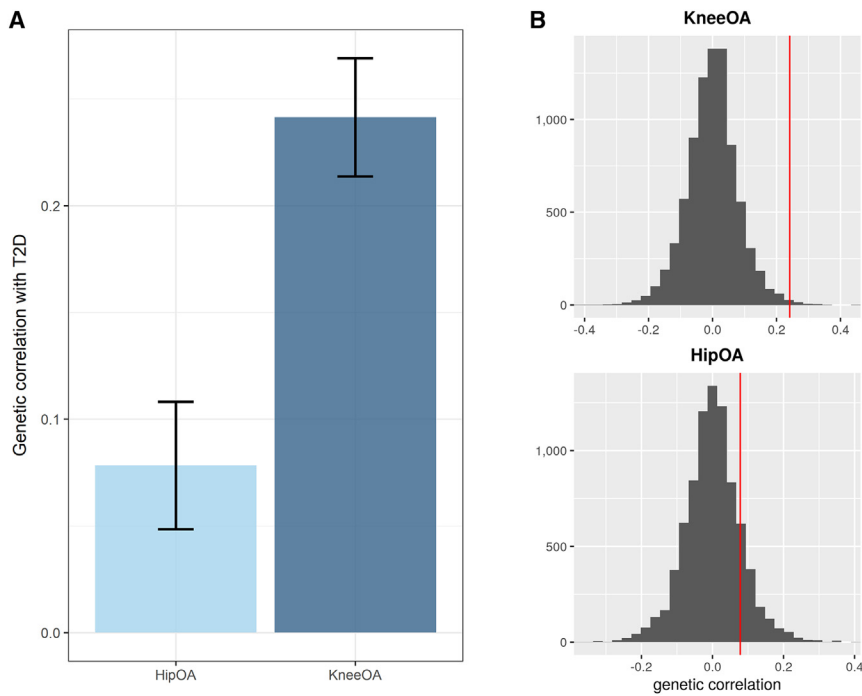


Figure 1. Stronger evidence for a genetic correlation between type 2 diabetes and knee osteoarthritis than between type 2 diabetes and hip osteoarthritis

(A) Genetic correlation (r_g) results between type 2 diabetes (T2D) and knee or hip osteoarthritis (OA). The error bars represent the standard error of the estimated genetic correlation.

(B) Permutation-based testing results for knee OA and hip OA, respectively. The red line is the actual correlation.

with the candidate genes. As a further line of evidence, we assessed whether the genes were differentially expressed in pancreatic islets from healthy versus diabetic individuals and intact versus degraded osteoarthritis cartilage. We searched the genes in databases for knockout mice and for rare and syndromic human diseases for association with pre-defined phenotypes related to type 2 diabetes and osteoarthritis. Finally, we included information on curated, previously defined effector genes for the individual diseases.

We defined 72 genes as likely effector genes for the type 2 diabetes-osteoarthritis comorbidity, as they displayed at least one line of supporting evidence for being involved in both diseases. Of the 72 likely effector genes, 19 showed at least three lines of evidence and were defined as high-confidence effector genes (Figure 4). These represent relevant candidates for further functional and clinical research. Knock-out mouse models for 17 out of the 19 high-scoring genes show phenotypes associated with both type 2 diabetes and osteoarthritis, which supports the role of those genes on the comorbidity. Eleven of these have not previously been defined as high-confidence effector genes for either disease based on a recent osteoarthritis study¹⁵ and the type 2 diabetes knowledge portal (<https://t2d.hugeamp.org/>). For two of the high-confidence genes, *APOE* and *WSCD2*, the 95% credible set for the causal variant from the colocalization analysis includes missense variants, namely rs429358 (c.466T>C [GenBank: NM_001302688.2] [p.Cys156Arg]) and rs3764002 (c.797C>T [GenBank: NM_014653.4] [p.Thr266Ile]). Six out of 19 high-confidence effector genes are the nearest gene to the lead variant of the respective colocalizing genomic locus: *WSCD2*, *TCF7L2*, *JADE2*, *GLIS3*, *FTO*, and *APOE* (Figure S3).

We searched the druggable genome for the druggability status of the likely effector genes for the type 2 diabetes and osteoarthritis comorbidity.³⁴ Sixteen out of 72 genes were included in the druggable genome (Table S4). Of these, six genes are tier 1 druggable targets (*GIPR*, *TPO*, *PAK1*, *SIGMAR1*, *CTSB*, *NOS3*), i.e., they are targets of drugs that have market authorization or are in clinical development. The

GIPR agonist tirzepatide was recently approved for the treatment of type 2 diabetes in adults. It has glucose-lowering effects and has been shown to increase insulin sensitivity.⁴² The *PAK1* inhibitor fostamatinib has been approved for the treatment of chronic immune thrombocytopenia.⁴³ It is also in clinical trial for the treatment of rheumatoid arthritis in order to alleviate the degree of inflammation of the joints.⁴⁴ *SIGMAR1* is a target of multiple approved drugs, including pentazocine, which is an analgesic used to treat moderate-to-severe pain. Naltrexone, an antagonist used in opioid overdose that also targets *SIGMAR1*, is being investigated for treating obesity.^{45,46} *TPO* encodes the thyroid peroxidase protein, which is the target of several approved drugs for the treatment of hyperthyroidism. One of these, the thyroid hormone dextrothyroxine, has been shown to lower serum levels of cholesterol in humans, but the interventional study has been discontinued due to serious adverse effects.⁴⁷

The 72 likely effector genes were enriched for several metabolic and cellular processes and for lipid localization and storage pathways. Hip osteoarthritis-related likely effector genes were enriched for bone-development pathways and metabolic processes. The 19 high-confidence effector genes were enriched for biological pathways related to diet and obesity (response to caloric restriction and *FTO* obesity variant mechanism) and for regulation of cell differentiation. The high-confidence genes related to hip osteoarthritis were enriched for the *FTO*-obesity-variant-mechanism pathway, regulation of lipid localization, and for a biological pathway related to skeletal formation (proximal/distal pattern formation) (Table S5 and Figure S4). These results provide biological support for the link between obesity and both diseases and for

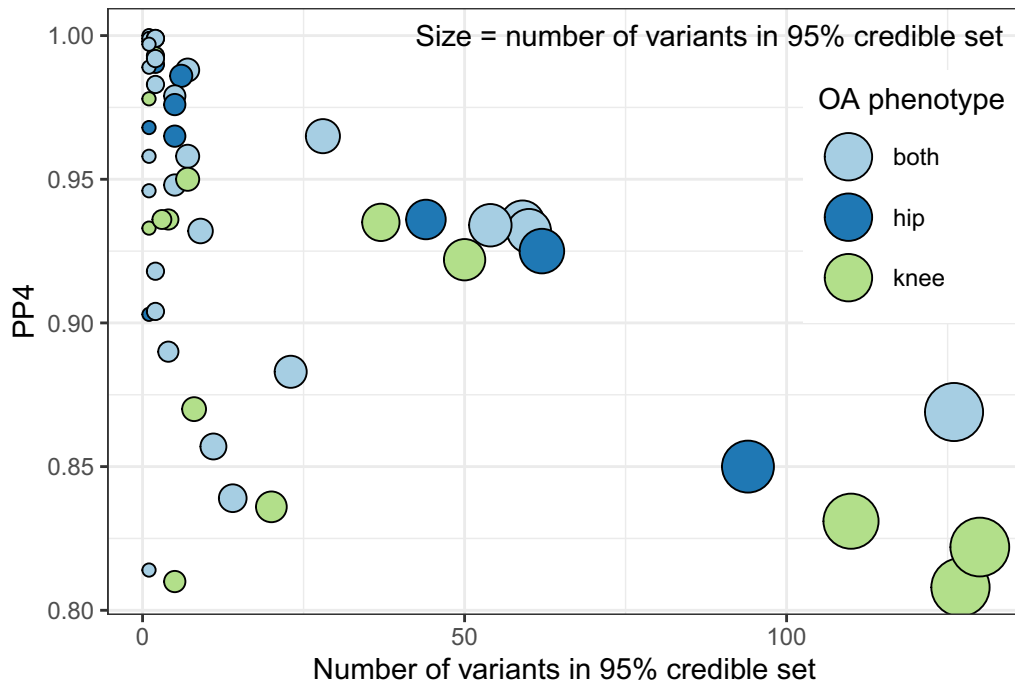


Figure 2. Overview of colocating regions between type 2 diabetes and osteoarthritis

The y axis depicts the posterior probability of a shared causal variant (PP4) and the x axis the number of variants in the 95% credible set for the causal variant. Each point represents a colocating signal between type 2 diabetes and one osteoarthritis (OA) phenotype. Point size is proportional to the number of variants in the colocalization analysis 95% credible set. We find strong statistical evidence for colocalization ($PP4 > 0.8$) for 51 signals. Some of those 51 signals colocalize between type 2 diabetes and more than one OA phenotype, resulting in 18 unique colocating genomic loci.

the association between bone development and hip osteoarthritis.⁷

Disentangling the effect of obesity

Obesity plays a causal role in both type 2 diabetes and osteoarthritis. To explore the role of obesity on the co-occurrence of type 2 diabetes and osteoarthritis, we studied four different measures that capture different aspects of obesity and adiposity: BMI, WHR, whole-body fat mass, and body fat percentage. Sixteen out of the 18 genomic regions that colocalized between type 2 diabetes and osteoarthritis show evidence of association or colocalization ($PP4 > 0.8$) with at least one adiposity-related trait (Table S11). Four high-confidence effector genes reside in the two genomic regions that do not show any evidence of colocalization or association with the analyzed measures of adiposity: *TMEM176A*, *RARRES2*, *SMARCD3*, and *GLIS3*. These may point to alternative biological mechanisms other than adiposity in the comorbidity between type 2 diabetes and osteoarthritis for these colocating signals.

We classified the high-confidence effector genes according to their level of association with obesity. Level 1 was defined as genes with variants directly associated with obesity or susceptibility to obesity in the OMIM database and included only one gene: *FTO*. Level 2 included nine genes for which variants in the gene were associated with the above-mentioned adiposity phenotypes on PhenoScanner or Ensembl. Finally, level 3 consisted of

the remaining nine genes, which were associated with obesity on the basis of our analysis (Table S13).

Next, we sought to investigate whether the adiposity measures that capture different aspects of obesity were causally associated with the expression of high-confidence effector genes in disease-relevant tissues. Within the constraints of the available instruments (material and methods), we found evidence of a causal relationship between several adiposity measures and nine high-confidence effector genes (Table S7). For example, we found that all measures of adiposity have a causal effect on higher expression of *IRX3* in synovium or pancreatic islets and on lower expression of *RTN2* in osteoarthritis cartilage. For the high-confidence effector genes located in the two genomic loci that did not show evidence of association or colocalization with adiposity, the direction of effect was not consistent across the different measures employed.

We assessed the causal role of BMI-associated variants with tissue-specific effects, selected based on evidence of their colocalization with brain or subcutaneous adipose tissue eQTLs.³⁹ For type 2 diabetes, we replicated previous results and showed that BMI-associated variants influencing genes expressed in brain tissue exert a stronger effect on the disease than adipose tissue-related variants,³⁹ although CIs largely overlapped. For knee osteoarthritis, we observed the same trend (Figure S23). For hip osteoarthritis, the results of the causal inference analysis provide evidence for a stronger effect of BMI-associated variants

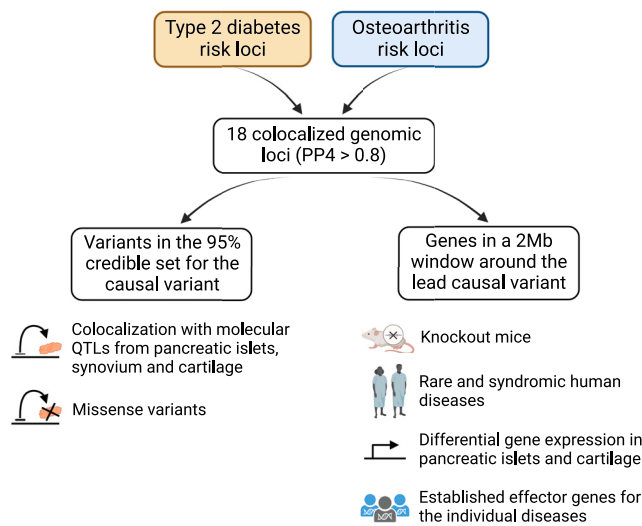


Figure 3. Study design to derive a list of high-confidence effector genes for the type 2 diabetes and osteoarthritis comorbidity

In each of the unique 18 genomic loci that colocalized between type 2 diabetes and osteoarthritis with a posterior probability of a single shared causal variant ($PP4 \geq 0.8$), we explored all genes in a 1-Mb window on either side of the lead variants of the 95% credible set for the causal variant of the colocalization analysis. For each gene, we searched databases for knockout mice and rare and syndromic human diseases for pre-defined type 2 diabetes- and musculoskeletal-related phenotypes. We also examined differentially expressed genes (DEGs) in pancreatic islets of healthy versus diabetic individuals and of degraded versus intact osteoarthritis cartilage. We also assessed whether the genes were already previously defined as established effector genes for the individual diseases. We examined all variants in the 95% credible set for the causal variant of each colocalization locus for missense variants within genes located in the colocalized genomic loci. We performed regional multi-trait colocalization analyses between type 2 diabetes, each osteoarthritis phenotype, and molecular QTLs from disease-relevant tissues.

that colocalize with adipose tissue eQTLs than with brain eQTLs (Table S12). Our results suggest a similar biological underpinning of the adiposity effect captured by BMI on type 2 diabetes and knee osteoarthritis but potentially different processes for hip osteoarthritis.

Insights gained from individual loci

FTO and *IRX3*

The obesity-related *FTO* locus colocalizes for type 2 diabetes and osteoarthritis with a posterior probability of a shared causal variant of over 92% (Figure 5A). The 95% credible set from the colocalization analysis consists of multiple variants in high LD with each other. The risk-increasing alleles for the lead causal variants are the same across type 2 diabetes and osteoarthritis. In addition to *FTO*, this locus is associated with a further high-confidence effector gene, *IRX3*. *IRX3* eQTLs in pancreatic islets colocalize with type 2 diabetes and osteoarthritis genetic signals with a $PP4 > 0.8$.

As shown above, adiposity is causally associated with an increase in *IRX3* expression in pancreatic islets and syno-

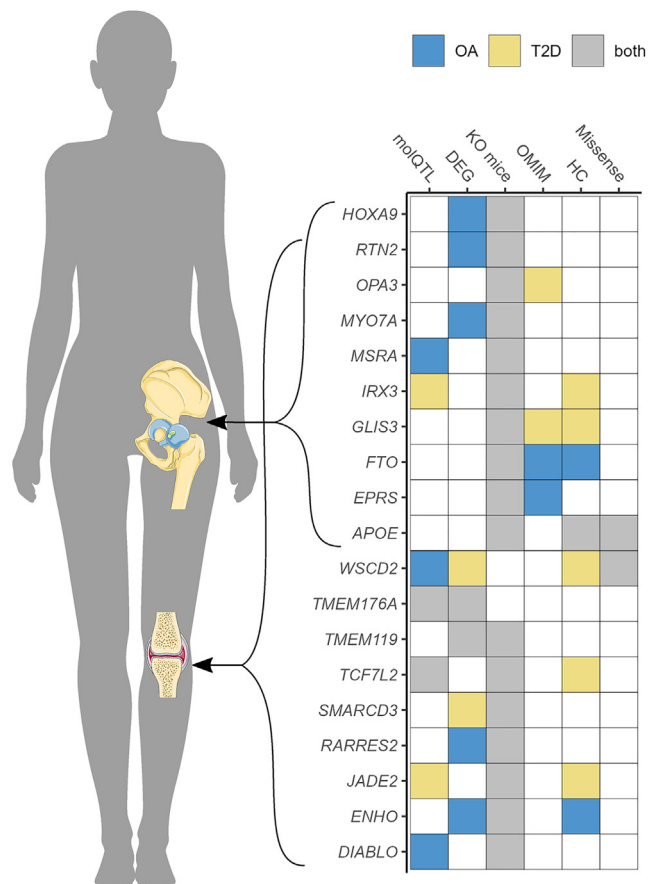


Figure 4. Overview of the 19 high-confidence effector genes for the type 2 diabetes and osteoarthritis comorbidity

Genes are stratified based on the joint affected by osteoarthritis. The scoring of the six biological lines of evidence is depicted on the right (material and methods). OA = osteoarthritis; T2D = type 2 diabetes; molQTLs = molecular quantitative trait loci; DEG = differentially expressed genes; KO mice = knockout mice; OMIM = Online Mendelian Inheritance in Man; HC = previously defined high-confidence effector genes; missense = missense variant.

vium (Table S7). Here, we performed causal inference analyses between the expression of high-confidence genes at this locus and type 2 diabetes or osteoarthritis. We find evidence for a causal effect of increased expression of *IRX3* in pancreatic islets on increased risk of type 2 diabetes (OR = 1.16, 95% CI = [1.08, 1.25], p value = 4.4×10^{-5} , F stat = 16.7).

FTO is a high-confidence osteoarthritis effector gene involved in skeletal development, adipogenesis, and neuronal function and development.¹⁵ It is also associated with syndromic human disease growth retardation, developmental delay, and facial dysmorphism (GDFD), which is a lethal autosomal-recessive multiple-congenital-anomaly syndrome.⁴⁸ *IRX3* is a known functional long-range target of *FTO* variants associated with obesity.⁴⁹ *FTO*- and *IRX3*-knockout mice show decreased body weight, decreased bone mineral density, and improved glucose tolerance (high bone mineral density is a risk factor for hip and knee osteoarthritis⁵⁰). We expected genes with an adiposity-driven effect to be involved in the shared genetic

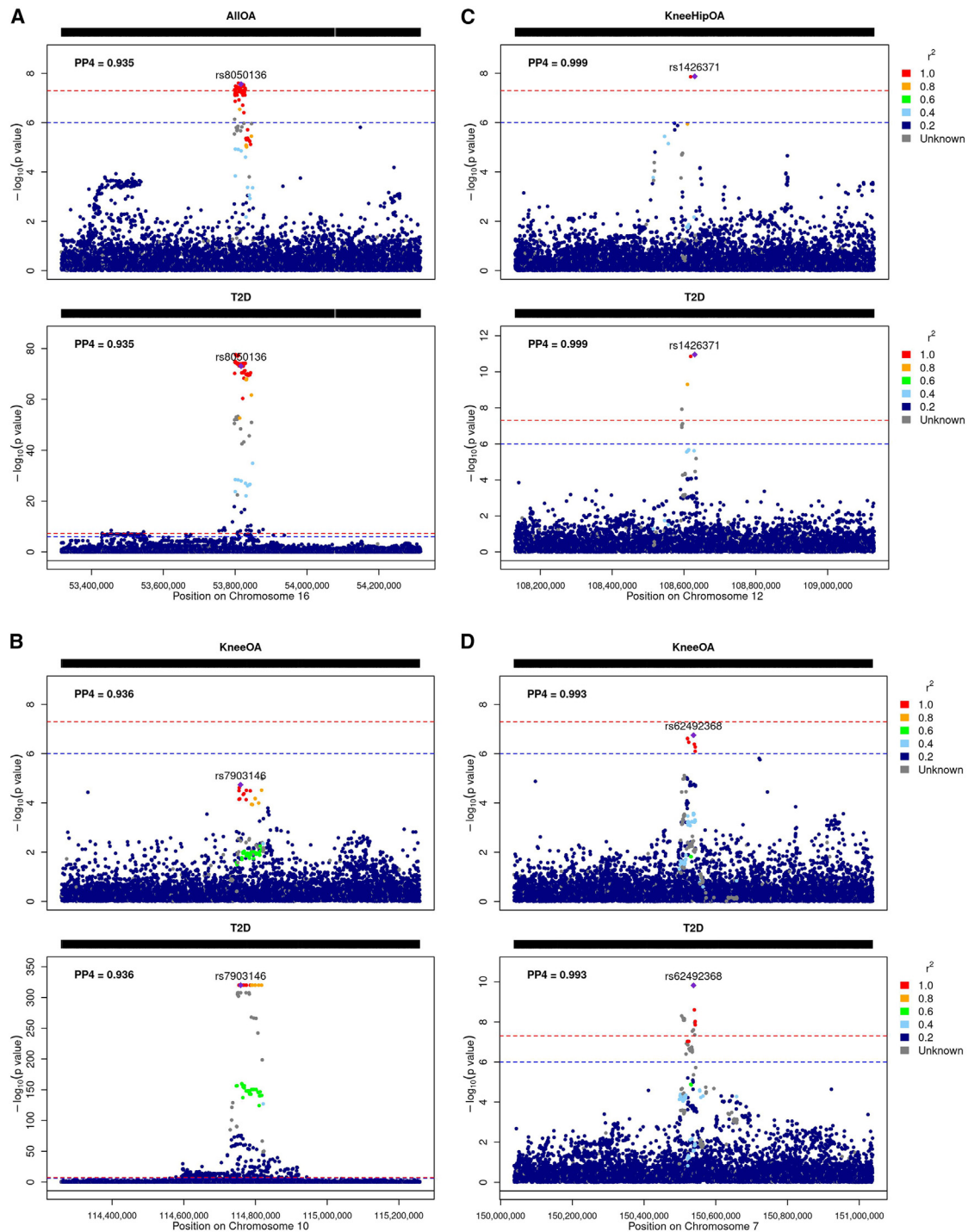


Figure 5. Regional association plots of the highlighted colocalizing regions between type 2 diabetes and osteoarthritis (A–D) *FTO* and *IRX3* region; (B) *TCF7L2* region; (C) *WSCD2* and *TMEM119* region; (D) *TMEM176A* region. The plots are colored on the basis of linkage disequilibrium between the lead causal variant of the colocalization and all other variants in the region. The red dashed line represents the genome-wide significance threshold (p value = 5×10^{-8}), and the blue dashed line represents a suggestive association threshold (p value = 10^{-6}). PP4 = posterior probability of a single shared causal variant; OA = osteoarthritis; T2D = type 2 diabetes; all OA = osteoarthritis at any site.

etiology of the type 2 diabetes-osteoarthritis comorbidity because obesity constitutes a common risk factor.⁸

TCF7L2

TCF7L2 is one of the highest-scoring effector genes. This genomic locus colocalizes for type 2 diabetes and knee

osteoarthritis with a posterior probability of a shared causal variant of over 93% (Figure 5B). Here, the 95% credible set for the causal variant from the colocalization analysis consists of three variants, which have opposite risk-increasing alleles for type 2 diabetes and knee osteoarthritis (Table S8).

Variants in *TCF7L2* have not been associated with osteoarthritis at genome-wide significance levels in the recent osteoarthritis GWAS (all variants in the credible set of the colocalization analysis achieve nominal significance in the recent knee osteoarthritis GWAS¹⁵). Genetic variants associated with *TCF7L2* expression in pancreatic islets and in osteoarthritis cartilage colocalize with this association signal. For the variants in both 95% credible sets from these colocalization analyses, the risk-increasing alleles for type 2 diabetes are associated with a lower BMI, a lower risk of knee osteoarthritis, and an increased expression of *TCF7L2* in pancreatic islets and osteoarthritic cartilage.

We found evidence that increased BMI causes decreased expression of *TCF7L2* in intact and degraded cartilage (Table S7). Additionally, we found evidence that an increase in *TCF7L2* expression in pancreatic islets causes an increase in type 2 diabetes risk (OR = 5.1, 95% CI = [4.7, 5.4], p value $< 1 \times 10^{-300}$, F stat = 2,018) and a decrease in knee osteoarthritis risk (OR = 0.83, 95% CI = [0.77, 0.91], p value = 1.77×10^{-5} , F stat = 18.4). These results are in line with the evidence shown above in support of an opposite effect of the genetic variants associated with the expression of *TCF7L2* in knee osteoarthritis and type 2 diabetes risk.

TCF7L2 is among the leading signals for type 2 diabetes risk and persists as a top signal after adjustment for BMI.⁵¹ Our results suggest that *TCF7L2* exerts an effect that goes through an alternative biological pathway to increased BMI. It has been shown that isoforms of *TCF7L2* regulate the expression of genes related to cartilage destruction in human chondrocytes.⁵² *TCF7L2* is a key effector gene of the Wnt/ β -catenin signaling pathway. This pathway plays a role in both type 2 diabetes, through glucose homeostasis, and osteoarthritis, through cartilage and bone formation.^{53,54}

TMEM119 and WSCD2

Two high-confidence effector genes, *WSCD2* and *TMEM119*, reside in the same genomic locus, which colocalizes for type 2 diabetes and knee osteoarthritis with a posterior probability of 99.9% (Figure 5C). The 95% credible set consists of two variants: rs1426371 and rs3764002, an intronic and a missense variant (amino acid change: Thr 266Ile) within *WSCD2*, respectively. The risk-increasing alleles of both variants are concordant for osteoarthritis of the knee and type 2 diabetes. The variant with the highest posterior probability of being causal for osteoarthritis and type 2 diabetes, rs1426371, has reached genome-wide significance levels in the recent knee osteoarthritis GWAS meta-analysis.¹⁵ The missense variant, rs3764002, is associated with WHR,⁵⁵ type 2 diabetes,¹⁶ lean mass,⁵⁶ and anxiety and neuroticism.⁵⁷ The missense variant is predicted to alter protein function, and this alteration is predicted to be damaging by both SIFT (<https://sift.bii.a-star.edu.sg>) and PolyPhen (<http://genetics.bwh.harvard.edu/pph2/>).

WSCD2 eQTLs in degraded osteoarthritis cartilage colocalize with type 2 diabetes and knee osteoarthritis with a

posterior probability of 99%. The lead eQTLs are rs1426371 and rs3764002. The expression level-increasing alleles are the same as the risk-increasing alleles for both diseases (Table S9). *WSCD2* is also a differentially expressed gene (DEG) in pancreatic islets from individuals with diabetes versus healthy individuals and is downregulated in diabetic islets. Moreover, it has been previously shown that *WSCD2* is functionally associated with type 2 diabetes and positively correlated with insulin secretion.^{58,59} Conclusions from causal inference analysis were limited as a result of weak instruments (F statistic < 10), which can bias causal effect inference (Table S7). Further research is needed to better understand the biological mechanisms through which *WSCD2* influences the type 2 diabetes-osteoarthritis comorbidity.

TMEM119 is the second-most highly scoring high-confidence effector gene. It is a DEG in osteoarthritis cartilage and pancreatic islets and is more highly expressed in degraded compared with intact cartilage and in healthy compared with pancreatic islets. Knockout mice for *TMEM119* show phenotypes related to both osteoarthritis and type 2 diabetes, such as decreased body weight, impaired osteoblast differentiation, and decreased compact bone thickness. *TMEM119* is related to bone formation by promoting osteoblast differentiation.⁶⁰ Fewer osteoblasts can lead to a decrease in compact bone thickness, which is also observed in knockout mice.⁶¹ The overexpression of *TMEM119* in degraded cartilage from individuals with osteoarthritis supports the evidence of an increase in bone formation in later stages of the disease.⁶² However, the lower expression of *TMEM119* in diabetes compared with healthy pancreatic islets points to further potential mechanisms of effect in the comorbidity.

TMEM176A

Variants in *TMEM176A*, also a high-confidence effector gene for the investigated comorbidity, have not been previously identified as implicated in either osteoarthritis or type 2 diabetes in the previous recent GWAS for the individual diseases. This locus colocalizes for type 2 diabetes and knee osteoarthritis with a posterior probability of a shared causal variant of 99.3% (Figure 5D). The index variants are rs62492368 and rs7794796, both located in the intron of *AOC1*. rs62492368 is associated with type 2 diabetes,¹⁶ and rs7794796 is associated with appendicular lean mass.⁶³ Type 2 diabetes and osteoarthritis show opposite risk-increasing alleles for all variants in the 95% credible set from the colocalization analysis (Table S10). These variants colocalize with PP4 > 0.8 between the diseases and eQTL data from pancreatic islets and synovium. The index variants from the colocalization with eQTLs also have opposite risk-increasing alleles for both diseases. Similar to the *TCF7L2* locus case, our results suggest that the mechanism through which *TMEM176A* exerts an effect on osteoarthritis and type 2 diabetes may have contrary directions.

We found body fat percentage to be linked to a decrease in the expression of *TMEM176A* in synovium (Table S7), albeit

with weak instruments (F statistic < 10). Decreased expression of *TMEM176A* in intact osteoarthritis cartilage and pancreatic islets was associated with reduced risk of type 2 diabetes (pancreatic islets: OR = 1.14, 95% CI = [1.07, 1.21], p value = 8.5×10^{-5} , F stat = 15.4; cartilage: OR = 1.05, 95% CI = [1.03, 1.08], p value = 8.5×10^{-5} , F stat = 15.4) and increased risk of knee osteoarthritis (pancreatic islets: OR = 0.93, 95% CI = [0.87, 0.99], p value = 0.048, F stat = 3.9; cartilage: OR = 0.97, 95% CI = [0.94, 0.99], p value = 0.048, F stat = 3.9). This genomic locus is one of the two colocalizing regions that do not show any evidence of statistical colocalization between type 2 diabetes, osteoarthritis, and the adiposity measures studied here. This suggests that this region, and possibly *TMEM176A*, acts through an alternative biological path to adiposity.

Results of the causal inference analysis mirror the output of differential expression analyses conducted in pancreatic islets and osteoarthritic cartilage.^{25,26} While *TMEM176A* is more highly expressed in pancreatic islets of diabetic individuals than in healthy pancreatic islets, an increase of its expression in the same tissue has a causal effect on increased risk of type 2 diabetes. Similarly, while *TMEM176A* was found to have lower expression in degraded compared with intact cartilage, the reduced expression of this gene in intact cartilage has a causal effect on increased risk of TKR.

Discussion

We present a genetic databased approach to disentangle the shared genetic etiology between two co-occurring chronic diseases and applied it to a common comorbidity pair: type 2 diabetes and osteoarthritis. Studies have shown a stronger association of BMI with osteoarthritis of the knee than of the hip.¹² We find stronger statistical evidence of a genetic correlation between type 2 diabetes and knee osteoarthritis than between type 2 diabetes and hip osteoarthritis. By leveraging recent large-scale GWASs for both diseases, we find evidence of colocalization at 18 genomic loci, and by incorporating multi-omics and functional genomics information, we derive a list of 19 high-confidence effector genes for the comorbidity. The majority of genomic loci colocalize for type 2 diabetes and knee, rather than hip, osteoarthritis, in keeping with the genome-wide-correlation analysis results.

Our findings support the epidemiological link between obesity, osteoarthritis, and type 2 diabetes. In this case, only two of the 18 colocalized regions do not colocalize with measures of adiposity. Several of the high-confidence genes, including *FTO* and *IRX3*, are associated with obesity-related traits. We show that the identified high-confidence effector genes are enriched for biological pathways associated with adiposity. Stratifying the high-confidence effector genes into knee or hip osteoarthritis provides further insight into the biological mechanisms underlying the comorbidity. High-confidence effector genes mostly related to hip osteoarthritis are also enriched for biological

pathways of skeletal formation, which underlines the strong link between bone development and hip osteoarthritis.⁷ Given that two-thirds of FDA-approved drugs are supported by genetic evidence, we explore the druggable potential of the prioritized genes.⁶⁴ We highlight approved drugs currently used for the treatment of diabetes, obesity, pain, and inflammation.

Observational studies report that the positive association between type 2 diabetes and osteoarthritis persists after adjusting for BMI.⁹ As BMI only captures a limited subset of the effect of adiposity on the comorbidity, this could be a source of residual confounding due to measurement error,⁶⁵ and observed attenuation of association can be underestimated. We performed in-depth analyses to disentangle the role of adiposity on comorbidity and find evidence that *TCF7L2* and *TMEM176A* exert an effect on type 2 diabetes and osteoarthritis through an alternative biological path. Further examination, including functional studies, is needed to dissect the precise way in which these genes affect comorbidity.

Type 2 diabetes and insulin resistance are known to be negatively correlated with bone strength and are also associated with bone fracture.⁶⁶ One possible link between bone and lipid metabolism is the fact that osteoblasts and adipocytes share a common progenitor cell in adult bone marrow with a degree of plasticity that can lead to an imbalance between the two cell lineages.⁶⁷ In support of this link, differentiation regulation of osteoblasts is highlighted by one of the identified high-confidence effector genes, *TMEM119*. In summary, we highlight three potential biological mechanisms underpinning the comorbidity between type 2 diabetes and osteoarthritis: obesity, imbalance between osteoblasts and adipocytes differentiation in adult bone marrow and the Wnt/ β -catenin signaling pathway.

We apply statistical colocalization analysis on regions where at least one of the studied diseases shows evidence of genome-wide association (p value < 5×10^{-8}) but not necessarily both. We have chosen this approach to overcome possible power issues from the individual GWASs due to different disease heterogeneity and sample sizes. Further, we embellish the colocalization results with a deep dive into biological lines of evidence for effector gene involvement in the colocalizing regions. The genetic and functional genomic data employed in this study are biased toward European populations. Going forward, it will be important to expand analyses to data from diverse populations. We have not performed analysis adjusted for BMI or other obesity-related phenotypes to avoid introducing collider bias. The eQTL data from pancreatic islets used in the analyses here comprise almost four times as many samples as the eQTL data from chondrocytes. Therefore, molecular QTL analyses for type 2 diabetes-relevant tissues were better powered than for osteoarthritis-relevant tissues. MR and the subsequent sensitivity analyses were conducted within the constraints of available instruments for expression of the high-confidence genes. This was

partially because the molecular QTL data used in this work include *cis* QTLs only, which restricts the analyses to variants within the vicinity of the genes or proteins of interest. Future studies should include a wider array of (as of now unavailable) genome-wide molecular QTLs, including at the single-cell level.

This work provides a proof of concept for the application of a study design that is relevant to any pair of comorbid diseases. We have studied one of the most frequently co-occurring pairs of complex diseases: type 2 diabetes and osteoarthritis. Our findings offer insights into the biological processes underpinning comorbidity and highlight potential drug repurposing opportunities in addition to new targets. As the world population life expectancy continues on an upward trajectory, the challenge of tackling multimorbidity will continue to be high on the healthcare agenda. Genomic-data-based approaches, as exemplified here for type 2 diabetes and osteoarthritis, can help improve our understanding of the co-occurrence of chronic conditions.

Data and code availability

- All original code produced for this manuscript has been deposited at Zenodo and is publicly available (<https://doi.org/10.5281/zenodo.7525171>).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request, Eleftheria Zeggini.

Supplemental information

Supplemental information can be found online at <https://doi.org/10.1016/j.ajhg.2023.06.010>.

Acknowledgments

We thank Lorraine Southam and Konstantinos Hatzikotoulas for their input and support. We also thank the Genetics of Osteoarthritis (GO) and the DIAMANTE and the InSPIRE consortia for providing their data.

Author contributions

E.Z. and A.L.A. conceived and designed the study and wrote the manuscript. A.L.A. and A.H. performed computation and statistical analyses and produced the figures. A.L.A., G.D.S., A.H., E.Z., and A.P.M. performed data interpretation. A.L.A., G.D.S., A.H., E.Z., G.K., and A.P.M. provided significant feedback on the methods, results, and manuscript.

Declaration of interests

A.H. and G.D.S. work in a unit funded by the Medical Research Council (MRC) and the University of Bristol (MC_UU_00011/1, MC_UU_00011/3).

Received: March 13, 2023

Accepted: June 20, 2023

Published: July 10, 2023

Web resources

coloc R package (v.3.2.1) (<https://chr1swallace.github.io/coloc/>)
ConsensusPathDB-human (<http://cpdb.molgen.mpg.de/>)
Drugbank (<https://www.drugbank.com>)
eQTLs and pQTLs for osteoarthritis cartilage and synovium¹⁷ (<https://msk.hugeamp.org/downloads.html>)
eQTLs for pancreatic islets¹⁸ (<https://zenodo.org/record/3408356>)
Gene Ontology (<http://geneontology.org/>)
GWAS catalog (<https://www.ebi.ac.uk/gwas/>)
GWAS summary statistics for BMI and WHR unadjusted for BMI¹⁹ (https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium)
GWAS summary statistics for osteoarthritis¹⁵ (<https://msk.hugeamp.org/downloads.html>)
GWAS summary statistics for type 2 diabetes¹⁶ (www.diagram-consortium.org/)
HyPrColoc R package (v.1.0.0) (<https://github.com/cnfoley/hyprcoloc>)
International Mouse Phenotyping Consortium (<https://www.mousephenotype.org/>)
KEGG (<https://www.genome.jp/kegg/>)
LD score regression (v.1.0.1) (<https://github.com/bulik/ldsc/>)
Mouse Genome Informatics (<http://www.informatics.jax.org/>)
OMIM database (<https://www.omim.org/>)
PhenoScanner V2 (<http://www.phenoscanter.medschl.cam.ac.uk/>)
PLINK 2.0 (<https://www.cog-genomics.org/plink/2.0/>)
PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>)
R statistical software (v.4.0.0) (<https://www.R-project.org/>)
Rat Genome Database (<https://rgd.mcg.edu/>)
Reactome (<https://reactome.org/>)
SIFT (<https://sift.bii.a-star.edu.sg>)
TwoSampleMR R package (v.0.5.6) (<https://mrcieu.github.io/TwoSampleMR/>)
Type 2 diabetes knowledge portal (<https://t2d.hugeamp.org/>)
WikiPathways (<https://www.wikipathways.org/>)

References

1. Nicholson, K., Makovski, T.T., Griffith, L.E., Raina, P., Stranges, S., and van den Akker, M. (2019). Multimorbidity and comorbidity revisited: refining the concepts for international health research. *J. Clin. Epidemiol.* *105*, 142–146.
2. Nguyen, H., Manolova, G., Daskalopoulou, C., Vitoratou, S., Prince, M., and Prina, A.M. (2019). Prevalence of multimorbidity in community settings: A systematic review and meta-analysis of observational studies. *J. Comorb.* *9*, 2235042X19870934.
3. Bezerra de Souza, D.L., Oliveras-Fabregas, A., Espelt, A., Bosque-Prous, M., de Camargo Cancela, M., Teixidó-Compañó, E., and Jerez-Roig, J. (2021). Multimorbidity and its associated factors among adults aged 50 and over: A cross-sectional study in 17 European countries. *PLoS One* *16*, e0246623.
4. Williams, M.F., London, D.A., Husni, E.M., Navaneethan, S., and Kashyap, S.R. (2016). Type 2 diabetes and osteoarthritis: A systematic review and meta-analysis. *J. Diabetes Complications* *30*, 944–950.
5. Sandoval-Rosario, M., Nayeri, B.M., Rascon, A., Rascon, A., Boring, M., Aseret-Manygoats, T., Helmick, C.G., Murphy, L.B., Hootman, J.M., Imperatore, G., and Barbour, K.E. (2018). Prevalence of Arthritis Among Adults with Prediabetes and Arthritis-Specific Barriers to Important Interventions for

- Prediabetes — United States, 2009–2016. *MMWR Morb. Mortal. Wkly. Rep.* *67*, 1238–1241.
6. Vos, T., Lim, S.S., Abbafati, C., Abbas, K.M., Abbasi, M., Abbafard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A., et al.; GBD 2019 Diseases and Injuries Collaborators (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* *396*, 1204–1222.
 7. Martel-Pelletier, J., Barr, A.J., Cicuttini, F.M., Conaghan, P.G., Cooper, C., Goldring, M.B., Goldring, S.R., Jones, G., Teichtahl, A.J., and Pelletier, J.-P. (2016). Osteoarthritis. *Nat. Rev. Dis. Primers* *2*, 1–18.
 8. Veronese, N., Cooper, C., Reginster, J.-Y., Hochberg, M., Branco, J., Bruyère, O., Chapurlat, R., Al-Daghri, N., Dennison, E., Herrero-Beaumont, G., et al. (2019). Type 2 diabetes mellitus and osteoarthritis Europe PMC Funders Group. *Semin. Arthritis Rheum.* *49*, 9–19.
 9. Louati, K., Vidal, C., Berenbaum, F., and Sellam, J. (2015). Association between diabetes mellitus and osteoarthritis: Systematic literature review and meta-analysis. *RMD Open* *1*, e000077.
 10. Sanderson, E., Glymour, M.M., Holmes, M.V., Kang, H., Morrison, J., Munafò, M.R., Palmer, T., Schooling, C.M., Wallace, C., Zhao, Q., and Davey Smith, G. (2022). Mendelian randomization. *Nat. Rev. Methods Primers* *2*, 6–21.
 11. Cui, Z., Feng, H., He, B., Xing, Y., Liu, Z., and Tian, Y. (2020). Type 2 Diabetes and Glycemic Traits Are Not Causal Factors of Osteoarthritis: A Two-Sample Mendelian Randomization Analysis. *Front. Genet.* *11*, 1726–1726.
 12. Zengini, E., Hatzikotoulas, K., Tachmazidou, I., Steinberg, J., Hartwig, F.P., Southam, L., Hackinger, S., Boer, C.G., Styrkarsdóttir, U., Gilly, A., et al. (2018). Genome-wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis. *Nat. Genet.* *50*, 549–558.
 13. Holmes, M.V., Lange, L.A., Palmer, T., Lanktree, M.B., North, K.E., Almqvora, B., Buxbaum, S., Chandrupatla, H.R., Elbers, C.C., Guo, Y., et al. (2014). Causal effects of body mass index on cardiometabolic traits and events: A Mendelian randomization analysis. *Am. J. Hum. Genet.* *94*, 198–208.
 14. Khor, A., Ma, C.A., Hong, C., Hui, L.L.Y., and Leung, Y.Y. (2020). Diabetes mellitus is not a risk factor for osteoarthritis. *RMD Open* *6*, e001030.
 15. Boer, C.G., Hatzikotoulas, K., Southam, L., Stefánsdóttir, L., Zhang, Y., Coutinho de Almeida, R., Wu, T.T., Zheng, J., Hartley, A., Teder-Laving, M., et al. (2021). Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell* *184*, 6003–6005.
 16. Mahajan, A., Taliun, D., Thurner, M., Robertson, N.R., Torres, J.M., Rayner, N.W., Payne, A.J., Steinthorsdóttir, V., Scott, R.A., Grarup, N., et al. (2018). Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* *50*, 1505–1513.
 17. Steinberg, J., Southam, L., Roumeliotis, T.I., Clark, M.J., Jayasuriya, R.L., Swift, D., Shah, K.M., Butterfield, N.C., Brooks, R.A., McCaskie, A.W., et al. (2021). A molecular quantitative trait locus map for osteoarthritis. *Nat. Commun.* *12*, 1309–1311.
 18. Viñuela, A., Varshney, A., van de Bunt, M., Prasad, R.B., Asplund, O., Bennett, A., Boehnke, M., Brown, A.A., Erdos, M.R., Fadista, J., et al. (2020). Genetic variant effects on gene expression in human pancreatic islets and their implications for T2D. *Nat. Commun.* *11*, 4912.
 19. Pulit, S.L., Stoneman, C., Morris, A.P., Wood, A.R., Glastonbury, C.A., Tyrrell, J., Yengo, L., Ferreira, T., Marouli, E., Ji, Y., et al. (2019). Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum. Mol. Genet.* *28*, 166–174.
 20. Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* *47*, 291–295.
 21. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. *Nature* *526*, 68–74.
 22. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., and Plagnol, V. (2014). Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet.* *10*, e1004383.
 23. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* *81*, 559–575.
 24. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O’Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* *562*, 203–209.
 25. Katsoula, G., Steinberg, J., Tuerlings, M., Coutinho de Almeida, R., Southam, L., Swift, D., Meulenbelt, I., Wilkinson, J.M., and Zeggini, E. (2022). A molecular map of long non-coding RNA expression, isoform switching and alternative splicing in osteoarthritis. *Hum. Mol. Genet.* *31*, 2090–2105.
 26. Wigger, L., Barovic, M., Brunner, A.D., Marzetta, F., Schöniger, E., Mehl, E., Kipke, N., Friedland, D., Burdet, F., Kessler, C., et al. (2021). Multi-omics profiling of living human pancreatic islet donors reveals heterogeneous beta cell trajectories towards type 2 diabetes. *Nat. Metab.* *3*, 1017–1031.
 27. Foley, C.N., Staley, J.R., Breen, P.G., Sun, B.B., Kirk, P.D.W., Burgess, S., and Howson, J.M.M. (2021). A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat. Commun.* *12*, 764.
 28. Kamburov, A., Pentchev, K., Galicka, H., Wierling, C., Leh-rach, H., and Herwig, R. (2011). ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Res.* *39*, D712–D717.
 29. Gillespie, M., Jassal, B., Stephan, R., Milacic, M., Rothfels, K., Senff-Ribeiro, A., Griss, J., Sevilla, C., Matthews, L., Gong, C., et al. (2022). The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* *50*, D687–D692.
 30. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* *44*, D457–D462.
 31. Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenter, D.N., Hanspers, K., A Miller, R., Digles, D., Lopes, E.N., Ehrhart, F., et al. (2021). WikiPathways: connecting communities. *Nucleic Acids Res.* *49*, D613–D621.
 32. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene Ontology: tool for the unification of biology. *Nat. Genet.* *25*, 25–29.

33. Kamat, M.A., Blackshaw, J.A., Young, R., Surendran, P., Burgess, S., Danesh, J., Butterworth, A.S., and Staley, J.R. (2019). PhenoScanner V2: an expanded tool for searching human genotype–phenotype associations. *Bioinformatics* 35, 4851–4853.
34. Finan, C., Gaulton, A., Kruger, F.A., Lumbers, R.T., Shah, T., Engmann, J., Galver, L., Kelley, R., Karlsson, A., Santos, R., et al. (2017). The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* 9, eaag1166.
35. Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J., et al. (2023). The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res.* 51, D977–D985.
36. Hemani, G., Zheng, J., Elsworth, B., Wade, K.H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., et al. (2018). The MR-base platform supports systematic causal inference across the human phenome. *Elife* 7, e34408.
37. Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* 44, 512–525.
38. Relton, C.L., and Davey Smith, G. (2012). Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int. J. Epidemiol.* 41, 161–176.
39. Leyden, G.M., Shapland, C.Y., Davey Smith, G., Sanderson, E., Greenwood, M.P., Murphy, D., and Richardson, T.G. (2022). Harnessing tissue-specific genetic variation to dissect putative causal pathways between body mass index and cardiometabolic phenotypes. *Am. J. Hum. Genet.* 109, 240–252.
40. Yengo, L., Sidorenko, J., Kemper, K.E., Zheng, Z., Wood, A.R., Weedon, M.N., Frayling, T.M., Hirschhorn, J., Yang, J., Visscher, P.M.; and GIANT Consortium (2018). Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum. Mol. Genet.* 27, 3641–3649.
41. Bowden, J., Davey Smith, G., Haycock, P.C., and Burgess, S. (2016). Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* 40, 304–314.
42. Frías, J.P. (2020). Tirzepatide: a glucose-dependent insulinotropic polypeptide (GIP) and glucagon-like peptide-1 (GLP-1) dual agonist in development for the treatment of type 2 diabetes. *Expert Rev. Endocrinol. Metab.* 15, 379–394.
43. Braselmann, S., Taylor, V., Zhao, H., Wang, S., Sylvain, C., Baluom, M., Qu, K., Herlaar, E., Lau, A., Young, C., et al. (2006). R406, an orally available spleen tyrosine kinase inhibitor blocks fc receptor signaling and reduces immune complex-mediated inflammation. *J. Pharmacol. Exp. Ther.* 319, 998–1008.
44. Tanaka, Y., Millson, D., Iwata, S., and Nakayamada, S. (2021). Safety and efficacy of fostamatinib in rheumatoid arthritis patients with an inadequate response to methotrexate in phase II OSKIRA-ASIA-1 and OSKIRA-ASIA-1X study. *Rheumatology* 60, 2884–2895.
45. Ray, L.A., Chin, P.F., and Miotto, K. (2010). Naltrexone for the treatment of alcoholism: clinical findings, mechanisms of action, and pharmacogenetics. *CNS Neurol. Disord.: Drug Targets* 9, 13–22.
46. Onakpoya, I.J., Lee, J.J., Mahtani, K.R., Aronson, J.K., and Heneghan, C.J. (2020). Naltrexone–bupropion (Mysimba) in management of obesity: A systematic review and meta-analysis of unpublished clinical study reports. *Br. J. Clin. Pharmacol.* 86, 646–667.
47. Gauthier, B.R., Sola-García, A., Cáliz-Molina, M.Á., Lorenzo, P.I., Cobo-Vuilleumier, N., Capilla-González, V., and Martín-Montalvo, A. (2020). Thyroid hormones in diabetes, cancer, and aging. *Aging Cell* 19, e13260.
48. Boissel, S., Reish, O., Proulx, K., Kawagoe-Takaki, H., Sedgwick, B., Yeo, G.S.H., Meyre, D., Golzio, C., Molinari, F., Kadhom, N., et al. (2009). Loss-of-Function Mutation in the Dioxygenase-Encoding FTO Gene Causes Severe Growth Retardation and Multiple Malformations. *Am. J. Hum. Genet.* 85, 106–111.
49. Smemo, S., Tena, J.J., Kim, K.H., Gamazon, E.R., Sakabe, N.J., Gómez-Marín, C., Aneas, I., Credidio, F.L., Sobreira, D.R., Wasserman, N.F., et al. (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* 507, 371–375.
50. Hartley, A., Sanderson, E., Granell, R., Paternoster, L., Zheng, J., Smith, G.D., Southam, L., Hatzikotoulas, K., Boer, C.G., van Meurs, J., et al. (2022). Using multivariable Mendelian randomization to estimate the causal effect of bone mineral density on osteoarthritis risk, independently of body mass index. *Int. J. Epidemiol.* 51, 1254–1267.
51. Del Bosque-Plata, L., Martínez-Martínez, E., Espinoza-Camacho, M.Á., and Gagnoli, C. (2021). The Role of TCF7L2 in Type 2 Diabetes. *Diabetes* 70, 1220–1228.
52. Rockel, J.S., Yu, C., Whetstone, H., Craft, A.M., Reilly, K., Ma, H., Tsumishima, H., Puvindran, V., Al-Jazrawe, M., Keller, G.M., and Alman, B.A. (2016). Hedgehog inhibits β -catenin activity in synovial joint development and osteoarthritis. *J. Clin. Invest.* 126, 1649–1663.
53. Jin, T. (2016). Current Understanding on Role of the Wnt Signaling Pathway Effector TCF7L2 in Glucose Homeostasis. *Endocr. Rev.* 37, 254–277.
54. Blom, A.B., van Lent, P.L., van der Kraan, P.M., and van den Berg, W.B. (2010). To Seek Shelter from the Wnt in Osteoarthritis? Wnt-Signaling as a Target for Osteoarthritis Therapy. *Curr. Drug Targets* 11, 620–629.
55. Lotta, L.A., Wittemans, L.B.L., Zuber, V., Stewart, I.D., Sharp, S.J., Luan, J., Day, F.R., Li, C., Bowker, N., Cai, L., et al. (2018). Association of Genetic Variants Related to Gluteofemoral vs Abdominal Fat Distribution With Type 2 Diabetes, Coronary Disease, and Cardiovascular Risk Factors. *JAMA* 320, 2553–2563.
56. Pei, Y.-F., Liu, Y.-Z., Yang, X.-L., Zhang, H., Feng, G.-J., Wei, X.-T., and Zhang, L. (2020). The genetic architecture of appendicular lean mass characterized by association analysis in the UK Biobank study. *Commun. Biol.* 3, 608.
57. Hill, W.D., Weiss, A., Liewald, D.C., Davies, G., Porteous, D.J., Hayward, C., McIntosh, A.M., Gale, C.R., and Deary, I.J. (2020). Genetic contributions to two special factors of neuroticism are associated with affluence, higher intelligence, better health, and longer life. *Mol. Psychiatry* 25, 3034–3052.
58. Mahajan, A., Wessel, J., Willems, S.M., Zhao, W., Robertson, N.R., Chu, A.Y., Gan, W., Kitajima, H., Taliun, D., Rayner, N.W., et al. (2018). Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nat. Genet.* 50, 559–571.
59. Taneera, J., Fadista, J., Ahlqvist, E., Atac, D., Ottosson-Laakso, E., Wollheim, C.B., and Groop, L. (2015). Identification of novel genes for glucose metabolism based upon expression

- pattern in human islets and effect on insulin secretion and glycemia. *Hum. Mol. Genet.* *24*, 1945–1955.
60. Mizuhashi, K., Kanamoto, T., Ito, M., Moriishi, T., Muranishi, Y., Omori, Y., Terada, K., Komori, T., and Furukawa, T. (2012). OBIF, an osteoblast induction factor, plays an essential role in bone formation in association with osteoblastogenesis. *Dev. Growth Differ.* *54*, 474–480.
 61. Harada, S.I., and Rodan, G.A. (2003). Control of osteoblast function and regulation of bone mass. *Nature* *423*, 349–355.
 62. Maruotti, N., Corrado, A., and Cantatore, F.P. (2017). Osteoblast role in osteoarthritis pathogenesis. *J. Cell. Physiol.* *232*, 2957–2963.
 63. Hernandez Cordero, A.I., Gonzales, N.M., Parker, C.C., Sokolof, G., Vandenberg, D.J., Cheng, R., Abney, M., Sko, A., Douglas, A., Palmer, A.A., et al. (2019). Genome-wide Associations Reveal Human-Mouse Genetic Convergence and Modifiers of Myogenesis, CPNE1 and STC2. *Am. J. Hum. Genet.* *105*, 1222–1236.
 64. Ochoa, D., Karim, M., Ghossaini, M., Hulcoop, D.G., McDonagh, E.M., and Dunham, I. (2022). Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. *Nat. Rev. Drug Discov.* *21*, 551.
 65. Fewell, Z., Davey Smith, G., and Sterne, J.A.C. (2007). The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am. J. Epidemiol.* *166*, 646–655.
 66. Picke, A.K., Campbell, G., Napoli, N., Hofbauer, L.C., and Rauer, M. (2019). Update on the impact of type 2 diabetes mellitus on bone metabolism and material properties. *Endocr. Connect.* *8*, R55–R70.
 67. Berendsen, A.D., and Olsen, B.R. (2014). Osteoblast-adipocyte lineage plasticity in tissue development, maintenance and pathology. *Cell. Mol. Life Sci.* *71*, 493–497.

The American Journal of Human Genetics, Volume 110

Supplemental information

**Genetic underpinning of the comorbidity
between type 2 diabetes and osteoarthritis**

Ana Luiza Arruda, April Hartley, Georgia Katsoula, George Davey Smith, Andrew P. Morris, and Eleftheria Zeggini

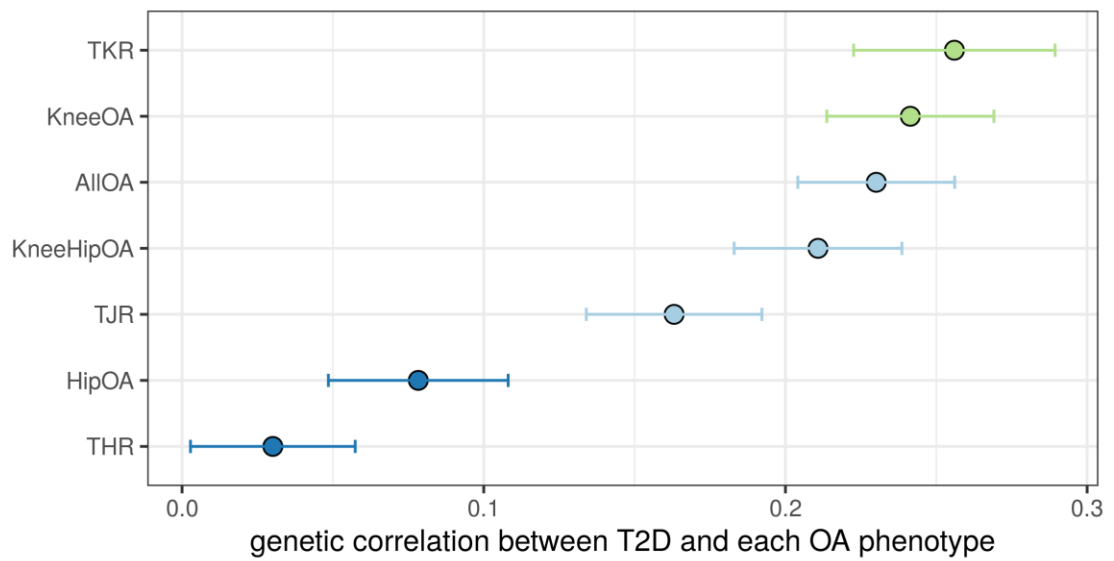
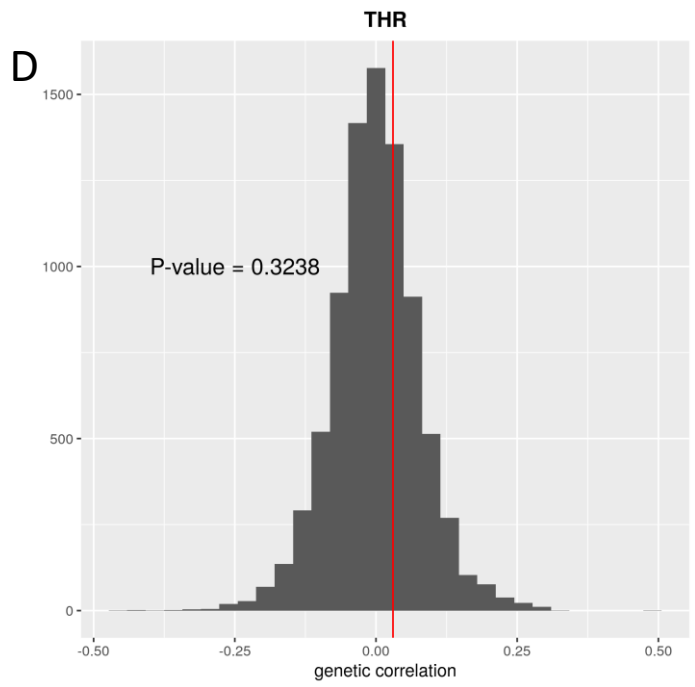
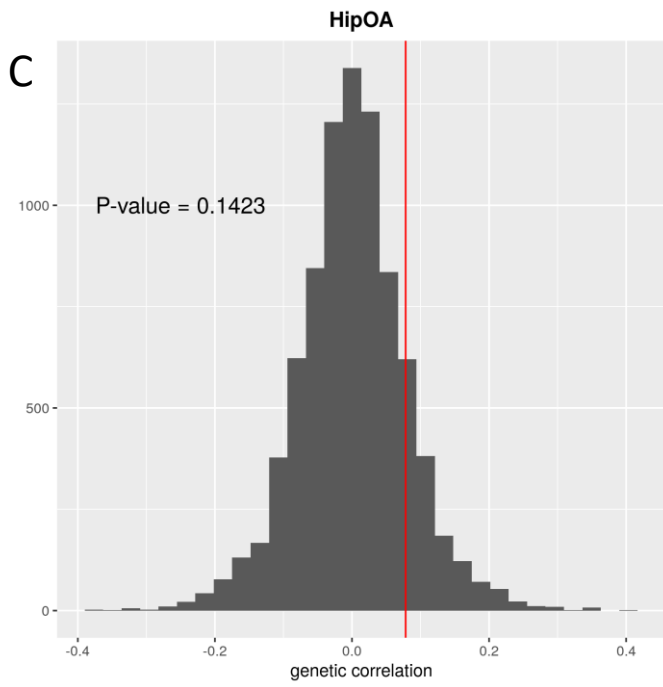
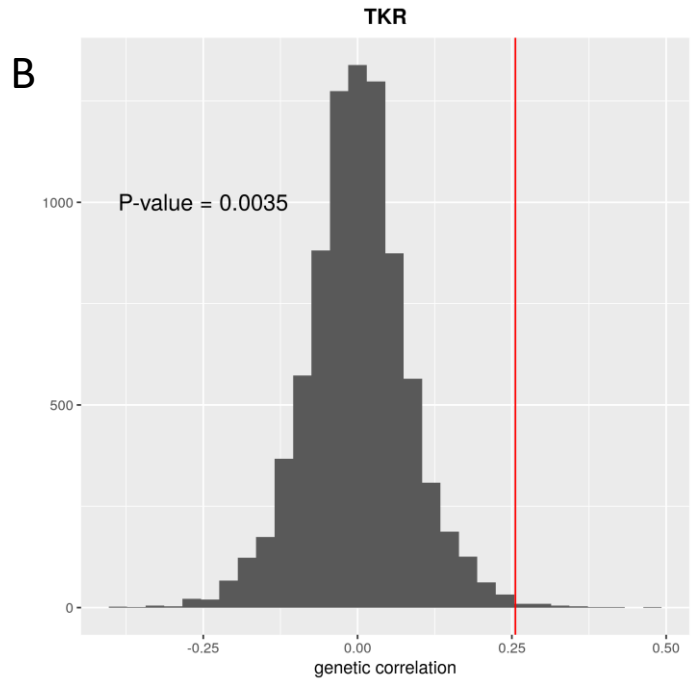
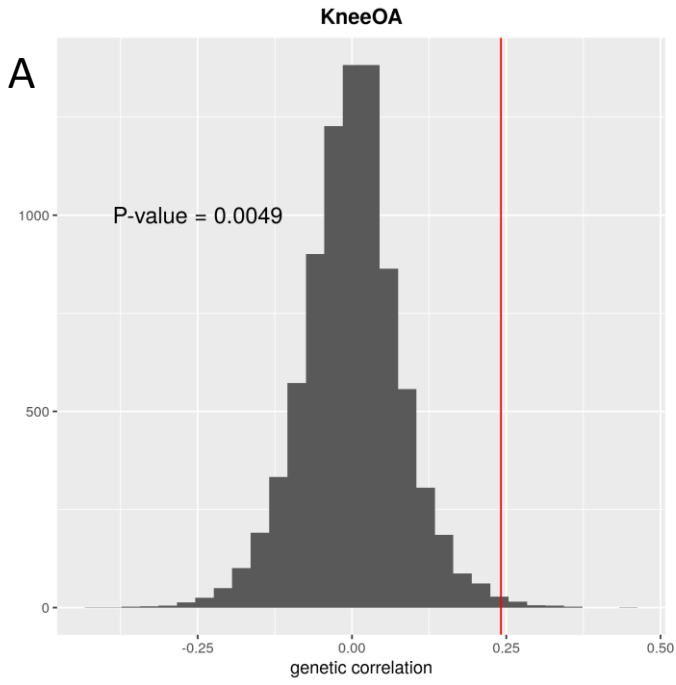


Figure S1: Results of genetic correlation analysis between type 2 diabetes and osteoarthritis. The y-axis depicts the osteoarthritis phenotype and the x-axis the respective genetic correlation. TKR = total knee replacement, THR = total hip replacement, AllOA = osteoarthritis at any site. The error bars represent the standard error of the estimated genetic correlation.



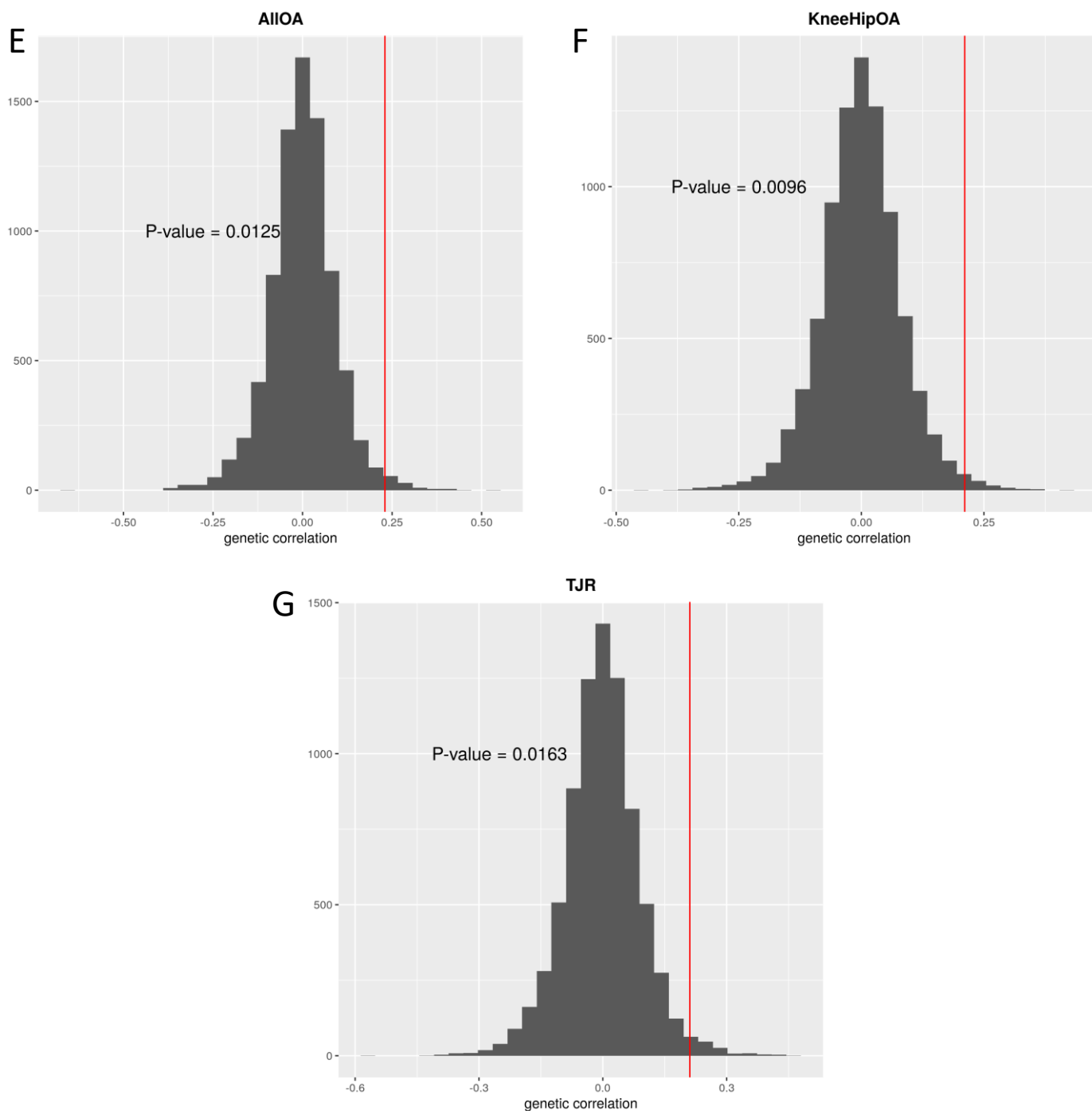


Figure S2: Permutation-based testing of genetic correlation between type 2 diabetes and osteoarthritis. The y-axis depicts the number of tests and the x-axis the genetic correlation of each test. The red line represents the actual genetic correlation. Subfigures: A) genetic correlation between type 2 diabetes and knee osteoarthritis, B) genetic correlation between type 2 diabetes and total knee replacement (TKR), C) genetic correlation between type 2 diabetes and hip osteoarthritis, D) genetic correlation between type 2 diabetes and total hip replacement (THR), E) genetic correlation between type 2 diabetes and osteoarthritis at any site, F) genetic correlation between type 2 diabetes and knee and/or hip osteoarthritis, G) genetic correlation between type 2 diabetes and total joint replacement (TJR).

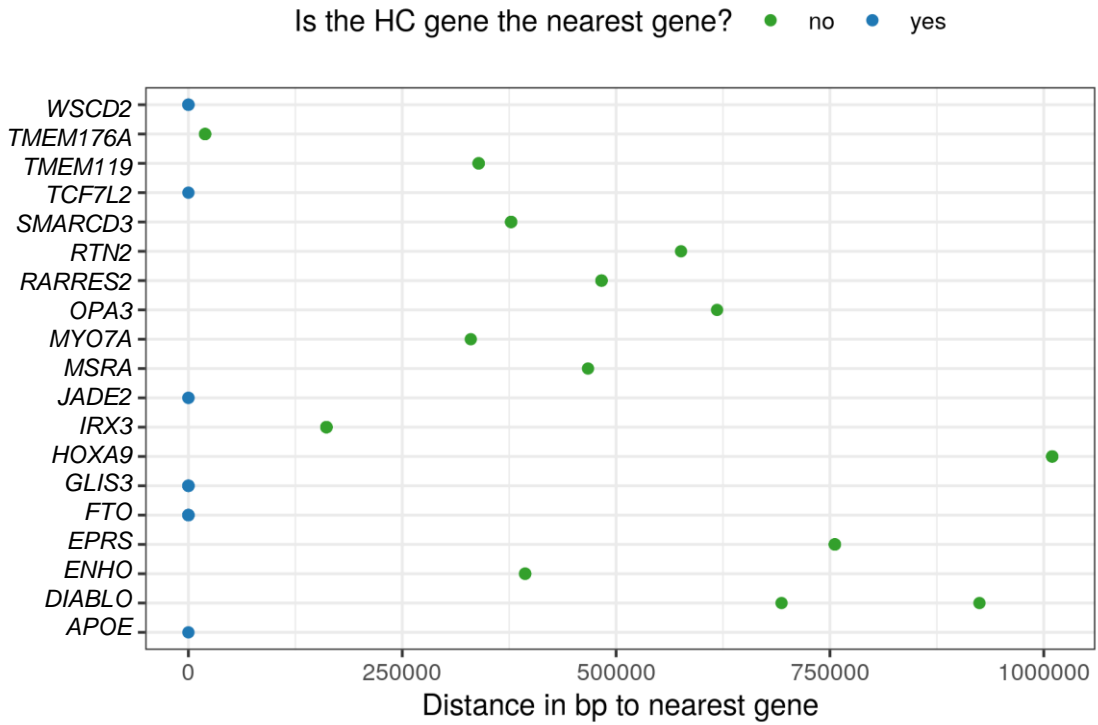


Figure S3: Distance in base pairs (bp) between each high confidence gene and the nearest gene to the lead causal variant in each colocalized region. Blue dots indicate that the high confidence gene is the nearest gene in the region.

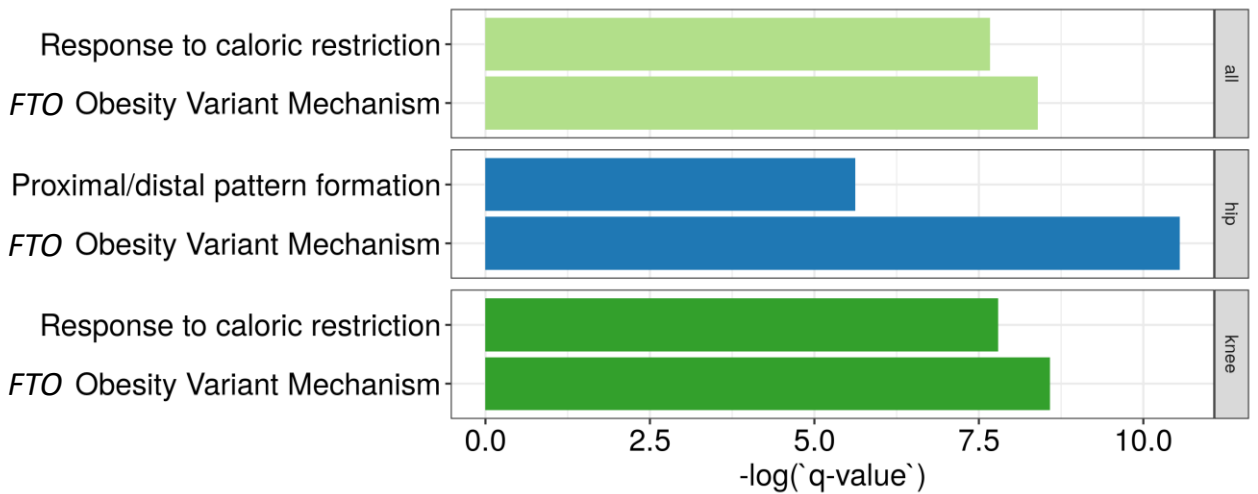
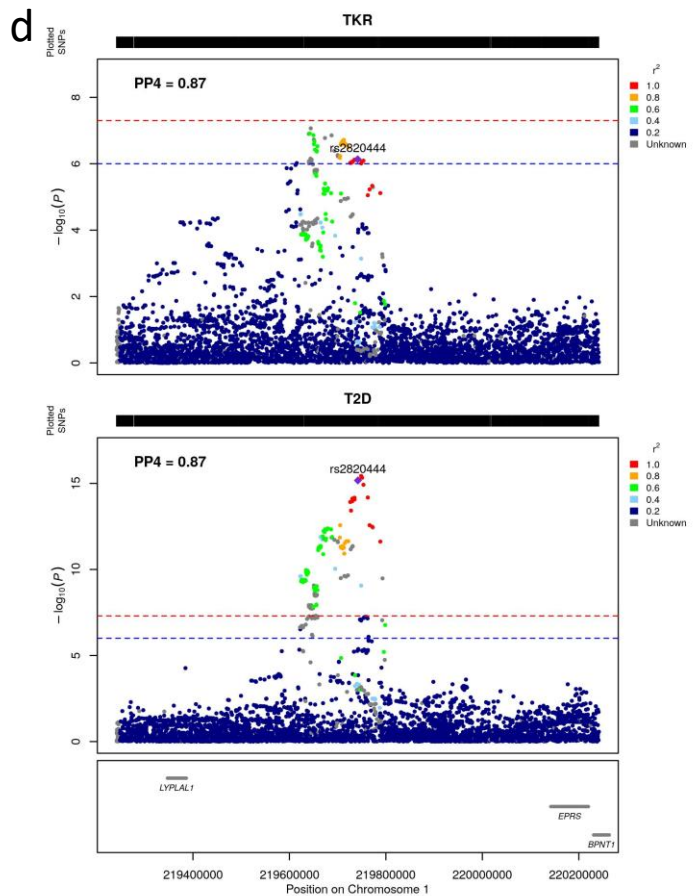
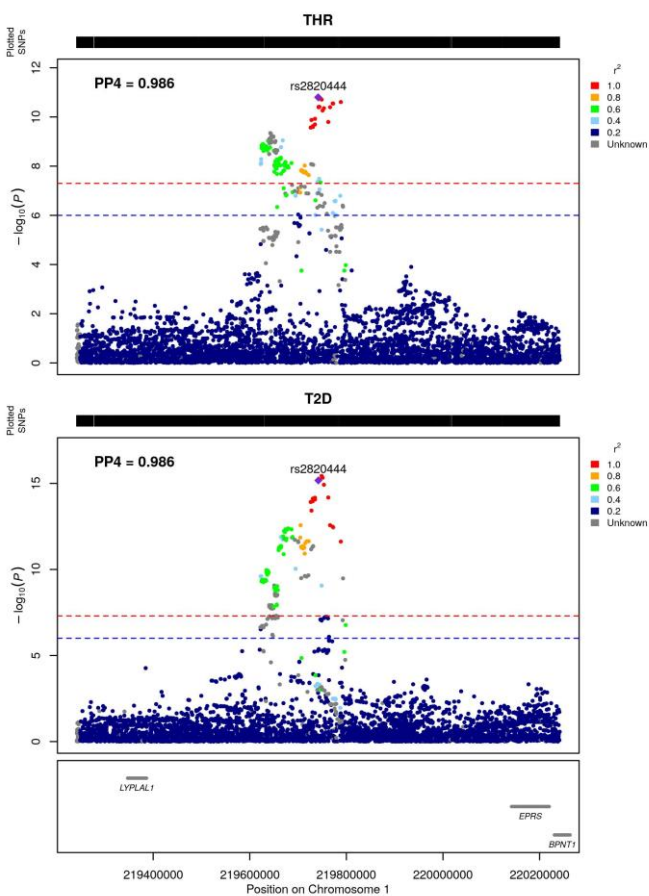
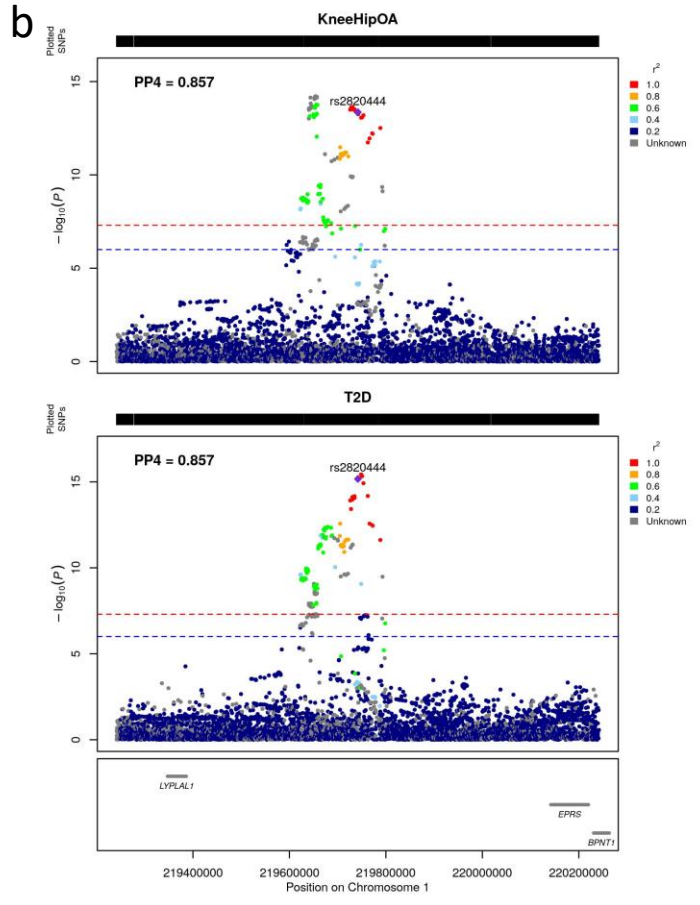
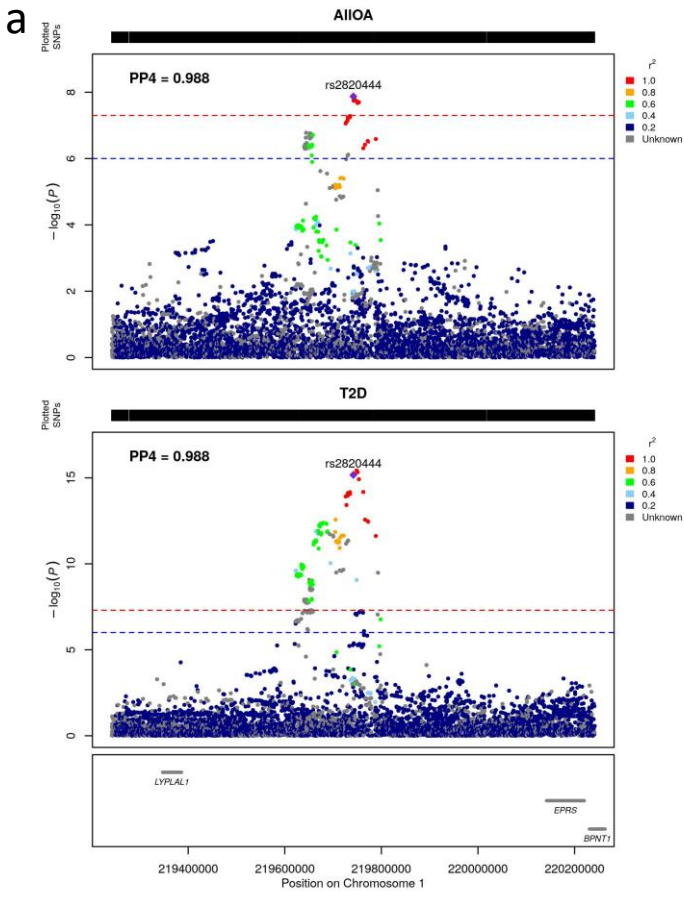


Figure S4: Top pathways enriched for the high confidence genes sets. The y-axis depicts the name of the enriched biological pathways for the different sets of genes: all = 19 high confidence genes, hip = 10 hip-related high confidence genes, knee = 18 knee-related high confidence genes. The x-axis depicts the negative logarithms of the FDR adjusted p-value (q-value).



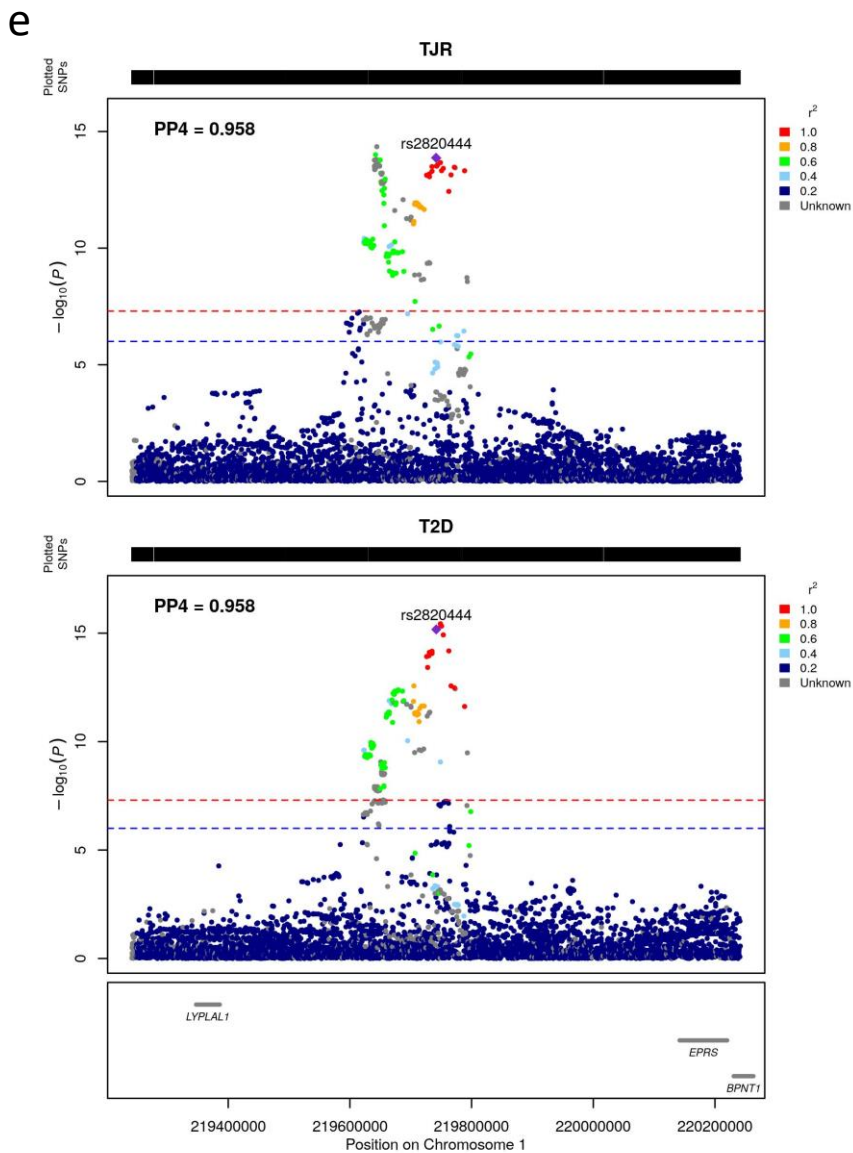


Figure S5: Colocalized region 1. This region colocalizes between type 2 diabetes and a) osteoarthritis at any site , b) knee and/or hip osteoarthritis, c) total hip replacement, d) total knee replacement, e) total joint replacement.

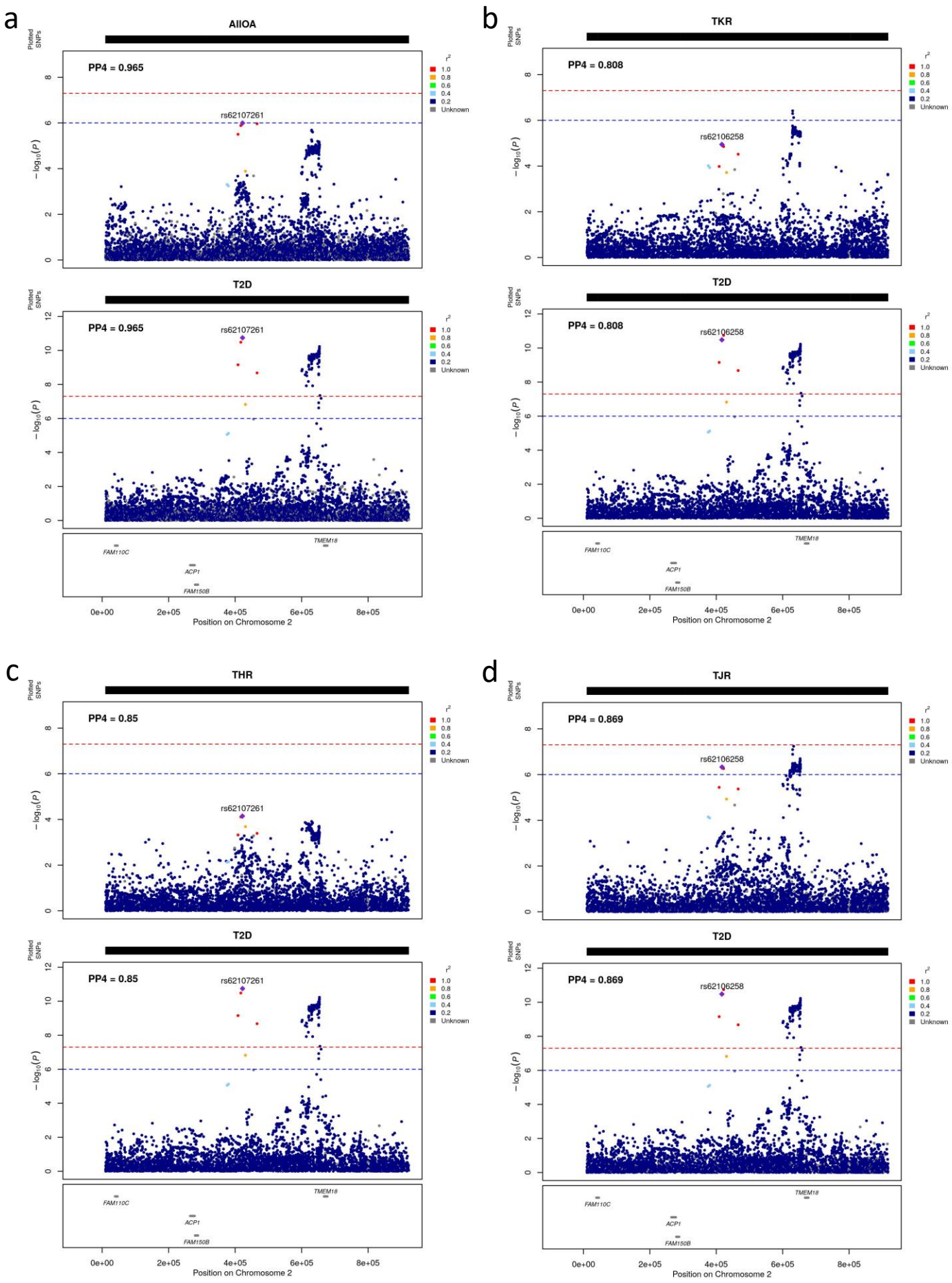


Figure S6: Colocalized region 2. This region colocalizes between type 2 diabetes and a) osteoarthritis at any site , b) total knee replacement, c) total hip replacement, d) total joint replacement.

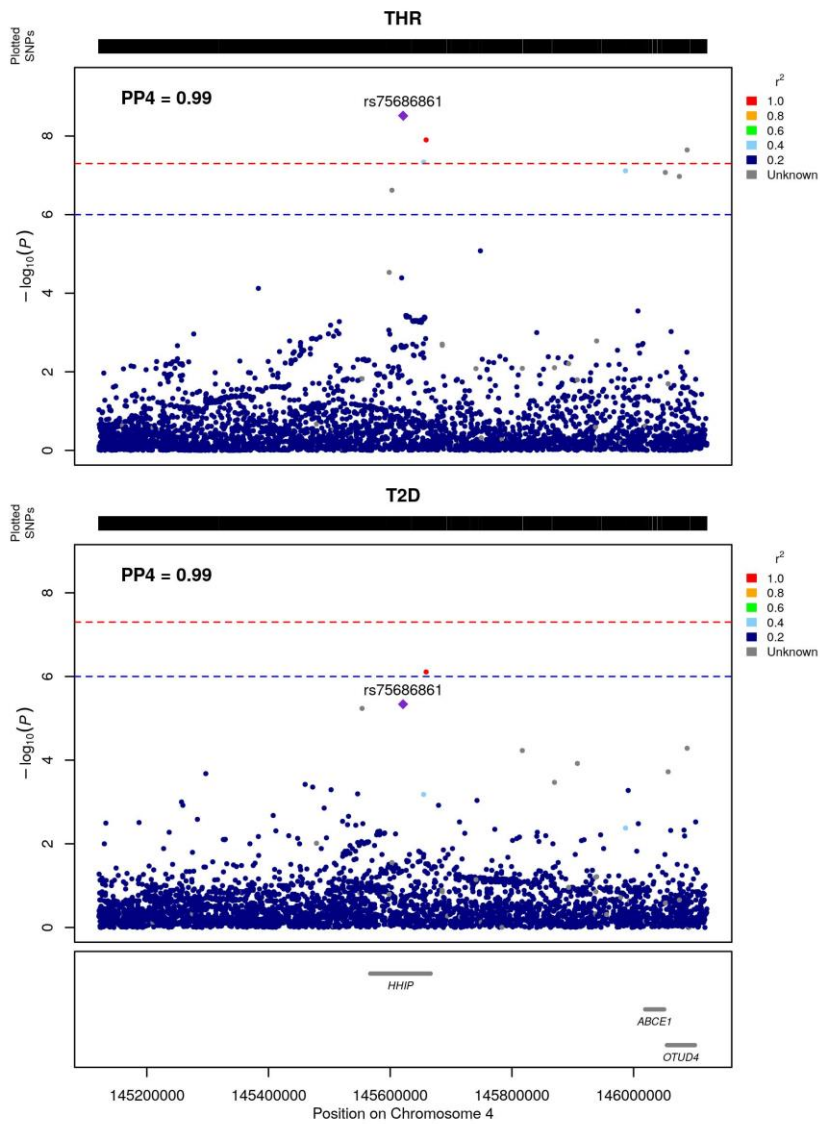


Figure S7: Colocalized region 3. This region colocalizes between type 2 diabetes and total hip replacement.

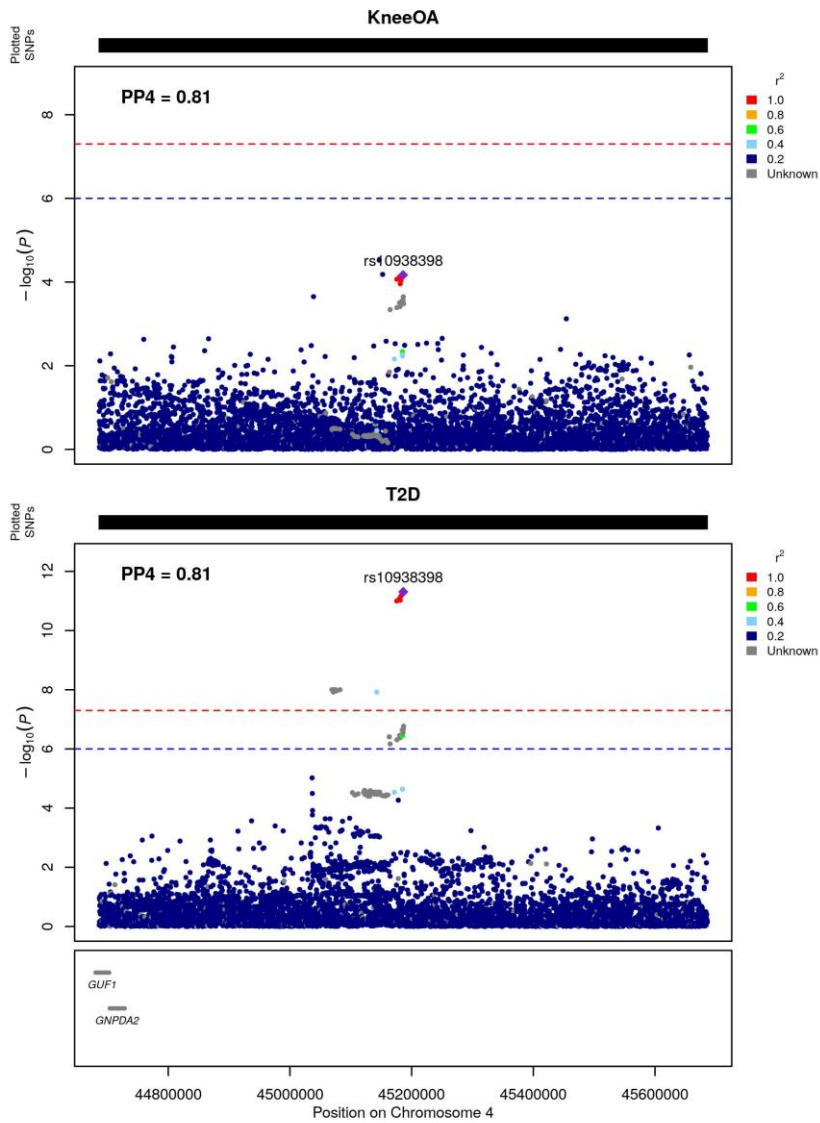


Figure S8: Colocalized region 4. This region colocalizes between type 2 diabetes and knee osteoarthritis.

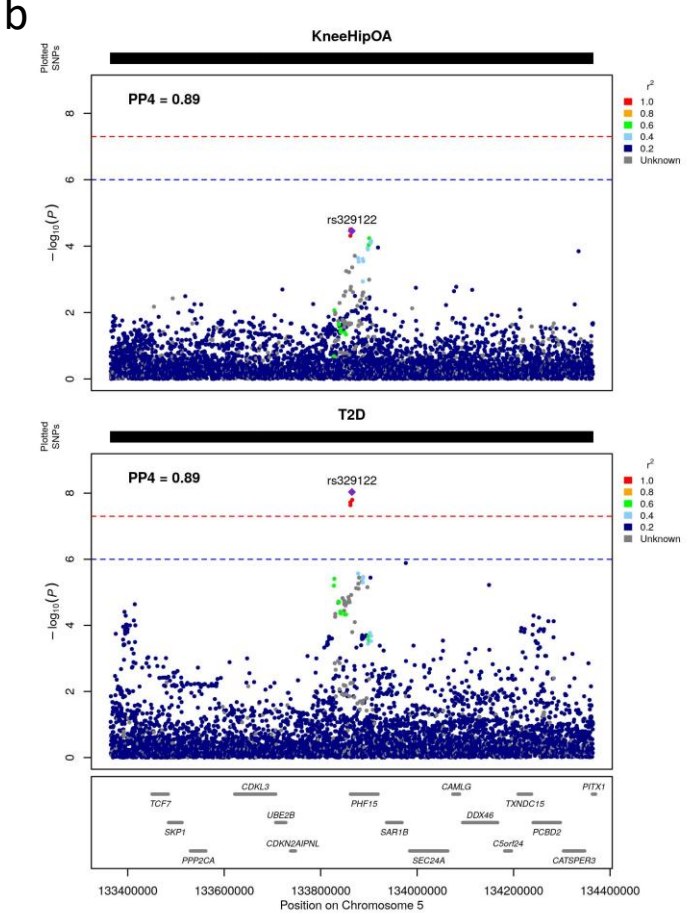
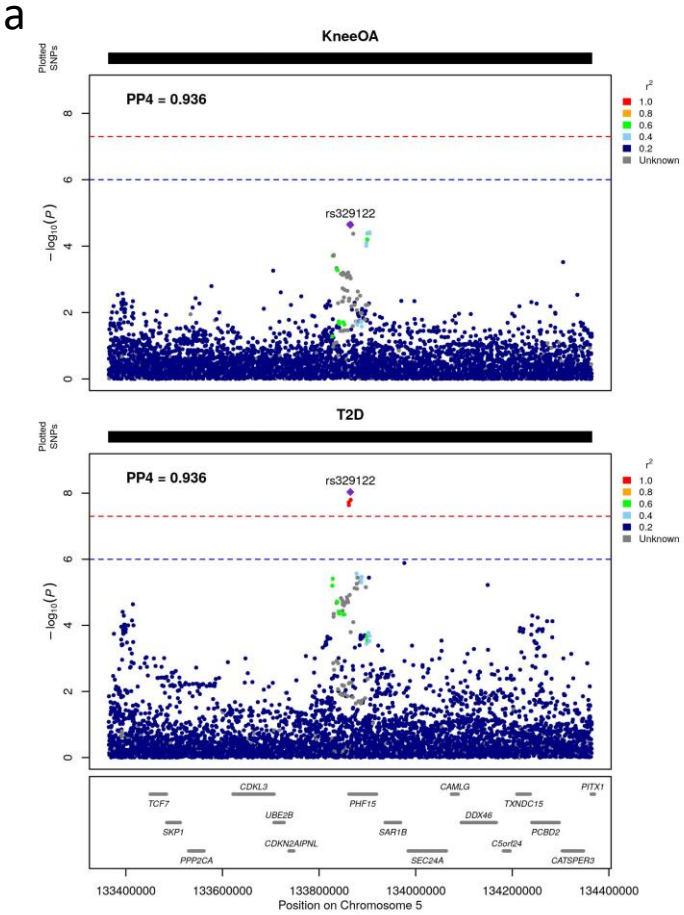


Figure S9: Colocalized region 5. This region colocalizes between type 2 diabetes and a) knee osteoarthritis, b) knee and/or hip osteoarthritis.

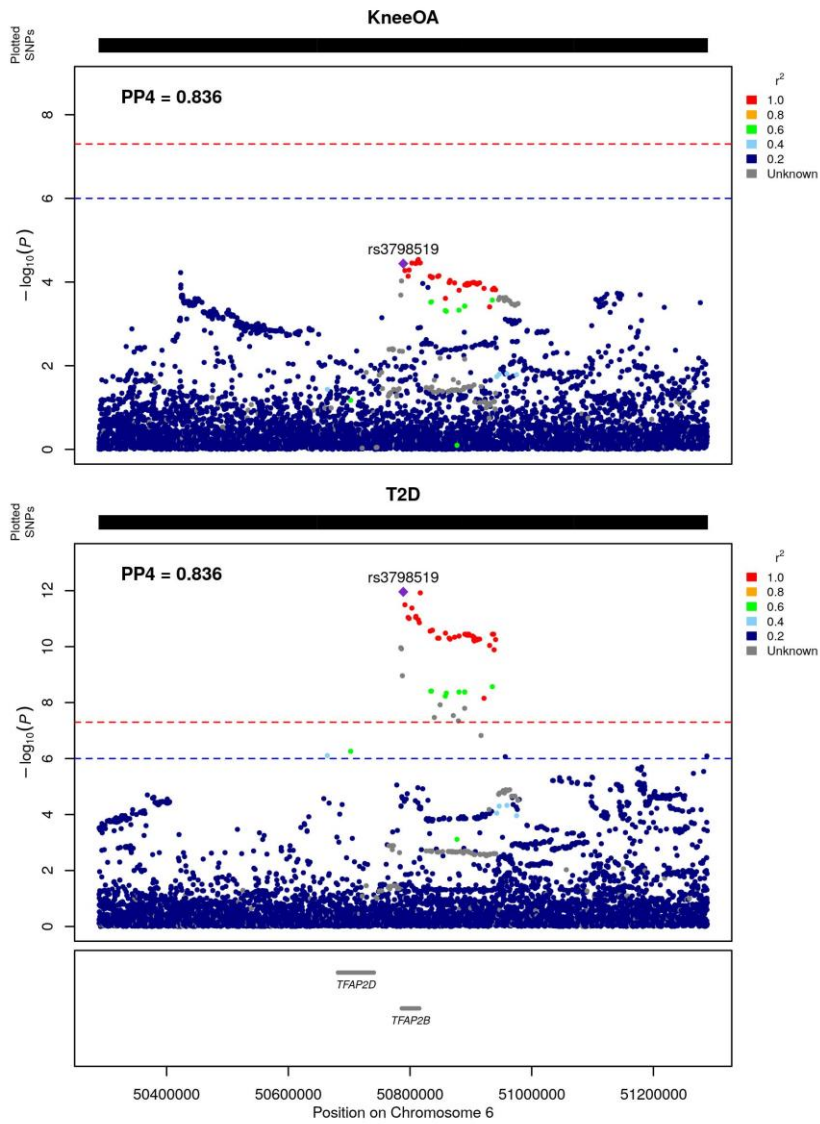


Figure S10: Colocalized region 6. This region colocalizes between type 2 diabetes and knee osteoarthritis.

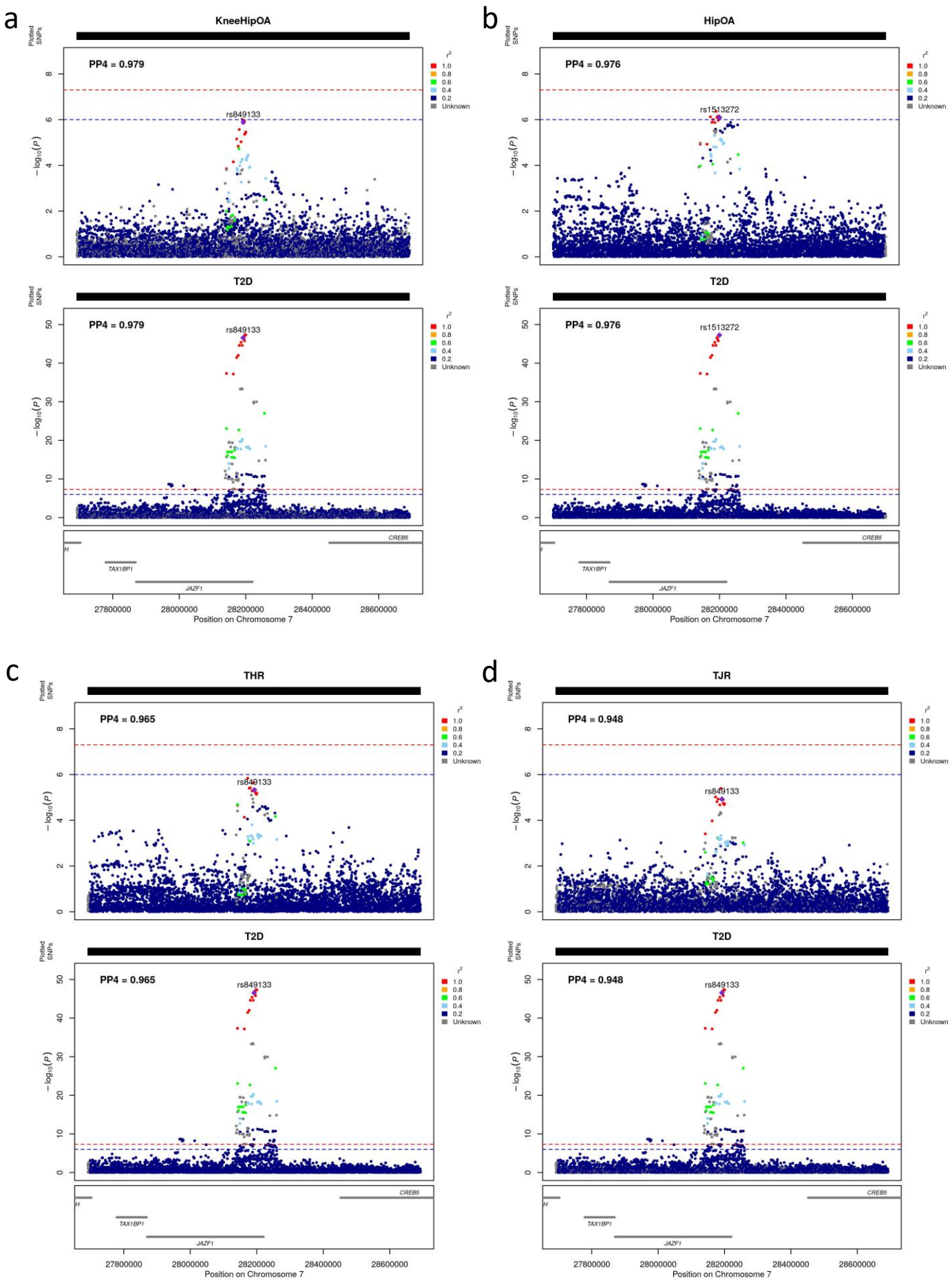


Figure S11: Colocalized region 7. This region colocalizes between type 2 diabetes and a) knee and/or hip osteoarthritis, b) hip osteoarthritis, c) total hip replacement, d) total joint replacement.

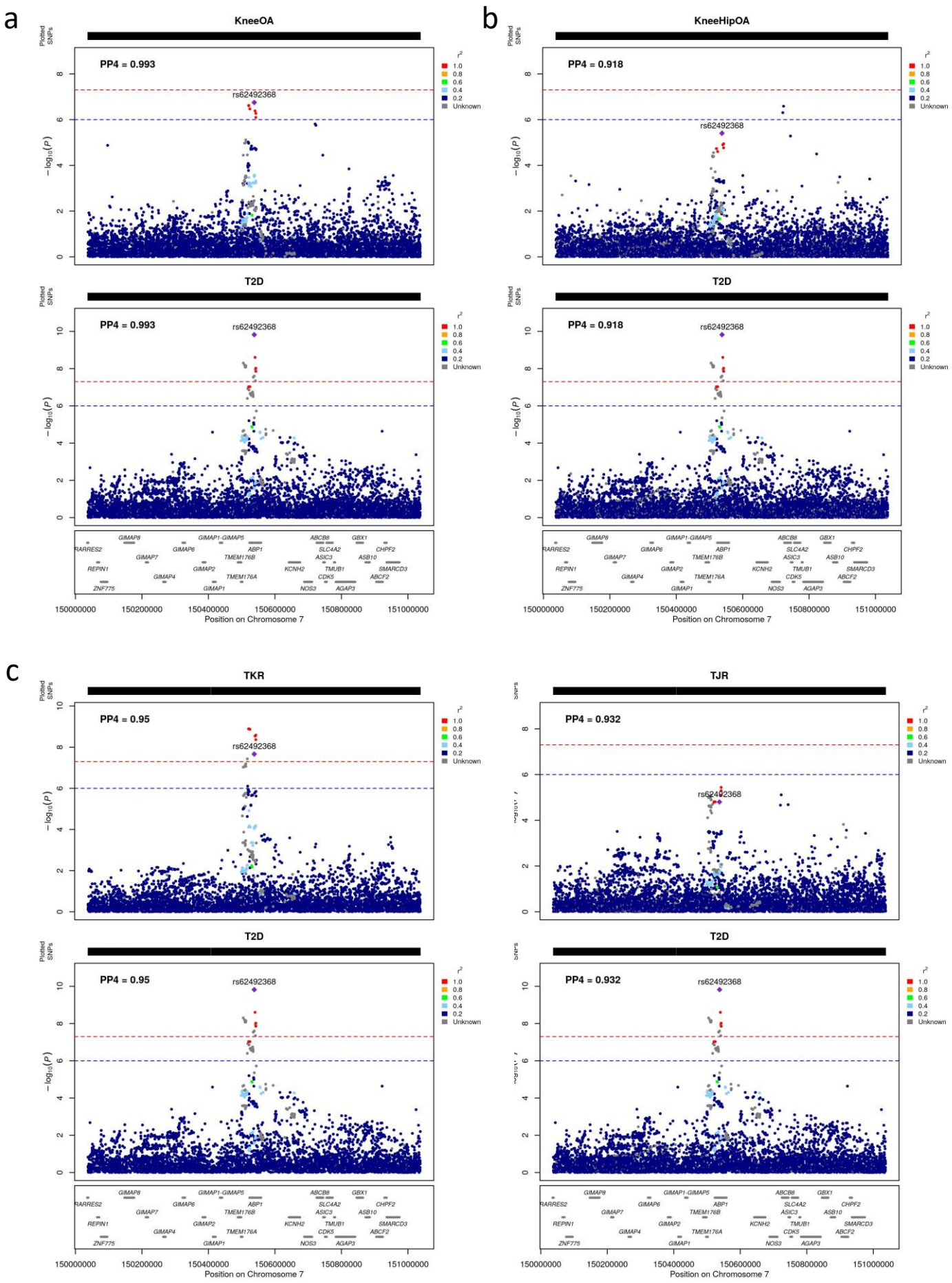


Figure S12: Colocalized region 8. This region colocalizes between type 2 diabetes and a) knee osteoarthritis, b) knee and/or hip osteoarthritis, c) total knee replacement, d) total joint replacement.

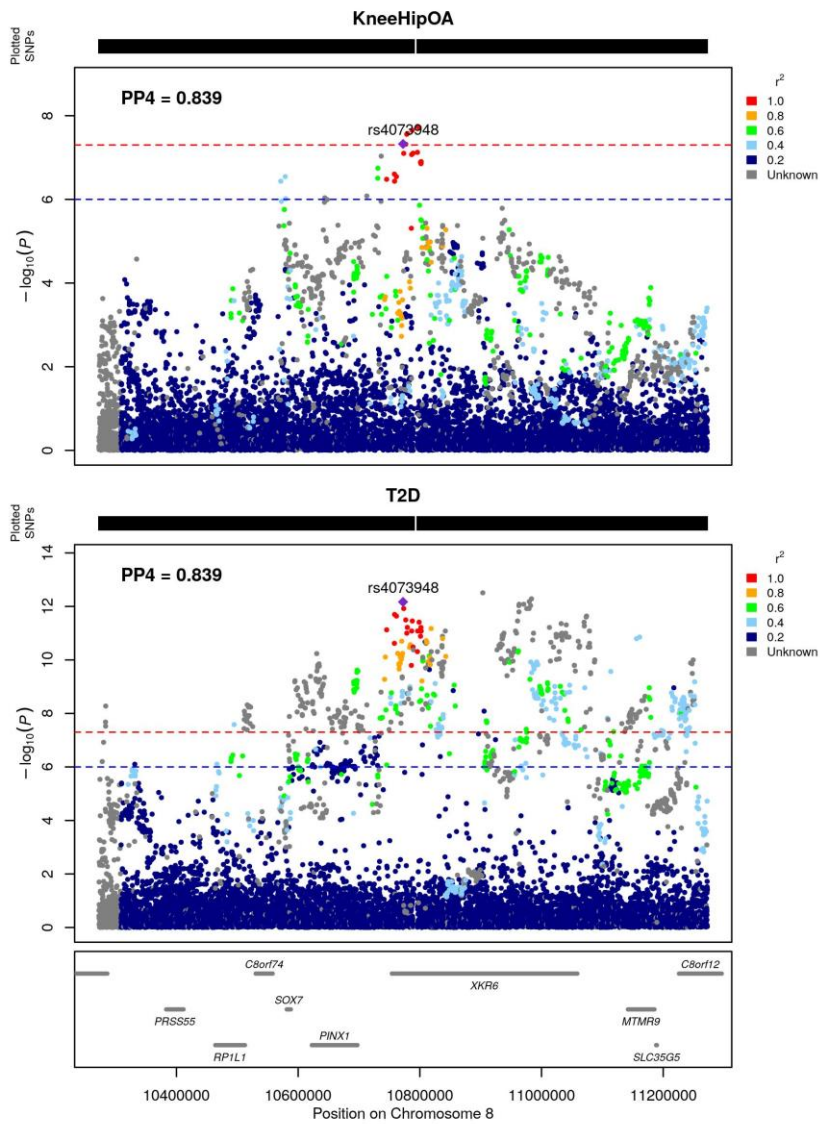
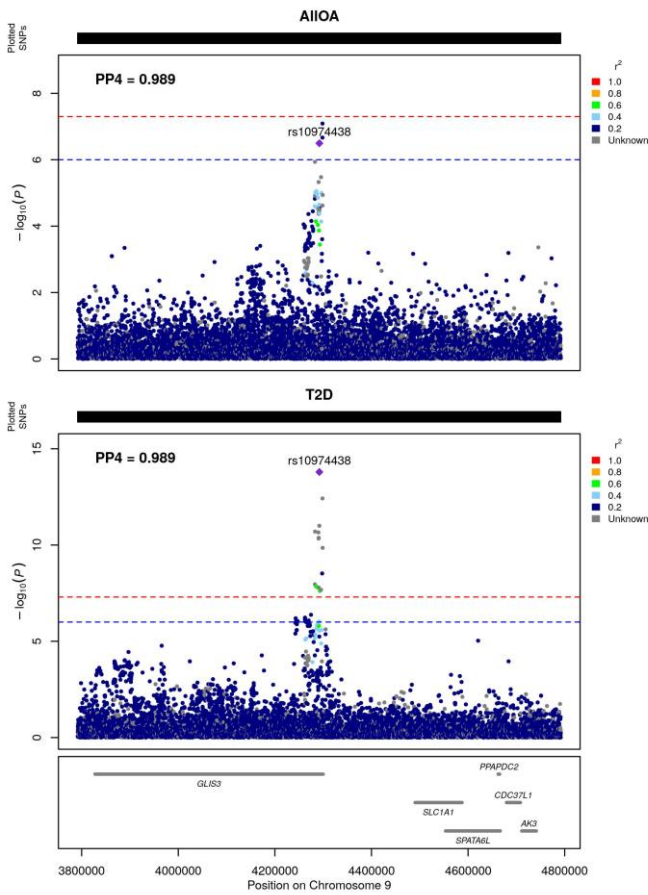
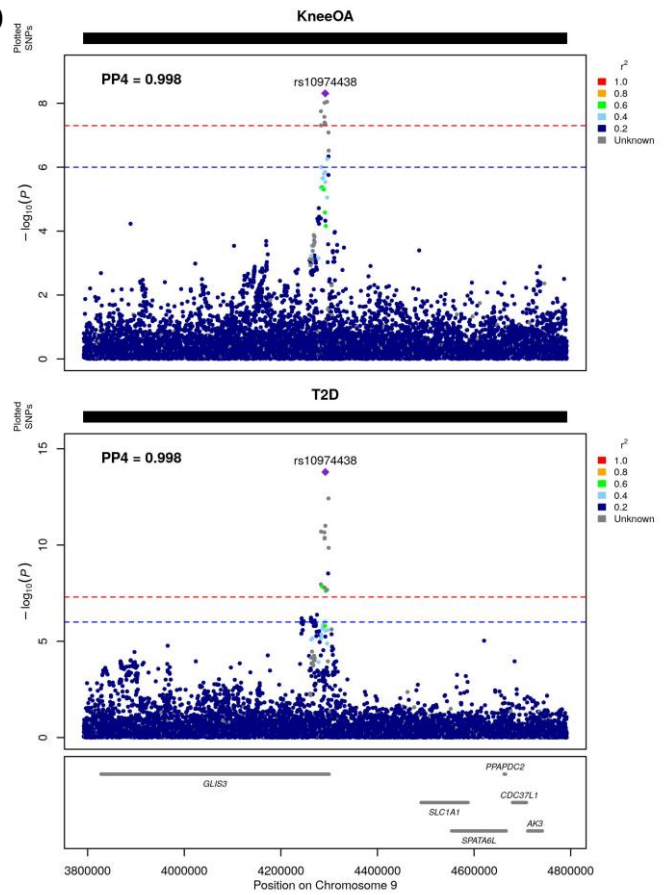
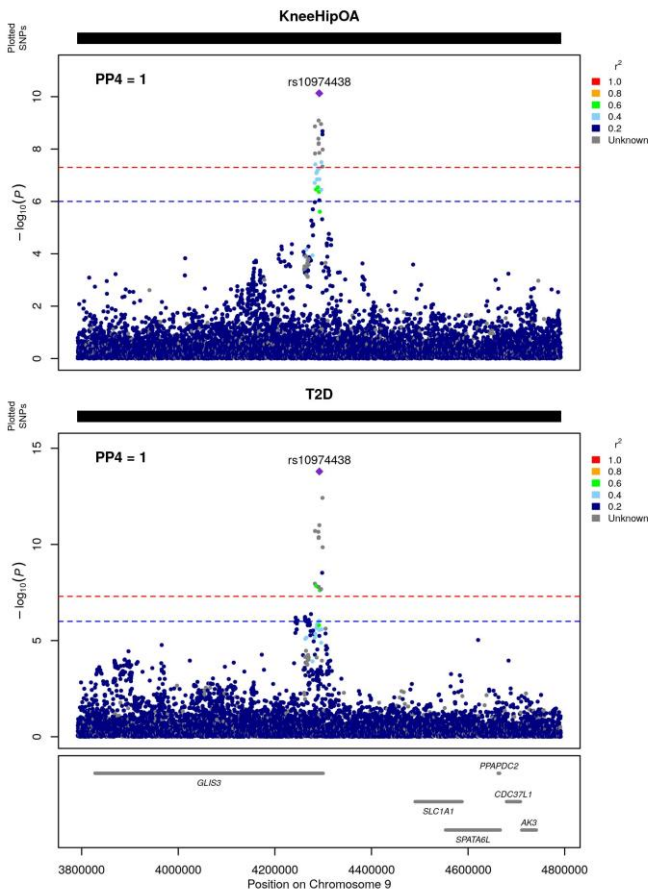
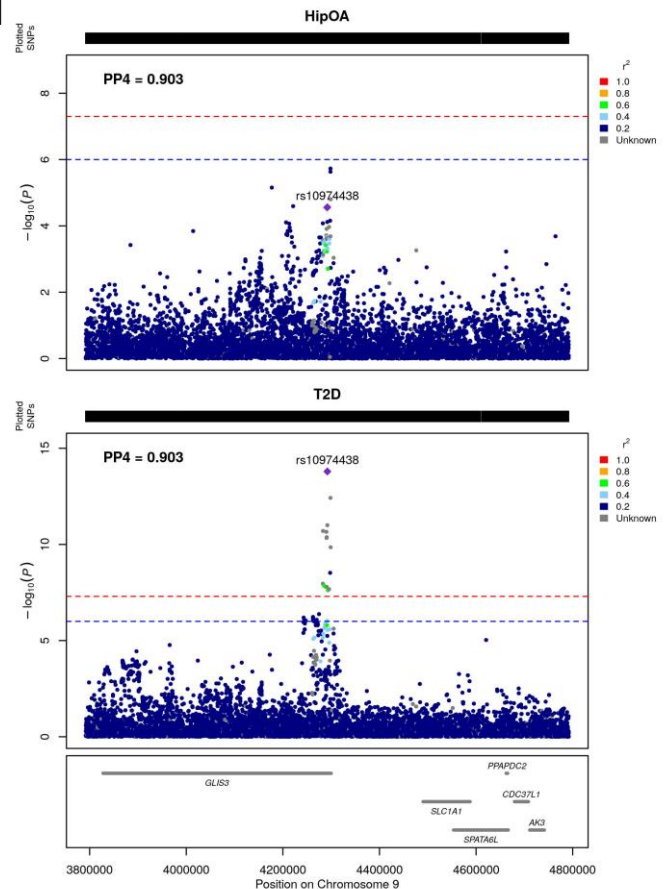


Figure S13: Colocalized region 9. This region colocalizes between type 2 diabetes and knee and/or hip osteoarthritis.

a**b****c****d**

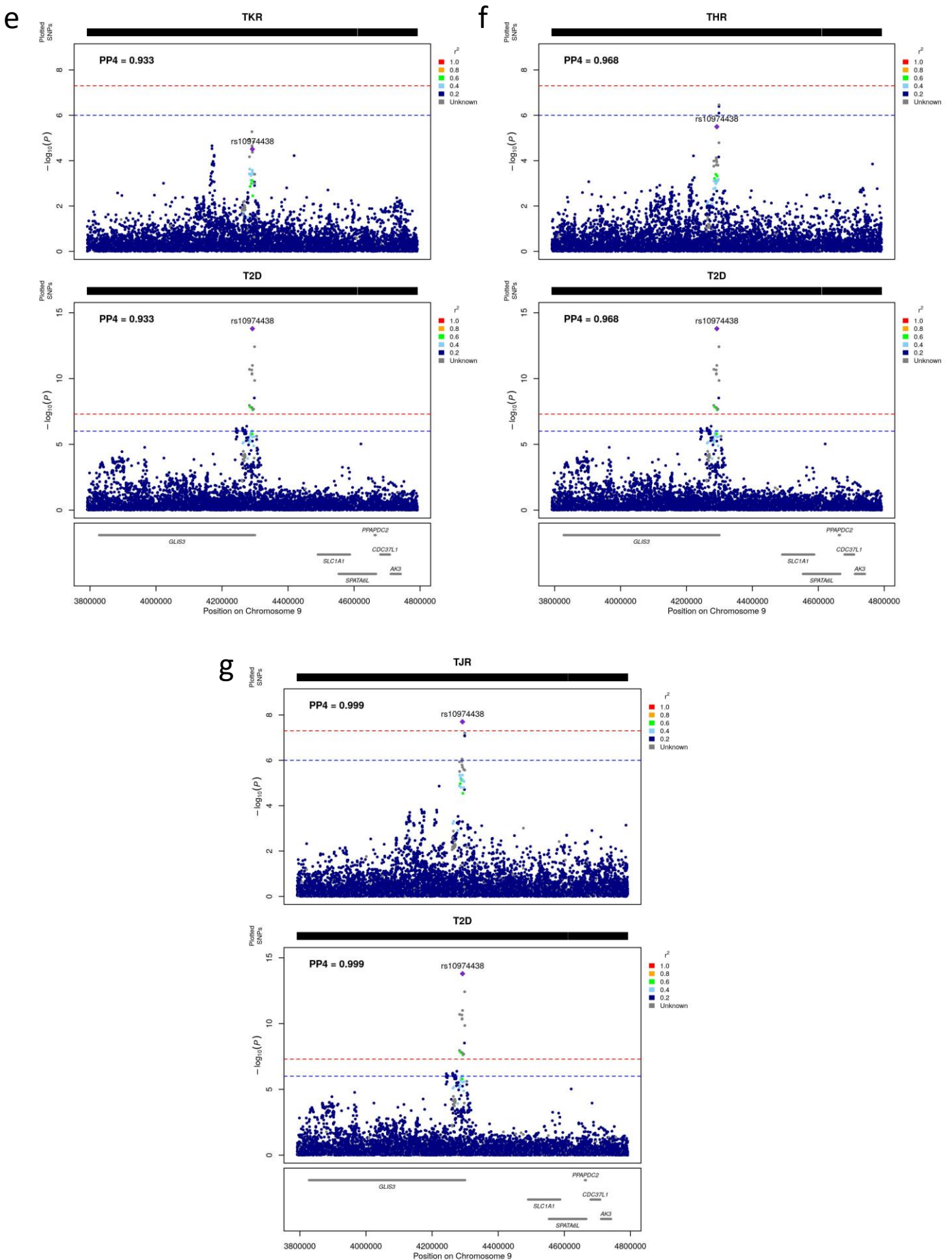


Figure S14: Colocalized region 10. This region colocalizes between type 2 diabetes and a) osteoarthritis at any site, b) knee osteoarthritis, c) knee and/or hip osteoarthritis, d) hip osteoarthritis, e) total knee replacement, f) total hip replacement, g) total joint replacement.

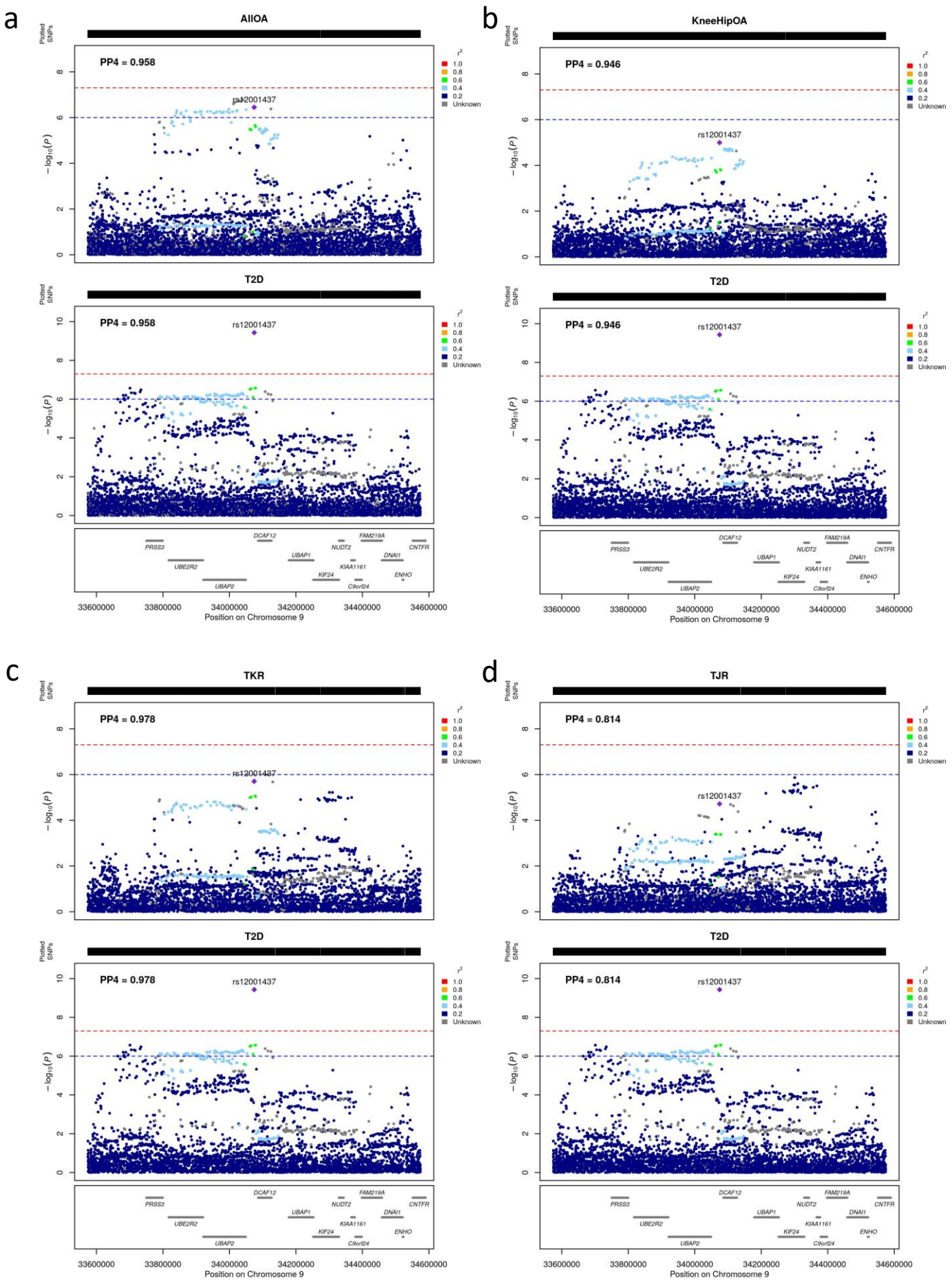


Figure S15: Colocalized region 11. This region colocalizes between type 2 diabetes and a) osteoarthritis at any site, b) knee and/or hip osteoarthritis, c) total knee replacement, d) total joint replacement.

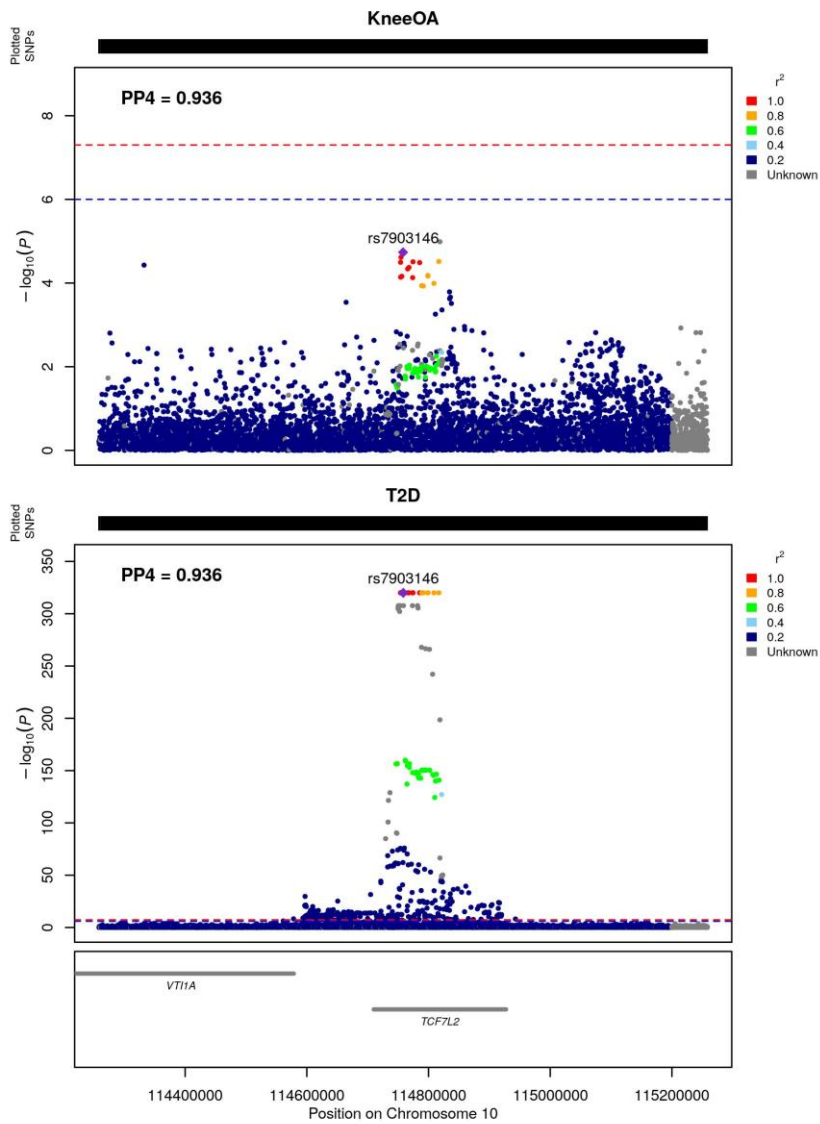


Figure S16: Colocalized region 12. This region colocalizes between type 2 diabetes and knee osteoarthritis.

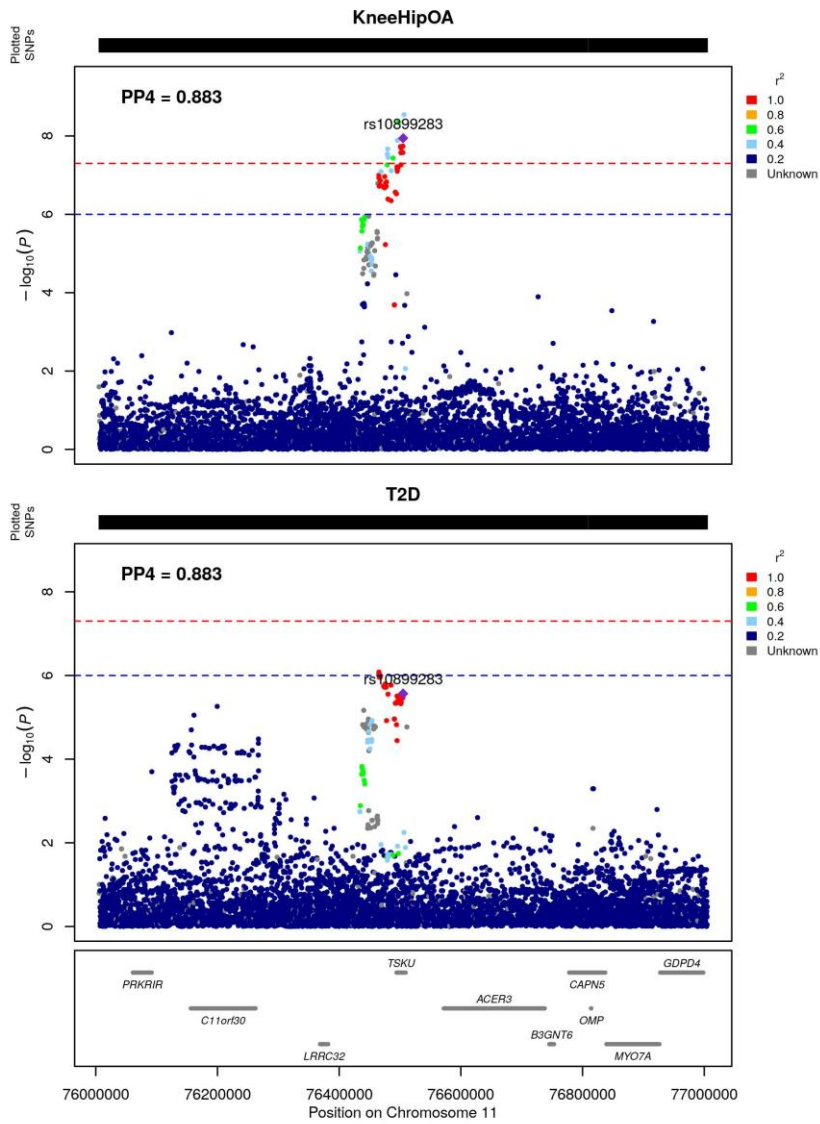


Figure S17: Colocalized region 13. This region colocalizes between type 2 diabetes and knee and/or hip osteoarthritis.

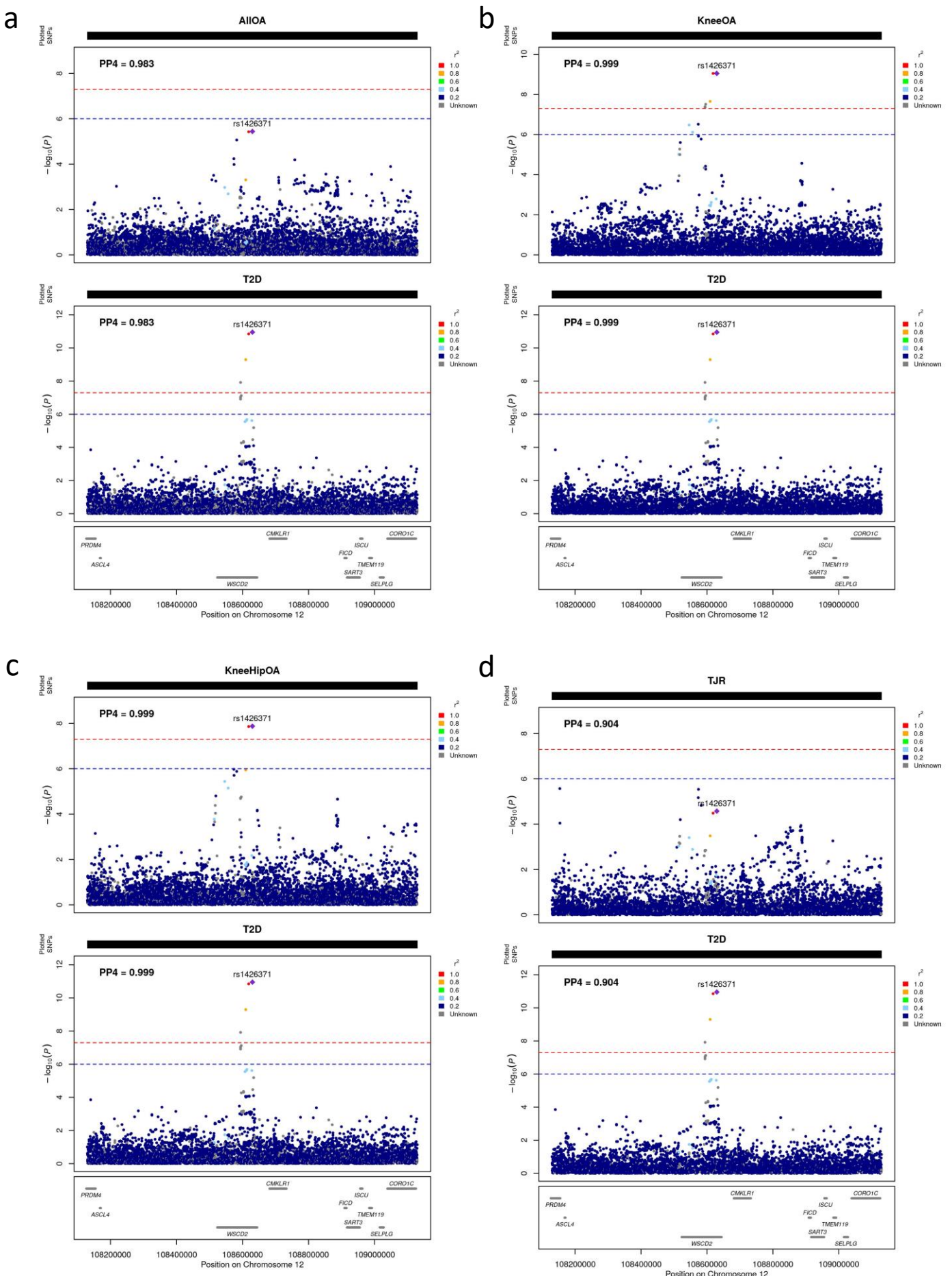


Figure S18: Colocalized region 14. This region colocalizes between type 2 diabetes and a) osteoarthritis at any site, b) knee osteoarthritis, c) knee and/or hip osteoarthritis, d) total joint replacement.

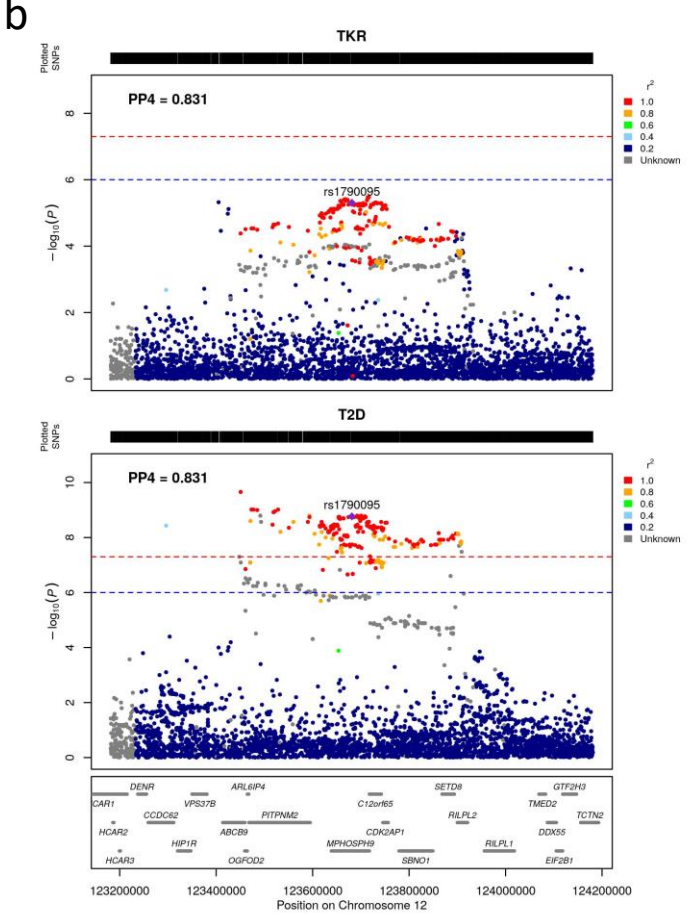
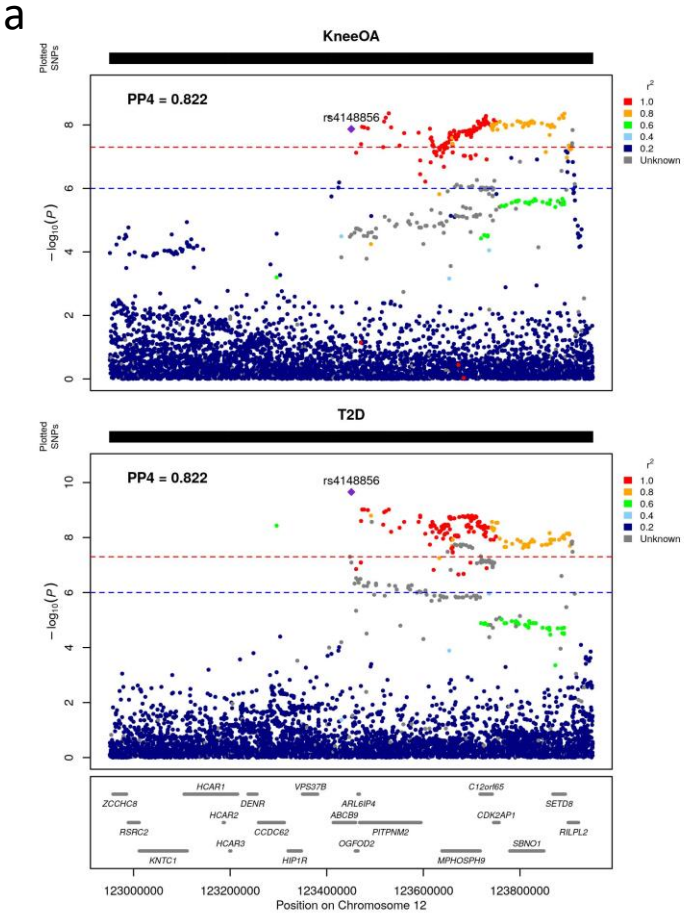


Figure S19: Colocalized region 15. This region colocalizes between type 2 diabetes and a) knee osteoarthritis, b) total knee replacement.

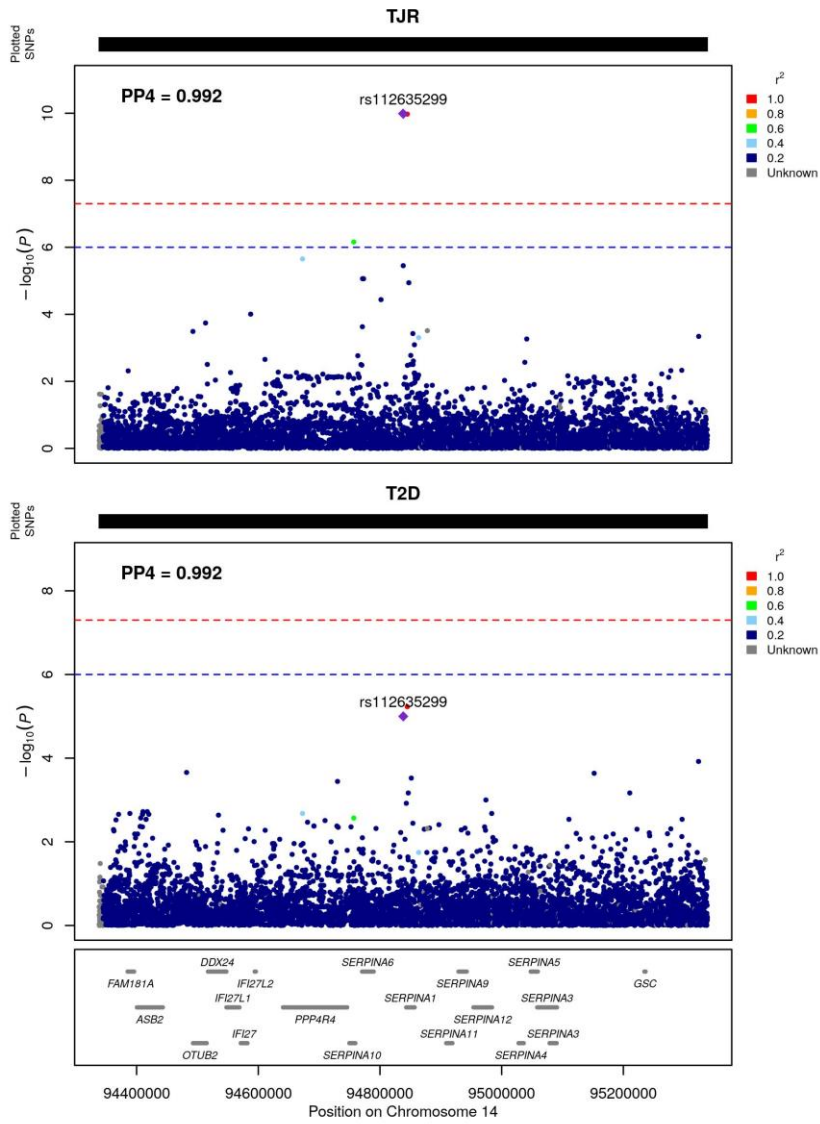
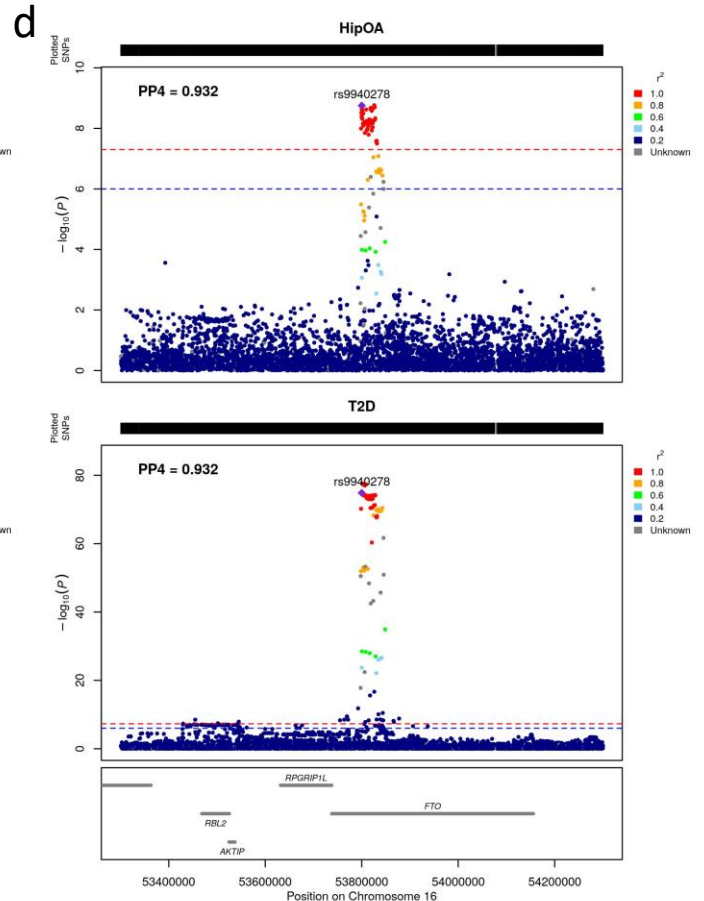
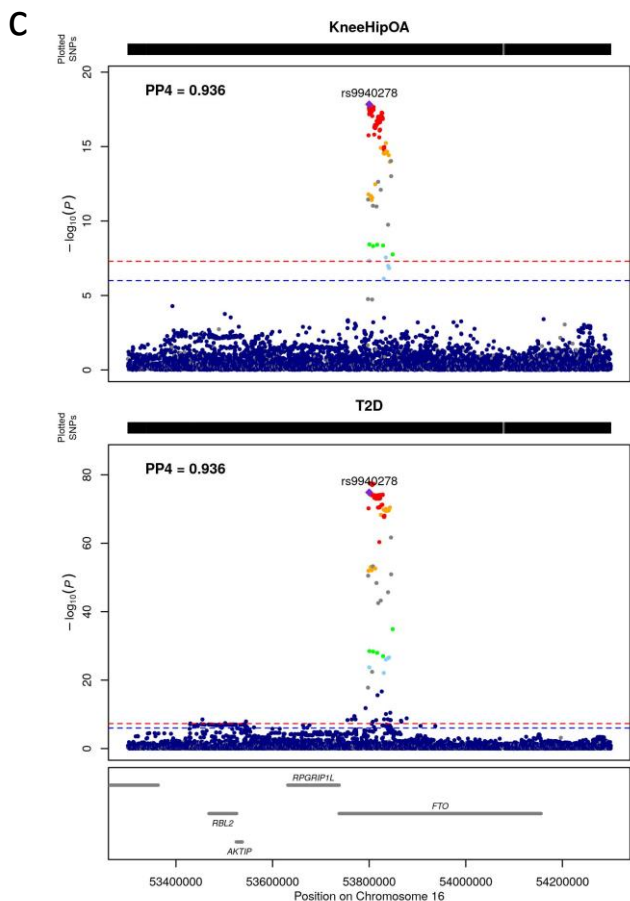
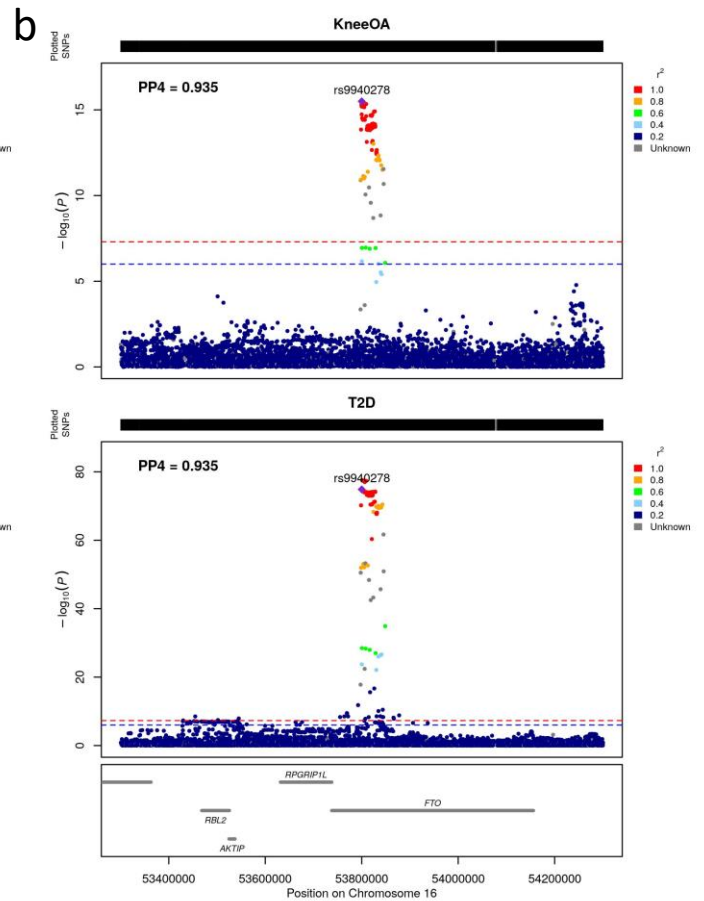
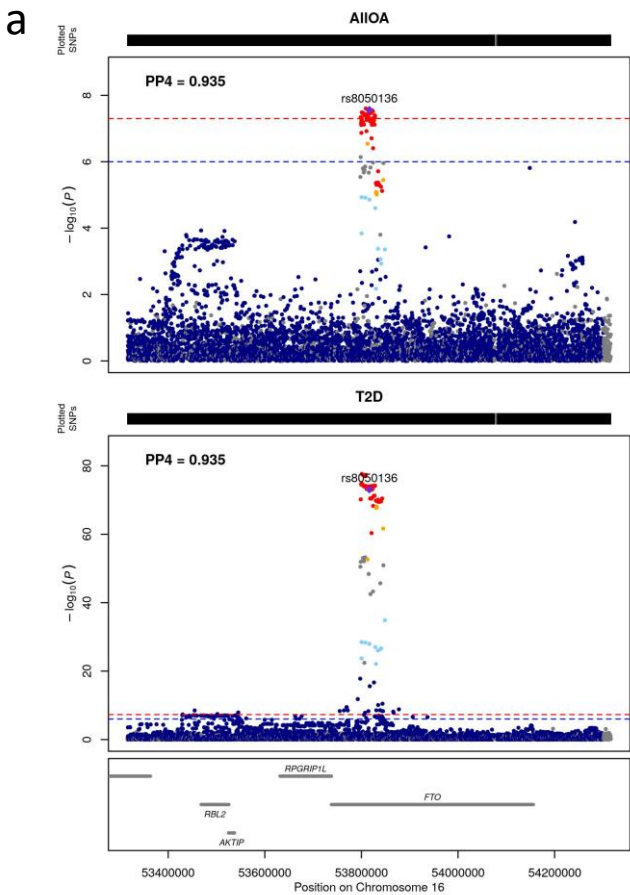


Figure S20: Colocalized region 16. This region colocalizes between type 2 diabetes and total joint replacement.



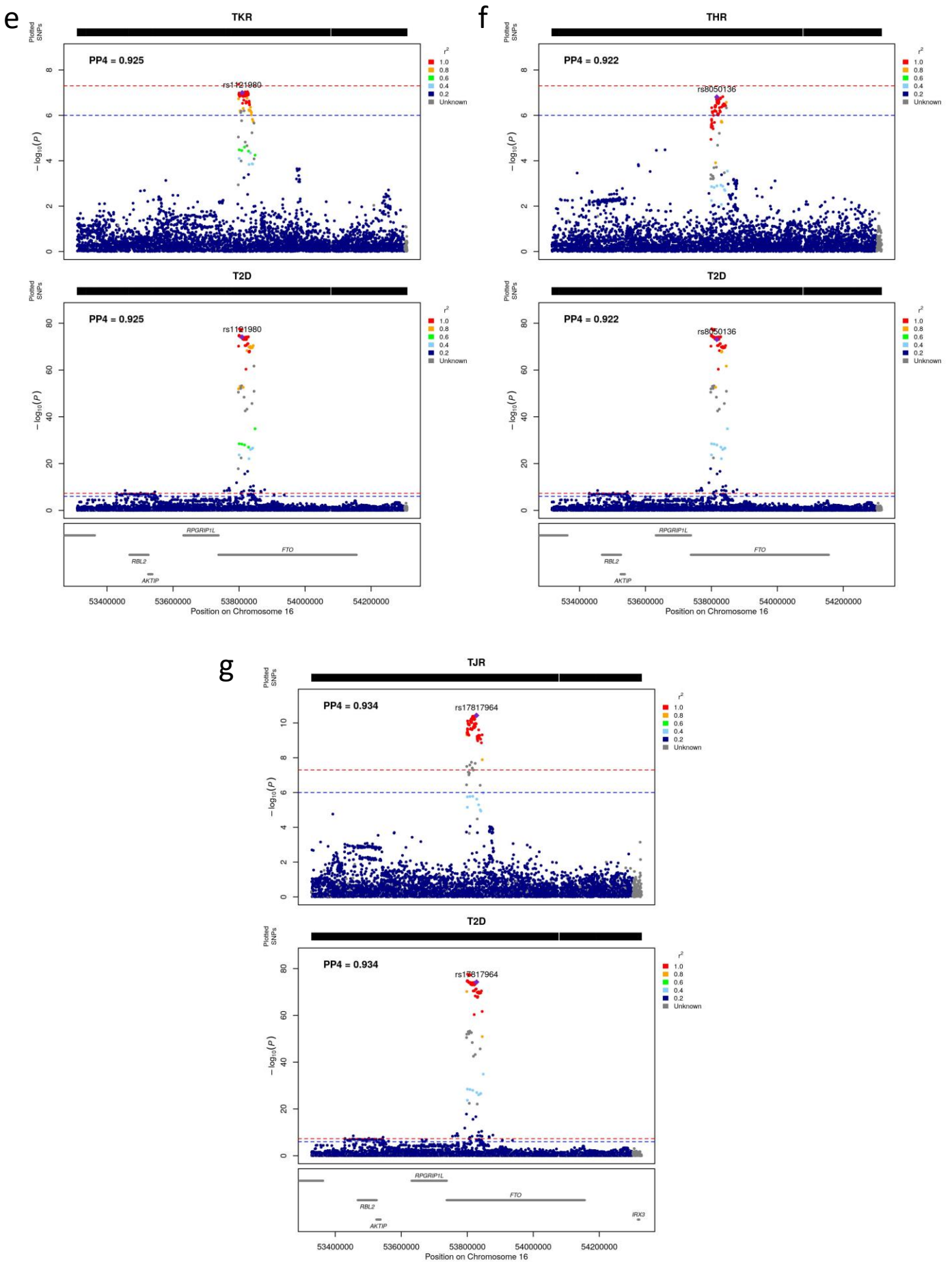


Figure S21: Colocalized region 17. This region colocalizes between type 2 diabetes and a) osteoarthritis at any site, b) knee osteoarthritis, c) knee and/or hip osteoarthritis, d) hip osteoarthritis, e) total knee replacement, f) total hip replacement, g) total joint replacement.

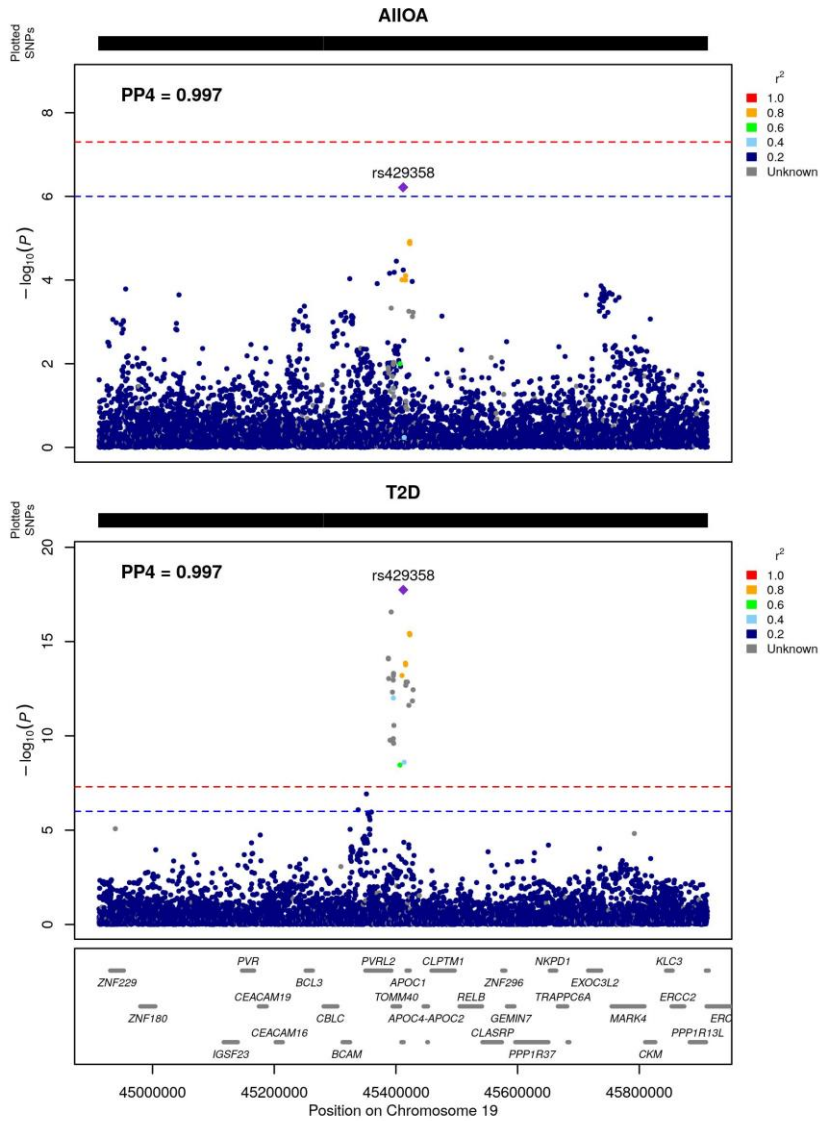


Figure S22: Colocalized region 18. This region colocalizes between type 2 diabetes and osteoarthritis at any site.

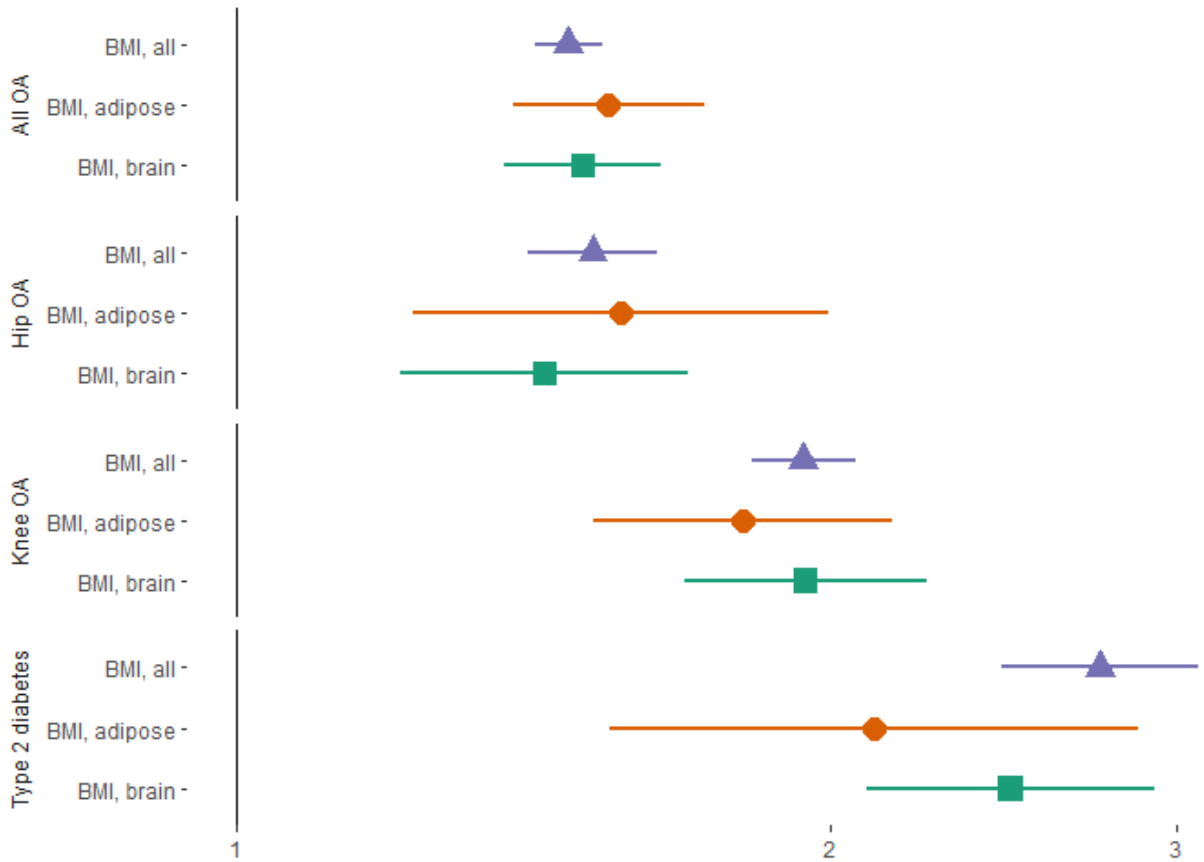


Figure S23: Results of causal inference analyses with tissue-specific BMI instruments. Instruments for the Mendelian randomization analysis were defined as BMI-associated variants that colocalize with eQTLs from the brain and subcutaneous adipose tissue, respectively. The y-axis depicts the magnitude of effect (odds ratio) along with the 95% confidence interval.

	rs7903146			rs35198068			rs34872471		
	EA	beta	pval	EA	beta	pval	EA	beta	pval
Knee OA	T	-0.035	1.8e-5	T	0.035	2.4e-5	T	0.034	3.2e-5
T2D	T	0.31	< 1e-300	T	-0.31	< 1e-300	T	-0.31	< 1e-300

Table S8: Overview of effect size of the variants in the 95% credible set of the *TCF7L2* region for type 2 diabetes and osteoarthritis.

	rs1426371			rs3764002		
	EA	beta	pval	EA	beta	pval
Knee OA	A	-0.0515	8.86E-10	T	-0.0508	8.90E-10
Knee and/or hip OA	A	-0.0408	1.34E-08	T	-0.0402	1.39E-08
All OA	A	-0.0246	3.62E-06	T	-0.0242	3.77E-06
TJR	A	-0.0412	2.67E-05	T	-0.0403	3.27E-05
T2D	A	-0.05	1.10E-11	T	-0.049	1.40E-11

Table S9: Overview of effect size of the variants in the 95% credible set of the *WSCD2* and *TMEM119* region for type 2 diabetes and osteoarthritis.

	rs62492368			rs7794796		
	EA	beta	pval	EA	beta	pval
Knee OA	A	-0.0397	1.77E-07	T	-0.1132	0.01828
Knee and/or hip OA	A	-0.0303	3.96E-06	T	-0.0281	1.26E-05
TKR	A	-0.074	2.17E-08	T	-0.0769	2.88E-09
TJR	A	-0.04	1.58E-05	T	-0.0404	8.06E-06
T2D	A	0.044	1.50E-10	T	0.04	2.50E-09

Table S10: Overview of effect size of the variants in the 95% credible set of the *TMEM176A* region for type 2 diabetes and osteoarthritis.

T2D = type 2 diabetes; OA = osteoarthritis; TJR = total joint replacement; TKR = total knee replacement; EA = effect allele; beta = effect size (log(odds ratio)); pval = p-value