Supplementary

Table S1 Corresponding SMR significance thresholds for each trait and tissue

Trait	Tissue	Number of genes	P_{SMR} (significance thresholds)
SR	Muscle-skeletal	5,474	9.13E–06 (0.05/5,474)
	Whole blood	4,542	1.10E–05 (0.05/4,542)
HD	Muscle-skeletal	5,479	9.13E–06 (0.05/5,479)
	Whole blood	4,545	1.10E–05 (0.05/4,545)
HD-hip	Muscle-skeletal	5,488	9.11E–06 (0.05/5,488)
	Whole blood	4,554	1.10E–05 (0.05/4,554)
HD-knee	Muscle-skeletal	5,484	9.11E–06 (0.05/5,484)
	Whole blood	4,547	1.10E–05 (0.05/4,547)
HD-hip/knee	Muscle-skeletal	5,480	9.12E–06 (0.05/5,480)
	Whole blood	4,545	1.10E–05 (0.05/4,545)

SMR, summary data-based Mendelian randomization; HD, hospital diagnosis; HD-knee, hospital diagnosis of knee osteoarthritis; HD-hip, hospital diagnosis of hip osteoarthritis; HD-hip/knee, hospital diagnosis of hip or knee osteoarthritis; SR, self-reported.

https://dx.doi.org/10.21037/atm-22-4471

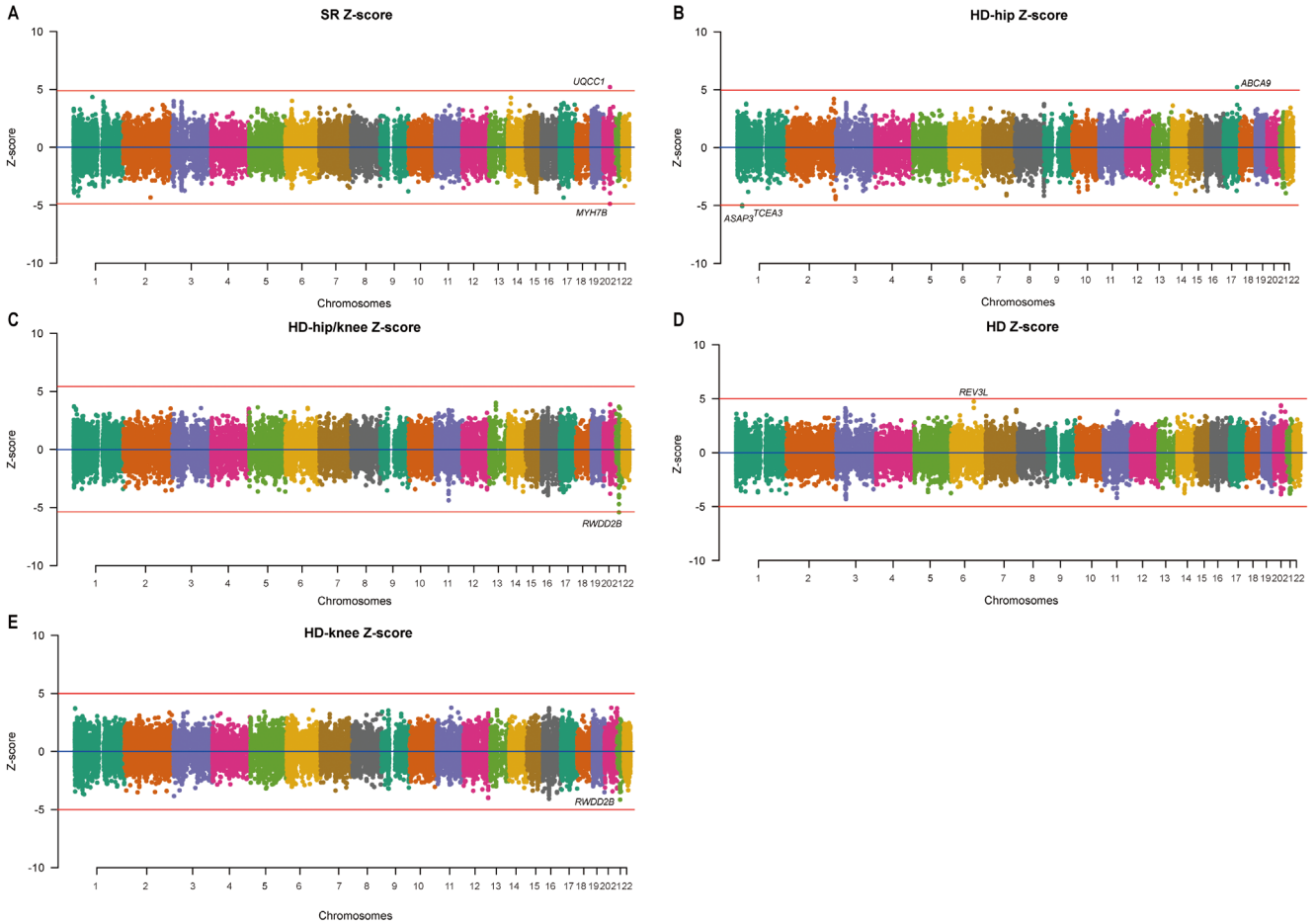

Figure S1 Manhattan plot showing TWAS results (Z-score). The y-axis indicates the z-score for each gene tested on all autosomal and single nucleotide polymorphism weight sets. The x-axis indicates the chromosomal position corresponding to the gene. The magenta horizontal line indicates the threshold of significance (|Z| =4.87, P=1.72E–06) within the transcriptome used in this study. The blue line indicates that Z-score is equal to 0. (A) Using a dataset from the self-reported status, *UQCC1* and *MYH7B* were identified as significant risk genes for OA, with *UQCC1* significantly upregulated and *MYH7B* significantly down-regulated in OA. (B) Using a dataset from the hospital diagnosed hip arthritis data, *ABCA9*, *TCEA3*, and *ASAP3* were identified as significant risk genes for OA, with *ABCA9* significantly upregulated and *TCEA3*, *ASAP3* significantly downregulated in OA. (C) Using a dataset from the hospital diagnosed hip and knee arthritis data, *RWDD2B* was identified as a significant risk gene for OA, with *RWDD2B* significantly downregulated in OA. (D) Using a dataset from a hospital diagnosis, where *REV3L* is the gene closest to the threshold of significance in this dataset. (E) Using a dataset from a hospital diagnosis of knee osteoarthritis, where *RWDD2B* is the gene closest to the threshold of significance in this dataset. TWAS, transcriptome-wide association studies; OA, osteoarthritis; HD, hospital diagnosis; HD, hospital diagnosis; HD-hip, hospital diagnosis of hip osteoarthritis; HD-hip/knee, hospital diagnosis of hip or knee osteoarthritis; SR, self-reported. OA, Osteoarthritis.

https://dx.doi.org/10.21037/atm-22-4471

Table S2 The genes closest to the significance threshold in the two GWAS datasets for the traits HD and HD-knee

Trait	Gene	Chr	Best GWAS ID[a]	eQTL ID[b]	TWAS Z[c]	TWAS P	Tissue
HD	*REV3L*	6	rs3851225	rs6568686	4.7	2.12E–06	Whole blood
HD-knee	*RWDD2B*	21	rs2832155	rs2150403	–4.1	3.35E–05	Whole blood

[a], the SNP showed the most significant association with OA in this locus; [b], the SNP showed the most significant association with gene expression in this locus; [c], the Z statistic reflects the association strength between this gene and OA. Z<0 suggests that this gene was predicted to be down-regulated in OA compared with controls, and vice versa. GWAS, genome-wide association studies; HD, hospital diagnosis; HD-knee, hospital diagnosis of knee osteoarthritis; Chr, chromosome; eQTL, expression quantitative trait loci; TWAS, transcriptome-wide association studies; SNP, single-nucleotide polymorphism. OA, osteoarthritis.

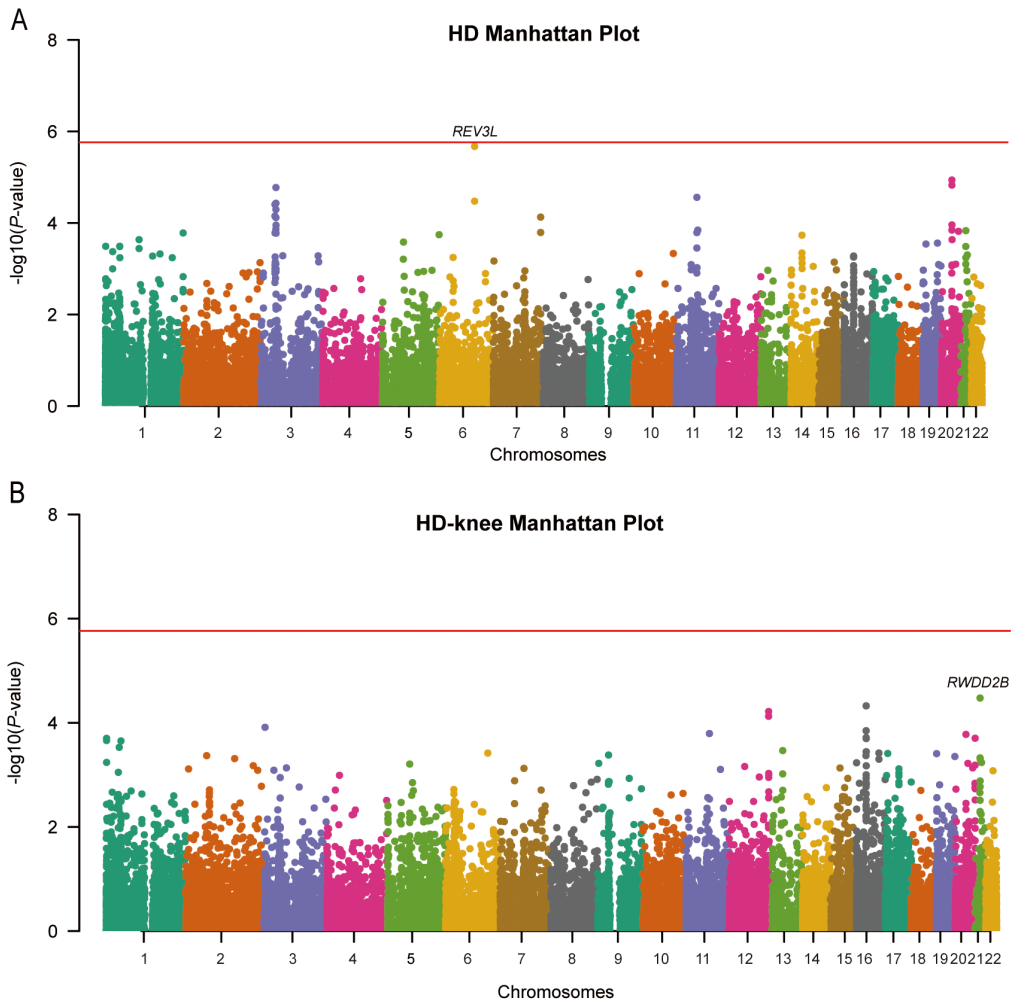

Figure S2 Manhattan plot showing TWAS results (P values) for both HD and HD-knee GWAS datasets. The y-axis shows the $-\log_{10}$ (P value). The x-axis indicates the chromosomal position corresponding to the gene, and the magenta horizontal line indicate the threshold of significance (P=1.72E–06) within the transcriptome used in this study. (A) Using a dataset from a hospital diagnosis, where *REV3L* is the gene closest to the threshold of significance in this dataset. (B) Using a dataset from a hospital diagnosis of knee osteoarthritis, where *RWDD2B* is the gene closest to the threshold of significance in this dataset. TWAS, transcriptome-wide association studies; HD, hospital diagnosis; HD-knee, hospital diagnosis of knee osteoarthritis; GWAS, genome-wide association studies.

https://dx.doi.org/10.21037/atm-22-4471

Table S3 Conditional analysis findings: jointly significant features per locus

Trait	Location	Jointly sign. features (SNP-weight set)	Top TWAS P value	Top GWAS P value	Variance explained
SR	chr20:32564078-34994192	*UQCC1* (muscle-skeletal), *MYH7B* (whole blood)	2.00E−07	8.17E−09	1.000
HD-hip	chr17:66063101-68052951	*ABCA9* (whole blood)	1.78E−07	1.01E−06	0.988
	chr1:22768784-24749706	*ASAP3* (muscle-skeletal), *TCEA3* (whole blood)	4.24E−07	4.24E−07	1.000
HD-hip/knee	chr21:29404679-31385507	*RWDD2B* (whole blood)	5.93E−08	5.39E−08	0.929
HD	chr6:110622462-112799342	*REV3L* (whole blood)	2.12E−06	8.51E−06	0.937
HD-knee	chr21:29404679-31385507	*RWDD2B* (whole blood)	3.35E−05	2.38E−05	0.914

The six risk genes identified in this study were mainly in three of the GWAS datasets, the other two datasets did not have risk genes that met the significance threshold, so only the genes closest to the significance threshold in these two GWAS datasets are listed. SNP, single-nucleotide polymorphism; GWAS, genome-wide association studies; TWAS, transcriptome-wide association studies; HD, hospital diagnosis; HD-knee, hospital diagnosis of knee osteoarthritis; HD-hip, hospital diagnosis of hip osteoarthritis; HD-hip/knee, hospital diagnosis of hip or knee osteoarthritis; SR, self-reported.

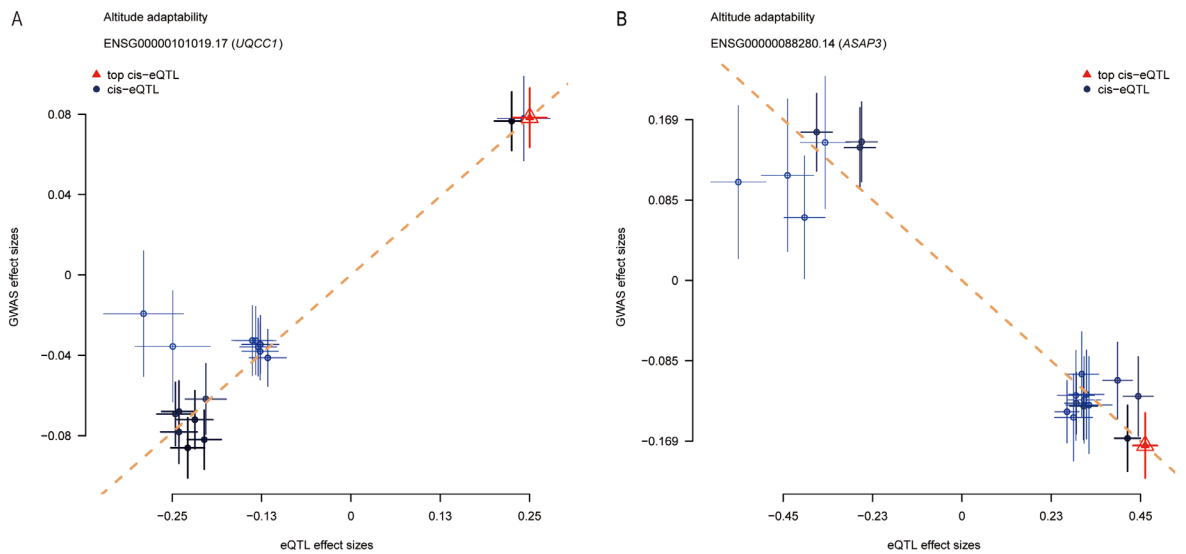

Figure S3 Effect relationships of related genes in GWAS and eQTL. The x-axis indicates the effect size of eQTL. The y-axis indicates the effect size of GWAS. The yellow dashed line is the slope of the regression using the Wald method. GWAS, genome-wide association studies; eQTL, expression quantitative trait loci.

https://dx.doi.org/10.21037/atm-22-4471

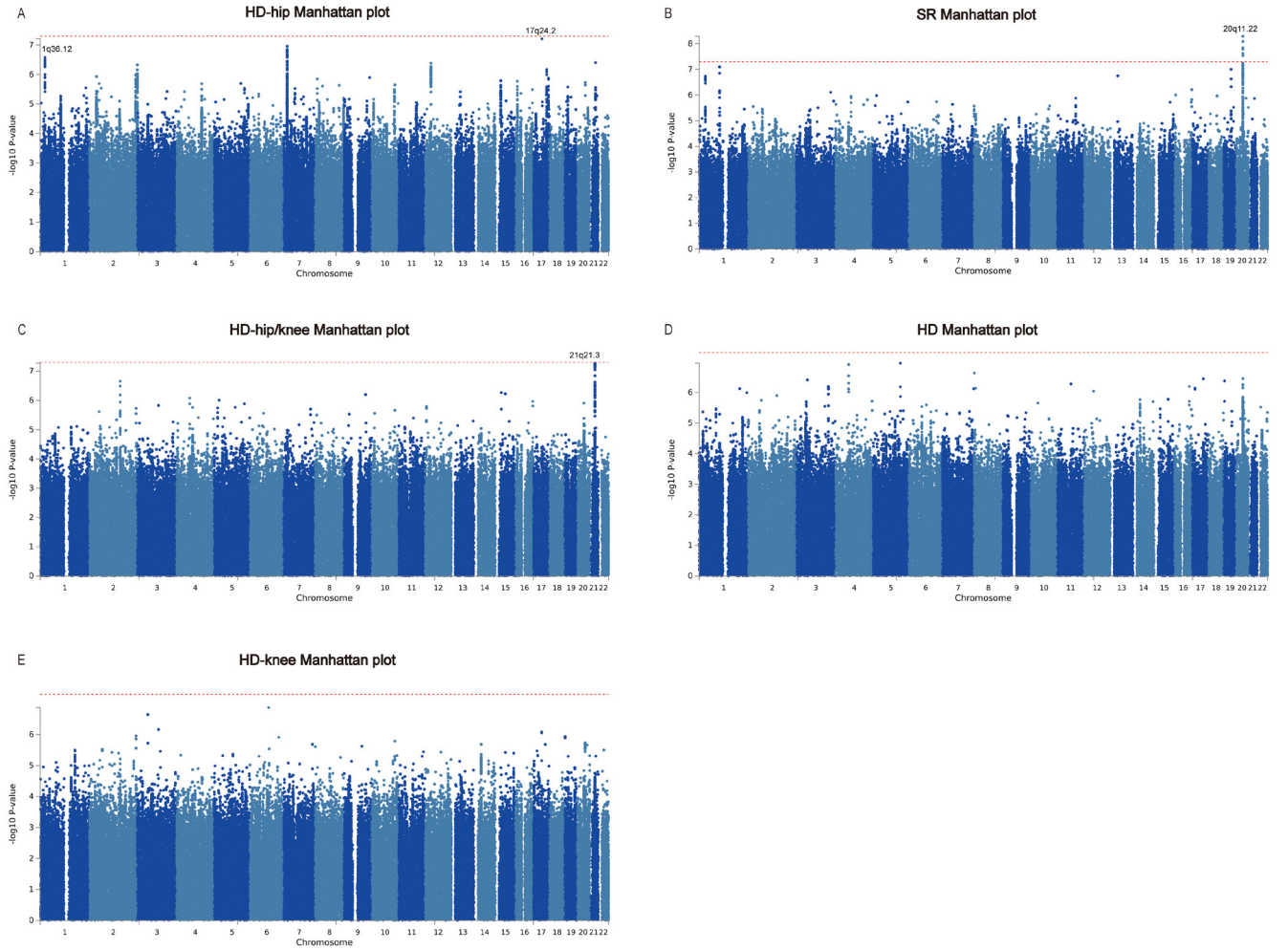

Figure S4 Manhattan plot of genome-wide P values in GWAS analysis. The x-axis represents genomic locations, the y-axis shows –log10 (P value), and the red dashed line indicates the genome-wide significance threshold of P=5E–08. (A) Manhattan plot for the HD-hip GWAS dataset. (B) Manhattan plot for the SR GWAS dataset. (C) Manhattan plot for the HD-hip/knee GWAS dataset. (D) Manhattan plot for the HD GWAS dataset. (E) Manhattan plot for the HD-knee GWAS dataset. HD, hospital diagnosis; HD-knee, hospital diagnosis of knee osteoarthritis; HD-hip, hospital diagnosis of hip osteoarthritis; HD-hip/knee, hospital diagnosis of hip or knee osteoarthritis; GWAS, genome-wide association studies; SR, self-reported.

https://dx.doi.org/10.21037/atm-22-4471

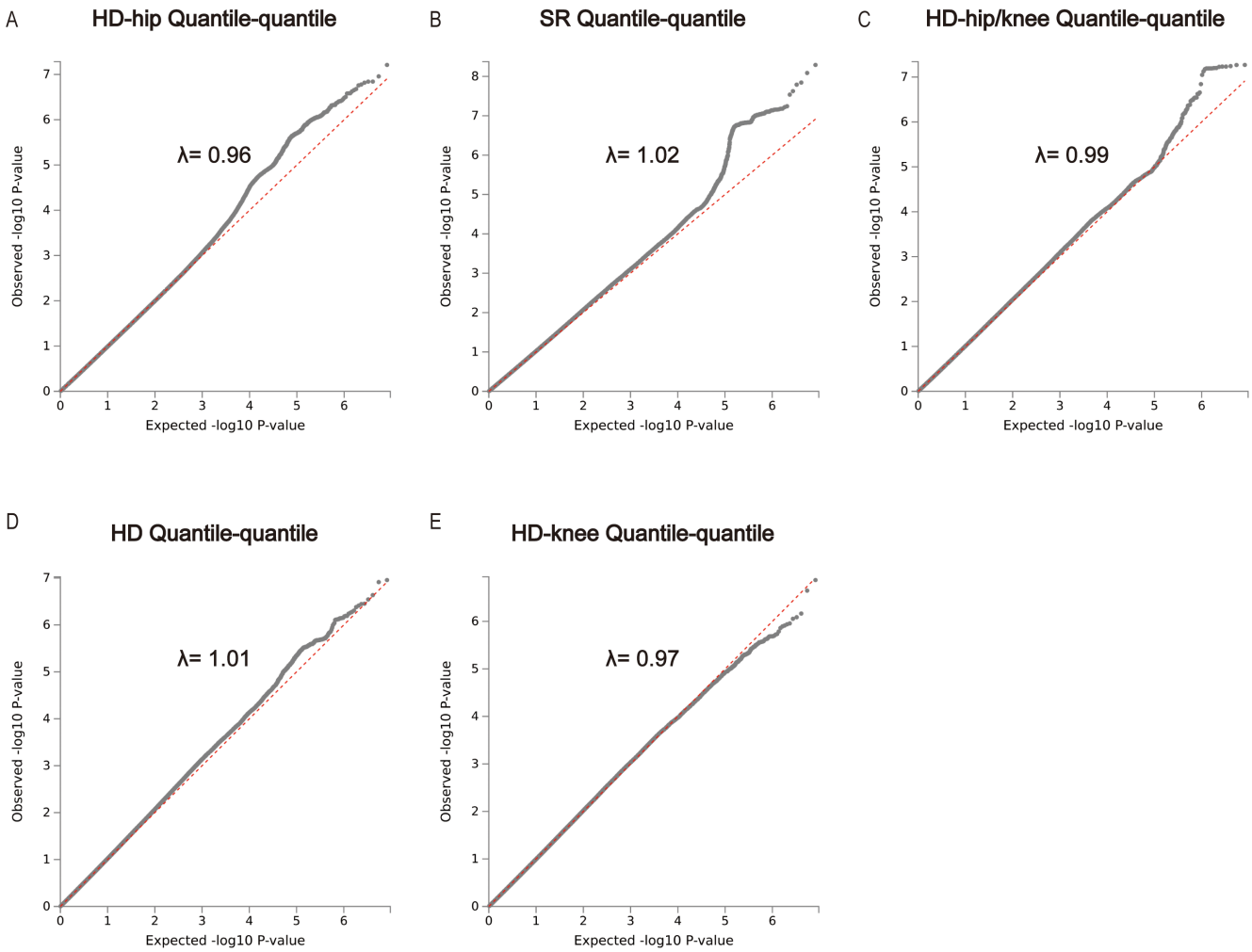

Figure S5 Quantile-quantile plot of the genome-wide P values in the GWAS analysis. The red dashed line represents the null hypothesis of no true association. The black dot with gradient λ (inflation coefficient) is fitted to the lower 90% of the distribution of the observed test statistics. (A) Quantile-quantile plot for the HD-hip GWAS dataset. The value of the inflation factor is 0.96. (B) Quantile-quantile plot for the SR GWAS dataset. The value of the inflation factor is 1.02. (C) Quantile-quantile plot for the HD-hip/knee GWAS dataset. The value of the inflation factor is 0.99. (D) Quantile-quantile plot for the HD GWAS dataset. The value of the inflation factor is 1.01. (E) Quantile-quantile plot for the HD-knee GWAS dataset. The value of the inflation factor is 0.97. HD, hospital diagnosis; HD-knee, hospital diagnosis of knee osteoarthritis; HD-hip, hospital diagnosis of hip osteoarthritis; HD-hip/knee, hospital diagnosis of hip or knee osteoarthritis; GWAS, genome-wide association studies; SR, self-reported.

https://dx.doi.org/10.21037/atm-22-4471