# Identification of Druggable Targets and Drug Prediction for Adolescent Idiopathic Scoliosis Based on Multi-Omics Mendelian Randomization Analysis

**Yangyi Shi**[1], **Weiming Fan**[2] **& Dongqi Guo**[3]

*Received June 23, 2026*
*Accepted August 3, 2026*
*Electronic access August 31, 2026*

**Objective** To identify druggable targets for adolescent idiopathic scoliosis (AIS), evaluate their safety, and predict candidate drugs targeting these genes.
**Methods:** Drug-target Mendelian randomization (MR) analysis was performed using cis-eQTLs of druggable genes from the eQTLGen consortium as exposures and AIS GWAS data from the GWAS Catalog as outcomes. GO and KEGG enrichment analyses were applied to the candidate genes. Protein QTLs (pQTLs) from FinnGen were then used as exposures to assess the targets at the protein level. A phenome-wide association study (PheWAS) was conducted via the PheWAS Portal to assess safety. Compounds were predicted using DSigDB, and molecular docking of targets and compounds was performed with QuickVina-W.
**Results:** Gene-expression MR identified 119 candidate druggable genes associated with AIS, enriched in processes such as immune effector regulation and collagen catabolism.Protein-level MR supported LTA and NT5E as candidate druggable targets, with LTA withstanding multiple-testing correction (adjusted $P = 0.031$), and PheWAS detected no significant adverse-association signals for either target across the phenotypes evaluated. DSigDB predicted 85 compound-target pairs, of which 78 completed docking; 61 pairs (78.2%) showed binding energy below -5.0 kcal·mol$^{-1}$ and 39 pairs (50.0%) below -7.0 kcal·mol$^{-1}$. After toxicological and drug-likeness triage, the highest-scoring candidate compounds included curcumin for LTA (-6.0 kcal·mol$^{-1}$) and enterolactone for NT5E (-9.5 kcal·mol$^{-1}$).
**Conclusions:** LTA and NT5E may serve as candidate druggable targets for AIS. The screened compounds showed strong in silico binding to these candidate targets and represent computationally nominated compounds that merit extensive validation and safety triage before therapeutic consideration.

**Keywords:** Adolescent idiopathic scoliosis; Mendelian randomization; Drug targets; Phenome-wide association study; Safety; Drug prediction .

## Introduction

Adolescent idiopathic scoliosis (AIS) has a global prevalence of approximately 1.65%–3% and can lead to trunk imbalance, back pain, and, in severe cases, cardiopulmonary dysfunction. AIS represents a major public health concern among children and adolescents[1]. Current clinical management follows a stepwise strategy based on the Cobb angle. Mild cases are typically managed through observation and rehabilitation, whereas moderate cases (20°–40°) are treated with bracing combined with rehabilitation. Severe deformities often require surgical correction[2,3]. However, these approaches have limitations. Early-stage disease lacks effective pharmacological interventions, while brace therapy is associated with poor compliance and psychosocial burden[4,5]. Surgical treatment, although effective, is invasive and carries risks of serious complications, including neurological injury and respiratory dysfunction[6]; furthermore, pediatric patients generally show poor adherence to rehabilitation therapy, resulting in poor adherence to systematic rehabilitation regimens. Therefore, there is an urgent need to explore safe and effective targeted small-molecule drugs to delay scoliosis progression at the molecular level[7] or to serve as synergistic regimens in combination with existing conservative treatments and surgical interventions. Such agents could improve patient outcomes and reduce surgery rates.

However, pediatric drug development faces substantial challenges, including limited predictive accuracy of traditional pharmacological models and increased susceptibility to adverse effects in developing individuals[8]. Human genetic approaches offer opportunities to improve target validation and

[1] *Epartment of Pathology*
[2] *Wangjing Hospital*
[3] *China Academy of Chinese Medical Sciences*

reduce drug development failure rates[9]. In terms of efficacy, Mendelian randomization (MR) uses expression or protein quantitative trait loci (eQTL/pQTL) as instrumental variables to overcome confounding bias in traditional studies and emulate the effects of drug-regulated targets on diseases[10]. Regarding safety, phenome-wide association studies (PheWAS) can assess the causal relationships between interventions targeting these loci and various adverse reactions, such as hepatorenal toxicity, cardiovascular events, and carcinogenicity. This not only overcomes the species differences inherent in animal experiments but also improves the predictive power and reliability of target safety assessments[11]. In this study, we integrated multi-omics data using MR, PheWAS, drug signature databases, and molecular docking to identify druggable targets and candidate compounds for AIS, providing a genetic framework for therapeutic development. We hypothesized that druggable genes whose genetically proxied transcript and protein levels are causally associated with AIS risk can be prioritized as candidate targets. The downstream phenome-wide screening, drug signature retrieval, and molecular docking analyses were planned as secondary, exploratory steps to nominate compound-level hypotheses, not as confirmatory tests of clinical efficacy or safety.

## Methods

The overall analytical workflow, from druggable-gene selection through Mendelian randomization to molecular docking and compound triage, is shown in Figure 1.

## Data Sources

### Exposure Datasets

The exposures in this study included the eQTLs and protein quantitative trait loci (pQTLs) of druggable genes. The summary eQTL data were obtained from the eQTLGen consortium (`https://www.eqtlgen.org`), which contains genome-wide cis-eQTL information from whole blood in 31,684 individuals of European ancestry. To ensure the clinical translational potential of the targets, this study referred to the Druggable Genome list defined by Finan et al.[12] and extracted the single-nucleotide polymorphisms (SNPs) of the cis-expression quantitative trait loci (cis-eQTLs) of these druggable genes as the primary exposure variables. The summary pQTL data used for validation were obtained from the FinnGen Data Freeze 10 v1 proteomics QTL release, a proteomic sub-release of the FinnGen biobank made publicly available on April 24, 2024 (`https://www.finngen.fi/en/access_results`). This release includes two assay-specific subsets: Olink and SomaScan. Because these platforms measure protein abundance using different technologies and generate non-interchangeable effect estimates, only the Olink subset, comprising 619 European-ancestry samples and 2,925 circulating proteins, was used in this study.

### Outcome Datasets

The summary data from the disease outcome genome-wide association study (GWAS) for AIS were derived from the GWAS Catalog database (GCST90179478, `https://www.ebi.ac.uk/gwas`), comprising 1,358 AIS patients and 12,507 healthy controls[13]. To avoid confounding bias from population stratification, this study selected large-scale AIS cohort data from European populations, including effect sizes, standard errors, and *P* values for all available SNPs.

### Selection of Instrumental Variables

To ensure the reliability and validity of the MR analysis, instrumental variables (IVs) included in the study were strictly screened[14]: first, a strict threshold was applied to retain only SNPs with *P* values below the genome-wide significance threshold ($5.0 \times 10^{-8}$) from the cis-eQTL and pQTL data, ensuring a strong association between the IVs and the exposures; to avoid the co-inheritance of adjacent SNPs, the PLINK software (version 1.90)[15] was used to remove linkage disequilibrium (LD), with the threshold set to LD $r^2 < 0.1$ within a 10,000 kb (kilobase pairs) range; finally, the F statistic for each IV was calculated to evaluate the strength of each instrumental variable, and SNPs with F ¡ 10 were excluded to avoid weak instrument bias[16].

### MR Analysis of Drug Targets

Two-sample MR analysis was performed using the TwoSampleMR package (0.5.1)[17] in R software (version 4.3.2). When multiple IVs were included, the inverse-variance-weighted (IVW) method was primarily used to evaluate the causal relationship between gene/protein expression levels and AIS; when only a single IV was available, the Wald ratio method was applied[18]. A $P < 0.05$ was considered the preliminary criterion for determining the existence of a causal relationship. To address multiple testing, Benjamini–Hochberg false discovery rate (FDR) and Bonferroni corrections were applied across the discovery-stage tests, and an exact binomial test assessed whether the number of nominally significant genes exceeded chance expectation; the same corrections were applied across the pQTL-covered candidates at the validation stage. Meanwhile, Cochran's Q test was used to assess heterogeneity; if $P > 0.05$, indicating no statistically significant heterogeneity, a fixed-effects model was adopted; otherwise, a random-effects model was used. The MR-Egger intercept test was used to examine horizontal pleiotropy; $P > 0.05$ indicated no significant horizontal pleiotropy. In addition to the

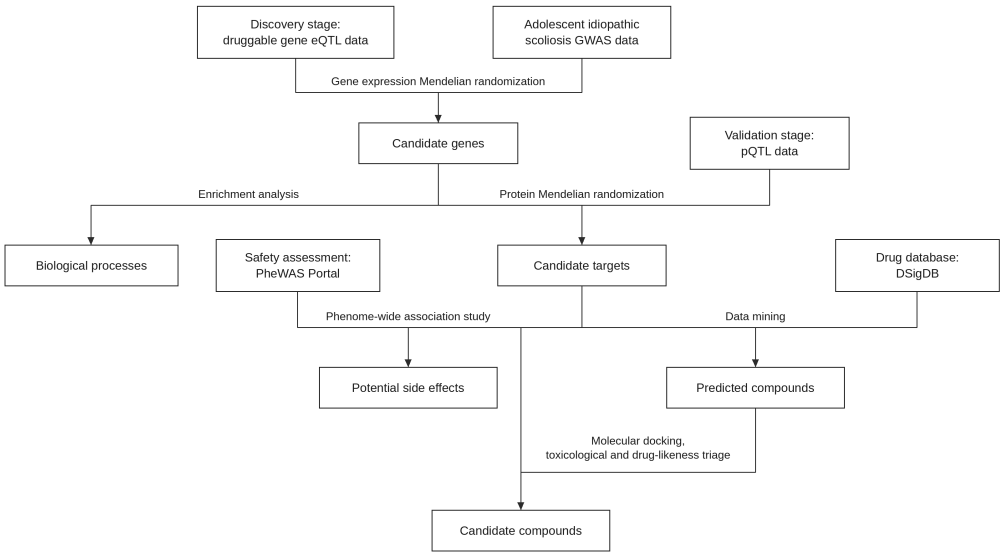


**Figure. 1** Analytical workflow of the study
Note: Candidate genes are the druggable genes retained at the cis-eQTL discovery stage; candidate targets are those additionally supported by cis-pQTL validation; candidate compounds are those retained after molecular docking and the toxicological and drug-likeness triage.

MR-Egger intercept test for horizontal pleiotropy, the MR-Egger slope and weighted-median estimators were reported as pleiotropy-robust sensitivity analyses for all genes with n IVs $\geq 3$. The qqman package (0.1.9) in R was used to draw Manhattan plots, and the Forestplotter package (1.1.2) was used to draw forest plots to display the MR results.

## Gene Functional Enrichment Analysis

The candidate genes for AIS, preliminarily identified through eQTL MR analysis, were input into R, and the ClusterProfiler package (4.10.1)[19] was used for Gene Ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. The GO analysis encompassed biological process (BP), cellular component (CC), and molecular function (MF), which were screened based on Bonferroni-corrected $P < 0.05$. The Enrichplot package (1.22.0) was used to draw lollipop charts to visualize the enrichment analysis results.

## Safety Evaluation Based on Phenome-Wide Association Study

The causal relationships between the AIS candidate targets and 17,361 binary phenotypes and 1,419 continuous phenotypes were evaluated through the PheWAS Portal database (`https://www.azphewas.com`)[20] to further assess the potential adverse reactions of the candidate targets. The PheWAS results were visualized using Manhattan plots generated with the qqman package (0.1.9).

## Drug Prediction

Using the candidate targets as input, potential candidate drugs were predicted in the DSigDB database (`http://dsigdb.tanlab.org/DSigDBv1.0/`)[21]. Based on the target-compound corresponding relationships, the Cytoscape 3.10.4 software[22] was utilized to construct the AIS target-compound network diagram. Cytoscape was used as a visualization tool for the target–compound relationships, and compounds shared between LTA and NT5E were identified by set intersection of their compound lists. The structural files of the predicted compounds were downloaded from the PubChem database (`https://pubchem.ncbi.nlm.nih.gov/`)[23]; for compounds providing only two-dimensional structures, OpenBabel-3.1.1[24] was used to convert their SDF files into three-dimensional conformations for subsequent molecular docking analysis.

## Molecular Docking

We performed molecular docking to evaluate binding energies and interaction characteristics between potential drugs

and their targets. The crystal structures of the target proteins were downloaded from the RCSB Protein Data Bank (PDB; https://www.rcsb.org/). LTA was represented by 1TNR (X-ray diffraction, 2.85 Å resolution, R-work 0.160) and NT5E by 6TVE (X-ray diffraction, 1.05 Å resolution, R-work 0.118, R-free 0.133). Receptor preprocessing included the removal of water molecules, hydrogen atoms, and co-crystallized ligands; the A chain was retained and subsequently converted into PDBQT format. Ligands were sequentially converted from SDF format to MOL2 and PDBQT formats. Subsequently, molecular docking was performed using QuickVina-W (v1.1)[25], with the following parameters: exhaustiveness = 8 and a random seed of 12345. Ultimately, the binding affinity between the ligand and the target protein was evaluated by the binding energy (kcal·mol$^{-1}$), with lower binding energy indicating stronger binding. Prior to docking, receptor binding pockets were predicted using the CB-Dock2 online server (https://cadd.labshare.cn/cb-dock2/). The search box for LTA was centered at (30, 32, 55) with dimensions 17 × 17 × 17 Å, and the search box for NT5E was centered at (24, 24, 31) with dimensions 35 × 28 × 17 Å. A compound–target pair was excluded from the docking calculation if no valid three-dimensional ligand conformation could be generated for PDBQT conversion, or if QuickVina-W failed to converge because the ligand contained non-standard atoms outside the Vina force field.

### Compound Triage

The compound-target pairs that completed docking were then triaged with two explicit filters before interpretation. A toxicology filter categorised as hazardous any compound with a Globally Harmonized System (GHS) carcinogenicity (H350) or germ-cell mutagenicity (H340) classification in PubChem, or belonging to a cytotoxic, immunosuppressive, hormonal, teratogenic, or environmental-toxicant class; because GHS statements describe bulk-chemical handling rather than clinical use, other GHS codes were not applied mechanically. A medicinal-chemistry filter then assessed oral drug-likeness by Lipinski's rule of five[26] using PubChem-computed descriptors (molecular weight ¿ 500 Da, XLogP3 ¿ 5, hydrogen-bond donors ¿ 5, acceptors ¿ 10; ≥2 violations = fail). Hazardous or failing compounds were not carried forward as candidate compounds.

## Results

### Selection and Screening of Instrumental Variables

To describe the source-library composition underlying this study, we examined the overlap among the druggable gene set, the eQTLGen cis-eQTL gene set, and the FinnGen DF10 Olink cis-pQTL gene set (Supplementary Figure S1). 851 druggable genes are shared across all three sources and thus possess both cis-eQTL and cis-pQTL coverage. For the analyses, discovery-stage MR was performed on druggable genes with available cis-eQTL instruments; after instrument selection (see Selection of Instrumental Variables), 1,800 genes carried at least one qualified cis-eQTL instrument and were taken forward into the discovery-stage MR analysis. The validation-stage MR analysis was subsequently restricted to the subset of candidate genes with available cis-pQTL coverage in the FinnGen DF10 Olink subset.

### Gene Expression MR Identifies Candidate Genes for AIS

In the discovery stage, a two-sample MR analysis was performed using cis-eQTL data from the eQTLGen consortium as exposures and the AIS GWAS summary data as the outcome. At a nominal screening threshold of $P$ ¡ 0.05, 119 candidate druggable genes showed causal associations with AIS risk and were carried forward for orthogonal cis-pQTL validation. This exceeded the 90 associations expected by chance among 1,800 genes (exact binomial $P = 1.5 \times 10^{-3}$), indicating genuine enrichment within the druggable genome, although no individual gene reached genome-wide significance after FDR or Bonferroni correction at this stage. As shown in the Manhattan plot, MR analysis identified genes such as MKNK1, CD1D, STK36, HLA-C, C4B, FKBP9, MATN2, HABP4, GPX2, and DNASE1 whose genetically proxied expression was causally associated with AIS risk, nominating them as candidate druggable genes for further validation (Figure 2).For all 119 candidate genes with ≥3 IVs, Cochran's Q and MR-Egger intercept tests were performed, with per-gene results provided in Supplementary Table S1.

### Enrichment Analysis of Candidate Genes for AIS

GO functional annotation and KEGG pathway analysis were performed on the 119 candidate druggable genes with significant causal relationships to AIS. The GO functional annotation results showed that in BP, the candidate genes exert their effects mainly by participating in processes such as regulation of immune effector process, regulation of cell killing, and collagen catabolic process; in CC, the proteins encoded by these genes are mainly located in the external side of plasma membrane, major histocompatibility complex (MHC) protein complex, and collagen-containing extracellular matrix; in terms of MF, their functions are closely related to cytokine receptor binding, antigen binding, and metalloendopeptidase activity. KEGG pathway analysis indicated that these candidate genes are mainly enriched in signaling pathways, including Cell adhesion molecule interaction, Phagosome, and Th1 and Th2 cell differentiation (Figure 3).

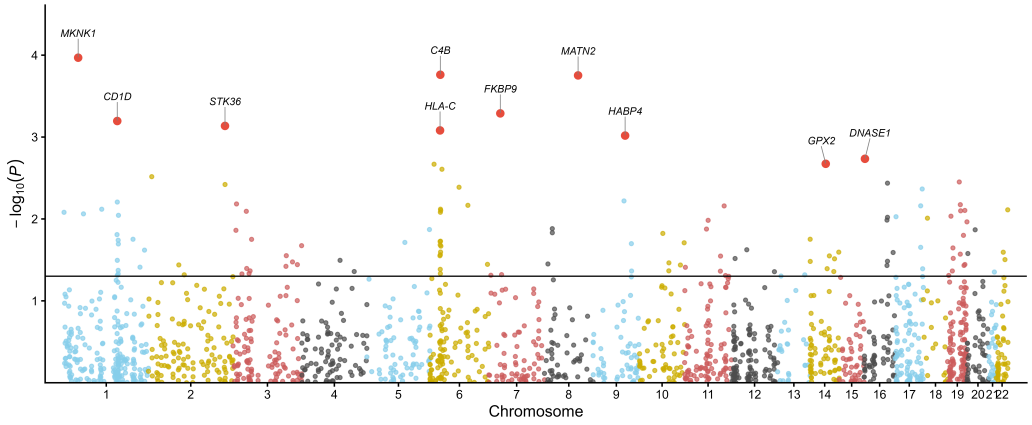


**Figure. 2** Manhattan plot of causal relationships between druggable genes and AIS

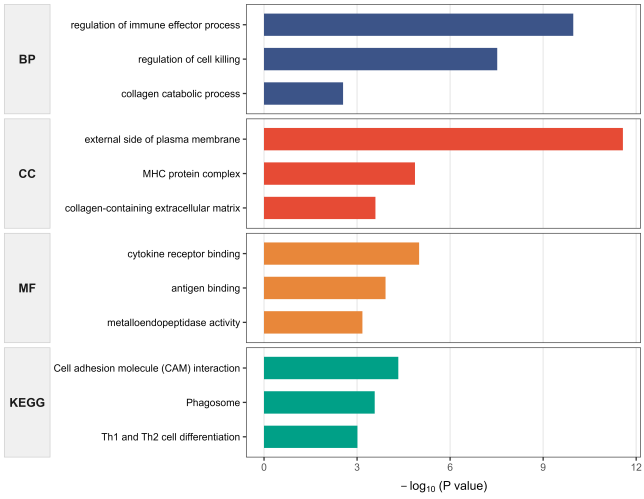


**Figure. 3** Bar chart of GO and KEGG enrichment analysis of AIS candidate genes

Note: The figure shows the top 3 most significant terms from the GO enrichment and KEGG pathway analyses. BP represents Biological Process, CC represents Cellular Component, MF represents Molecular Function, and KEGG represents Kyoto Encyclopedia of Genes and Genomes pathway.

### Protein MR Validation of AIS Candidate Targets

In the validation stage, MR analysis was performed using the FinnGen DF10 Olink cis-pQTL data as exposures and the same AIS GWAS data from the discovery stage as outcomes to verify the reliability of the results. In this stage, the 119 candidate genes were assessed at the protein level. A candidate gene was carried forward as a candidate target if it simultaneously satisfied three prespecified criteria: (i) cis-pQTL IVW effect direction concordant with the cis-eQTL stage on the AIS outcome; (ii) cis-pQTL IVW $P$ ¡ 0.05; and (iii) no MR-Egger intercept evidence of directional pleiotropy at $P$ ¡ 0.05. The results showed that 2 genes, LTA and NT5E, had concordant effect directions in Mendelian randomization analyses of eQTLs and pQTLs. Among them, LTA significantly reduced the risk of AIS at both the gene expression level (odds ratio [OR] = 0.668, 95% confidence interval [CI] = 0.475–0.940, $P$ = 0.0207) and the protein expression level (OR = 0.904, 95% CI = 0.853–0.959, $P$ ¡ 0.001); NT5E was also associated with a reduced risk of AIS at the gene expression level (OR = 0.825, 95% CI = 0.724–0.941, $P$ = 0.0041) and the protein expression level (OR = 0.928, 95% CI = 0.868–0.991, $P$ = 0.0258). For the two candidate targets, instrument strength was strong at both stages: LTA was proxied by 6 cis-eQTL and 17 cis-pQTL instruments (F-statistic ranges 35–207 and 20–212), and NT5E by 43 cis-eQTL and 20 cis-pQTL instruments (F ranges 31–1615 and 20–70); all F ¿ 10, indicating no weak-instrument bias. Across the 41 pQTL-covered candidates, LTA retained significance after multiple-testing correction (Bonferroni-adjusted $P$ = 0.031; FDR q ¡ 0.05), whereas NT5E remained nominal and is carried forward as a candidate meriting further functional validation. Neither gene showed significant heterogeneity ($P$ ¿ 0.05); LTA showed no evidence of directional pleiotropy at either stage, whereas for NT5E the protein-stage MR-Egger intercept indicated possible pleiotropy ($P$ = 0.015), for which pleiotropy-robust estimates (MR-Egger slope OR = 0.758, $P$ = 0.003; weighted median OR = 0.919, $P$ = 0.052) remained direction-concordant. Full per-stage MR diagnostics, including F-statistic ranges, MR-Egger and weighted-median sensitivity estimates, and Cochran's Q tests, are provided in Supplementary Table S2. The remaining genes were not carried forward owing to non-

concordant effect directions between the discovery and validation stages or nominal significance not being reached (Figure 4).

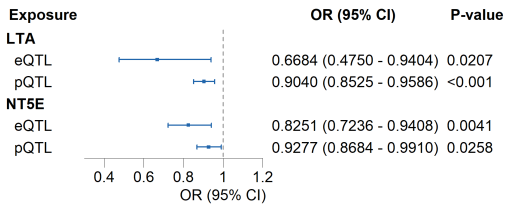


**Figure. 4** Forest plot of Mendelian randomization results for the expression levels of candidate druggable targets (LTA and NT5E) and AIS risk

### Safety Evaluation of AIS Candidate Targets by PheWAS

A phenome-wide association study was conducted using the PheWAS Portal database to identify potential adverse reactions associated with targeting druggable genes for AIS. Within the phenotype panel comprising 17,361 binary and 1,419 continuous phenotypes from the PheWAS Portal, neither LTA nor NT5E showed a phenome-wide significant association with any evaluated phenotype at $P ¡ 5 \times 10^{-8}$; the most strongly suggestive but non-significant associations per target are summarized in Supplementary Table S3. These findings indicate that targeting LTA and NT5E shows no significant adverse-association signals at the genome-wide significance threshold across the phenotypes examined, which merits confirmation in dedicated safety studies (Figure 5).

### Drug Prediction and Target Network Construction Based on DSigDB

To predict potential therapeutic agents for AIS, this study used the DSigDB database to identify and screen small-molecule compounds targeting the candidate targets (LTA, NT5E). The search returned 85 compound-target pairs. Visualization of the target–compound network showed that LTA (red diamond node) and NT5E (blue diamond node) each possess distinct potential regulatory compound libraries; in addition, six compounds (purple circular nodes), namely 1-NITROPYRENE, cyclosporin A, estradiol, indomethacin, vincristine, and acetaminophen, can regulate both LTA and NT5E (Supplementary Figure S2).

### Molecular Docking Validation of Targeted Binding Activity of Candidate Drugs

Of the 85 compound–target pairs nominated by DSigDB, 4 pairs were excluded during ligand preparation and 3 further pairs were excluded during docking execution; the remaining 78 pairs successfully completed the molecular docking calculation. The results showed that most pairs reached favorable binding energies: 61 pairs (78.2%) had a binding energy below -5.0 kcal·mol$^{-1}$, read as medium affinity; 39 pairs (50.0%) had a binding energy below -7.0 kcal·mol$^{-1}$, read as high affinity[27] (Supplementary Figure S3). Neither optimal binder was retained by the compound triage: 1-NITROPYRENE carries a GHS carcinogenicity classification (H350; Supplementary Table S4), and Rifaximin did not satisfy the drug-likeness filter (3 Lipinski violations; Supplementary Table S5). Among candidate compounds, the highest-scoring binder for LTA was curcumin (-6.0 kcal·mol$^{-1}$), and for NT5E enterolactone and co-dergocrine mesilate ranked jointly highest (both -9.5 kcal·mol$^{-1}$). The 3D docking models of the curcumin–LTA and enterolactone–NT5E complexes are shown in Figure 6, and the binding energies of all retained compound–target pairs are listed in Supplementary Table S6.

## Discussion

In this study, we systematically integrated the druggable genome with AIS GWAS data via cis-eQTL and cis-pQTL Mendelian randomization, identifying 119 candidate druggable genes associated with AIS risk at the discovery stage. Based on orthogonal cis-pQTL validation with effect-direction concordance, we further narrowed the candidate pool and identified LTA and NT5E as candidate druggable targets for AIS, warranting further functional and clinical investigation. Downstream gene functional enrichment, PheWAS, and DSigDB-based drug signature retrieval combined with molecular docking respectively characterized biological processes among the candidates, evaluated target-level safety signals, and nominated candidate compounds that merit extensive validation and safety triage before therapeutic consideration.

In addition to formal multiple-testing correction, this study uses a prespecified two-stage discovery–validation framework. The cis-eQTL nominal $P ¡ 0.05$ serves as a screening filter, and target retention additionally requires cis-pQTL replication at $P ¡ 0.05$ with effect-direction concordance. This orthogonal cross-omics requirement prioritised 2 of the 119 candidate genes (1.7%) for follow-up.

The two candidate targets map onto two established axes of AIS pathogenesis: NT5E onto bone-metabolic homeostasis, and LTA onto the immune–neuromuscular processes implicated in paravertebral muscle asymmetry. The NT5E gene encodes the CD73 protein (Ecto-5′-nucleotidase), a key rate-limiting enzyme that catalyzes the generation of adenosine from adenosine monophosphate (AMP). Our MR analysis showed that NT5E expression was significantly associated with a reduced risk of AIS at both eQTL and pQTL levels, suggesting a protective role. Previous in vitro and animal

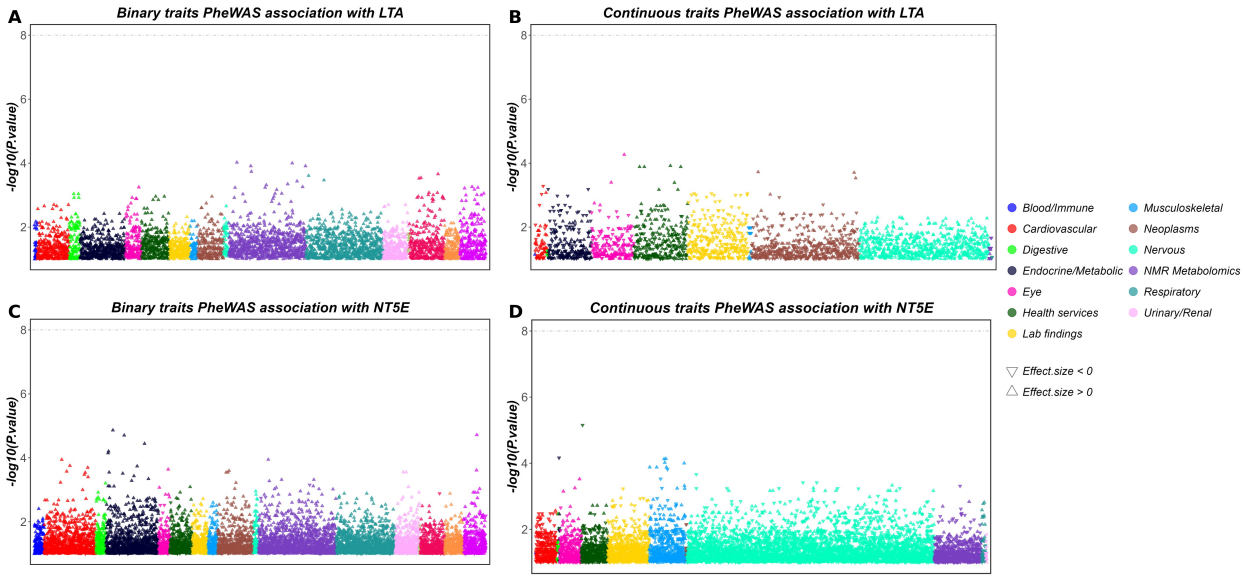


**Figure. 5** PheWAS analysis of candidate targets for AIS
Note: A. LTA-binary traits; B. LTA-continuous traits; C. NT5E-binary traits; D. NT5E-continuous traits.

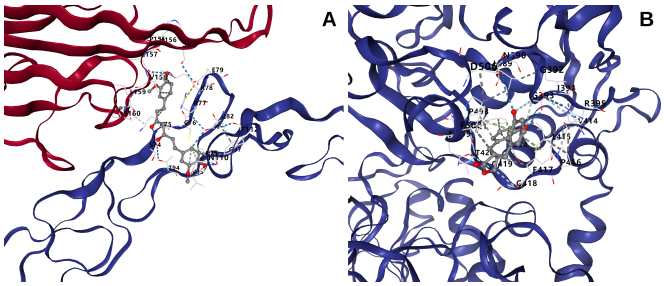


**Figure. 6** 3D molecular docking models of the highest-scoring candidate compound for each candidate target
Note: A. curcumin docked into LTA (PDB 1TNR); B. enterolactone docked into NT5E (PDB 6TVE).

model studies have explicitly demonstrated[28] that local adenosine generated by CD73 can promote osteoblast differentiation and bone formation. Given the systemic osteopenia and asymmetrical spinal bone remodeling observed in AIS[29], pharmacological activation of NT5E may plausibly improve bone microenvironment homeostasis and delay scoliosis progression by regulating local adenosine concentrations, warranting further functional validation.

LTA (lymphotoxin alpha) is a key member of the tumor necrosis factor superfamily and plays an important role in regulating the immune microenvironment. Our GO enrichment analysis indicated that our candidate genes for AIS are significantly enriched in the regulation of immune effector processes and the catabolic processes of the extracellular matrix (containing collagen). In recent years, local low-grade inflammation and autoimmune imbalance have been considered important driving factors for the asymmetrical fibrosis of paravertebral muscles and skeletal degeneration in AIS[30]. Although LTA is generally considered a pro-inflammatory factor, in the complex immune network, moderate expression of immune factors is essential for maintaining tissue repair and immune tolerance of the musculoskeletal system[31]. Our genetic data suggested a potential protective effect of LTA on AIS, providing a novel target for exploring immunomodulation in the intervention of spinal deformities.

In addition to confirming target efficacy, safety assessment is a critical barrier determining whether new drug development can successfully move toward clinical translation. AIS patients are in their peak period of growth and development. The potential toxic side effects of long-term medication on the endocrine system, epiphyseal growth, and hepatorenal function are therefore a critical concern in clinical practice[8]. Distinct from previous omics studies that relied solely on eQTL, this study strictly incorporated pQTL data in the validation stage. Given that the direct receptors for the vast majority of small-molecule drugs are proteins rather than mRNA, MR validation based on pQTL can more realistically simulate the physiological effects after a drug targets and binds to a protein. This effectively excludes false-positive interference caused by post-transcriptional modification or post-translational degradation, thereby providing a more solid pharmacological basis for the final targets screened[32].

In the candidate drug prediction and evaluation stage, the

molecular docking results are consistent with mechanisms proposed for AIS. For the candidate target NT5E, enterolactone ranked jointly highest among the candidate compounds (-9.5 kcal·mol$^{-1}$), and several dietary flavonoids, including chrysin, luteolin and apigenin, also scored between -9.2 and -9.0 kcal·mol$^{-1}$. The incidence of AIS shows significant gender differences, with females at far higher risk of progressing to severe scoliosis than males, and abnormal estrogen receptor signaling is recognized as one of the core pathogenic hypotheses of AIS[1,33]. NT5E is itself an estrogen receptor-regulated gene: in breast cancer cells, estrogen receptor signaling lowers its expression and reduces adenosine generation[34]. Our Mendelian randomization estimates place higher NT5E at lower AIS risk, so a mechanism that lowers NT5E would act in the direction of higher risk. Enterolactone is a lignan formed by gut microbiota from dietary precursors[35], and gut microbiota composition in AIS has itself been associated with aberrant bone homeostasis[36]. These docking results provide preliminary in silico evidence for further exploration of NT5E-targeted modulators involved in endocrine–bone metabolism regulation.

For the candidate target LTA, curcumin ranked highest among the candidate compounds (-6.0 kcal·mol$^{-1}$). Curcumin is a dietary polyphenol widely studied for anti-inflammatory activity, with meta-analyses of randomized trials reporting reductions in C-reactive protein, interleukin-6 and tumor necrosis factor-$\alpha$ in some analyses[37]. Its nomination is consistent with the immune effector processes enriched among the candidate genes and with the position of LTA in the tumor necrosis factor superfamily.

Previous AIS genetics provides context for these findings. Genome-wide association studies have established a set of AIS susceptibility loci, most consistently LBX1 at 10q24.3[38] and GPR126/ADGRG6 at 6q24.1[39], with further signals near PAX1, BNC2 and SOX9. These loci lie almost entirely in non-coding sequence and act on cartilage, muscle, bone, connective-tissue and intervertebral-disc programmes[40]. None appeared among our candidate genes, which follows from the design: the druggable-genome restriction excludes developmental transcription factors and matrix proteins, and blood-derived instruments cannot proxy genes expressed mainly in spinal tissues. The two approaches are therefore complementary, ours asking which druggable protein abundances associate with AIS risk. They nonetheless converge on connective-tissue biology, in that our candidate genes are enriched for collagen catabolic processes.

This study has several strengths in research design, target safety evaluation, and technical integration. First, to our knowledge, this is the first study to provide genetics-based causal evidence for candidate druggable targets in AIS, mitigating the confounding and reverse-causation biases of traditional observational studies. Second, the addition of orthogonal cis-pQTL validation, requiring effect-direction concordance with the cis-eQTL discovery stage, provides protein-level corroboration of the candidate targets. Third, the PheWAS evaluated LTA and NT5E across 17,361 binary and 1,419 continuous phenotypes from the PheWAS Portal, with no significantadverse-association signals detected for either target. This absence of detected adverse-association signals is particularly relevant for AIS patients in the growth and development phase who may require long-term intervention. Together, these components establish a genetics-anchored framework that integrates causal inference, multi-omics validation, and exploratory drug repositioning, providing a foundation for subsequent functional and translational investigation.

However, this study has several limitations. First, the etiology of AIS is extremely complex, involving multiple pathways such as genetics, biomechanics, and neuroendocrinology; although MR can indicate the causal direction of intervention targets, the clinical effect size achievable through single-target intervention may be limited. Second, molecular docking only quantifies the binding energy between small-molecule ligands and receptors based on spatial conformation and physicochemical properties; this "binding" is not fully equivalent to clinical "activation or inhibition" efficacy. Third, the GWAS and eQTL/pQTL summary data used in this study primarily originate from European populations. Given differences in genetic backgrounds across ethnic groups, the generalizability of these targets to other populations, such as Asians, still requires further validation. Allele frequencies and linkage-disequilibrium structure also differ across ancestries, so the cis instruments used here may not transfer directly; extrapolation of these findings to non-European populations therefore warrants caution pending ancestry-specific replication. The exposures are also blood-derived, whereas the pathological process of AIS develops in the spine and the surrounding musculoskeletal tissue. The estimates therefore describe whole-blood expression and circulating protein abundance rather than expression in vertebral bone, growth-plate cartilage, intervertebral disc or paravertebral muscle, where cis effects may differ in magnitude and direction. Future in vivo and in vitro experiments, such as using disease animal models, are urgently needed to confirm the true efficacy of these genetic targets and candidate drugs in treating AIS. In addition, given the modest size of the current AIS GWAS and the absence of independent outcome replication, individual discovery-stage associations should be regarded as hypothesis-generating and warrant confirmation in larger and functional studies. Finally, the PheWAS indicates that no adverse signal was detected rather than that safety has been established; power is limited for rare events and the contributing cohorts comprise adults, so paediatric and developmental outcomes merit dedicated assessment.

## Declarations

### *Data Availability*

The datasets analyzed in this study are publicly available from the eQTLGen consortium (https://www.eqtlgen.org), the GWAS Catalog (accession GCST90179478, https://www.ebi.ac.uk/gwas), and FinnGen Data Freeze 10 (https://www.finngen.fi/en/access_results).

### *Ethics Statement*

All data used were obtained from publicly available databases with prior ethical approval; no additional ethical approval was required for this study.

### *Competing interests*

The authors declare no competing interests

## References


1 S. Wang, M. Li, J. Ren, J. Tao, M. Fang, L. Kong Global prevalence and associated risk factors of scoliosis in children and adolescents: a systematic review and meta-analysis. BMC Public Health. Vol. 25, pg. 3640, 2025, *https://doi.org/10.1186/s12889-025-24905-4*.

2 S. Negrini, S. Donzelli, A. G. Aulisa, D. Czaprowski, S. Schreiber, J. C. de Mauroy, H. Diers, T. B. Grivas, P. Knott, T. Kotwicki, A. Lebel, C. Marti, T. Maruyama, J. O'Brien, N. Price, E. Parent, M. Rigo, M. Romano, L. Stikeleather, J. Wynne, F. Zaina 2016 SOSORT guidelines: orthopaedic and rehabilitation treatment of idiopathic scoliosis during growth. Scoliosis Spinal Disord. Vol. 13, pg. 3, 2018, *https://doi.org/10.1186/s13013-017-0145-8*.

3 S. L. Weinstein, L. A. Dolan, J. G. Wright, M. B. Dobbs Effects of bracing in adolescents with idiopathic scoliosis. N Engl J Med. Vol. 369, pg. 1512–1521, 2013, *https://doi.org/10.1056/NEJMoa1307337*.

4 L. Graham-Wisener, S. Sloan, J. McMullan, R. Waterworth, C. Close, M. Clarke, R. Gordon, P. Toner Identification of important outcomes for surgical and brace treatment of adolescent idiopathic scoliosis: qualitative analysis of the perspectives of young people, parents, and healthcare professionals. Bone Jt Open. Vol. 7, pg. 54–65, 2026, *https://doi.org/10.1302/2633-1462.71.BJO-2025-0218.R1*.

5 L. Rivett, A. Rothberg, A. Stewart, R. Berkowitz The relationship between quality of life and compliance to a brace protocol in adolescents with idiopathic scoliosis: a comparative study. BMC Musculoskelet Disord. Vol. 10, pg. 5, 2009, *https://doi.org/10.1186/1471-2474-10-5*.

6 A. R. Hariharan, S. A. Shah, J. Petfield, M. Baldwin, B. Yaszay, P. O. Newton, L. G. Lenke, B. S. Lonner, F. Miyanji, P. D. Sponseller, A. F. Samdani, Harms Study Group Complications following surgical treatment of adolescent idiopathic scoliosis: a 10-year prospective follow-up study. Spine Deform. Vol. 10, pg. 1097–1105, 2022, *https://doi.org/10.1007/s43390-022-00508-6*.

7 H. Sun, Y. Huang, H. Zhou, H. Chen, J. Dai, J. Zhao, X. Huang, H. Cai, J. Shen Asymmetrical paravertebral muscles fibrosis causes progression of adolescent idiopathic scoliosis via myostatin signalling in fibro-adipogenic progenitors. J Orthop Translat. Vol. 56, pg. 101027, 2026, *https://doi.org/10.1016/j.jot.2025.11.003*.

8 H. Watanabe, N. Nagano, Y. Tsuji, N. Noto, M. Ayusawa, I. Morioka Challenges of pediatric pharmacotherapy: a narrative review of pharmacokinetics, pharmacodynamics, and pharmacogenetics. Eur J Clin Pharmacol. Vol. 80, pg. 203–221, 2024, *https://doi.org/10.1007/s00228-023-03598-x*.

9 K. J. Carss, A. M. Deaton, A. Del Rio-Espinola, D. Diogo, M. Fielden, D. A. Kulkarni, J. Moggs, P. Newham, M. R. Nelson, F. D. Sistare, L. D. Ward, J. Yuan Using human genetics to improve safety assessment of therapeutics. Nat Rev Drug Discov. Vol. 22, pg. 145–162, 2023, *https://doi.org/10.1038/s41573-022-00561-w*.

10 I. Daghlas, D. Gill Mendelian randomization as a tool to inform drug development using human genetics. Camb Prism Precis Med. Vol. 1, pg. e16, 2023, *https://doi.org/10.1017/pcm.2023.5*.

11 R. N. Jerome, M. M. Joly, N. Kennedy, J. K. Shirey-Rice, D. M. Roden, G. R. Bernard, K. J. Holroyd, J. C. Denny, J. M. Pulley Leveraging human genetics to identify safety signals prior to drug marketing approval and clinical use. Drug Saf. Vol. 43, pg. 567–582, 2020, *https://doi.org/10.1007/s40264-020-00915-6*.

12 C. Finan, A. Gaulton, F. A. Kruger, R. T. Lumbers, T. Shah, J. Engmann, L. Galver, R. Kelley, A. Karlsson, R. Santos, J. P. Overington, A. D. Hingorani, J. P. Casas The druggable genome and support for target identification and validation in drug development. Sci Transl Med. Vol. 9, pg. eaag1166, 2017, *https://doi.org/10.1126/scitranslmed.aag1166*.

13 H. Yu, A. M. Khanshour, A. Ushiki, N. Otomo, Y. Koike, E. Einarsdottir, Y. Fan, L. Antunes, Y. H. Kidane, R. Cornelia, R. R. Sheng, Y. Zhang, J. Pei, N. V. Grishin, B. M. Evers, J. P. Y. Cheung, J. A. Herring, C. Terao, Y. Q. Song, C. A. Gurnett, P. Gerdhem, S. Ikegawa, J. J. Rios, N. Ahituv, C. A. Wise Association of genetic variation in COL11A1 with adolescent idiopathic scoliosis. Elife. Vol. 12, pg. RP89762, 2024, *https://doi.org/10.7554/eLife.89762*.

14 C. A. Emdin, A. V. Khera, S. Kathiresan Mendelian randomization. JAMA. Vol. 318, pg. 1925–1926, 2017, *https://doi.org/10.1001/jama.2017.17219*.

15 C. C. Chang, C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, J. J. Lee Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. Vol. 4, pg. 7, 2015, *https://doi.org/10.1186/s13742-015-0047-8*.

16 S. Burgess, S. G. Thompson, CRP CHD Genetics Collaboration Avoiding bias from weak instruments in Mendelian randomization studies. Int J Epidemiol. Vol. 40, pg. 755–764, 2011, *https://doi.org/10.1093/ije/dyr036*.

17 G. Hemani, J. Zheng, B. Elsworth, K. H. Wade, V. Haberland, D. Baird, C. Laurin, S. Burgess, J. Bowden, R. Langdon, V. Y. Tan, J. Yarmolinsky, H. A. Shihab, N. J. Timpson, D. M. Evans, C. Relton, R. M. Martin, G. Davey Smith, T. R. Gaunt, P. C. Haycock The MR-Base platform supports systematic causal inference across the human phenome. Elife. Vol. 7, pg. e34408, 2018, *https://doi.org/10.7554/eLife.34408*.

18 C. Qin, L. M. Diaz-Gallo, B. Tang, Y. Wang, T. D. Nguyen, A. Harder, Y. Lu, L. Padyukov, J. Askling, S. Hägg Repurposing antidiabetic drugs for rheumatoid arthritis: results from a two-sample Mendelian randomization study. Eur J Epidemiol. Vol. 38, pg. 809–819, 2023, *https://doi.org/10.1007/s10654-023-01000-9*.

19 S. Xu, E. Hu, Y. Cai, Z. Xie, X. Luo, L. Zhan, W. Tang, Q. Wang, B. Liu, R. Wang, W. Xie, T. Wu, L. Xie, G. Yu Using clusterProfiler to characterize multiomics data. Nat Protoc. Vol. 19, pg. 3292–3320, 2024, *https://doi.org/10.1038/s41596-024-01020-z*.

20 Q. Wang, R. S. Dhindsa, K. Carss, A. R. Harper, A. Nag, I. Tachmazidou, D. Vitsios, S. V. V. Deevi, A. Mackay, D. Muthas, M. Hühn, S. Monkley, H. Olsson, AstraZeneca Genomics Initiative, S. Wasilewski, K. R. Smith, R. March, A. Platt, C. Haefliger, S. Petrovski Rare variant contribution to human disease in 281,104 UK Biobank exomes. Nature. Vol. 597, pg. 527–532, 2021, *https://doi.org/10.1038/s41586-021-03855-y*.

21 M. Yoo, J. Shin, J. Kim, K. A. Ryall, K. Lee, S. Lee, M. Jeon, J. Kang, A. C. Tan DSigDB: drug signatures database for gene set analysis. Bioinformatics. Vol. 31, pg. 3069–3071, 2015, *https://doi.org/10.109*

3/bioinformatics/btv313.

22 P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. Vol. 13, pg. 2498–2504, 2003, https://doi.org/10.1101/gr.1239303.

23 S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E. E. Bolton PubChem 2023 update. Nucleic Acids Res. Vol. 51, pg. D1373–D1380, 2023, https://doi.org/10.1093/nar/gkac956.

24 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison Open Babel: an open chemical toolbox. J Cheminform. Vol. 3, pg. 33, 2011, https://doi.org/10.1186/1758-2946-3-33.

25 N. M. Hassan, A. A. Alhossary, Y. Mu, C. K. Kwoh Protein-ligand blind docking using QuickVina-W with inter-process spatio-temporal integration. Sci Rep. Vol. 7, pg. 15451, 2017, https://doi.org/10.1038/s41598-017-15571-7.

26 C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. Vol. 46, pg. 3–26, 2001, https://doi.org/10.1016/s0169-409x(00)00129-0.

27 L. Jin, Y. Guan, X. Li, M. Wang, Y. Shen, N. Wang, Z. He Combining network pharmacology, molecular docking and experimental validation to explore the effects and mechanisms of indirubin on acute lymphoblastic leukemia. Drug Des Devel Ther. Vol. 19, pg. 1083–1103, 2025, https://doi.org/10.2147/DDDT.S500249.

28 M. Takedachi, H. Oohara, B. J. Smith, M. Iyama, M. Kobashi, K. Maeda, C. L. Long, M. B. Humphrey, B. J. Stoecker, S. Toyosawa, L. F. Thompson, S. Murakami CD73-generated adenosine promotes osteoblast differentiation. J Cell Physiol. Vol. 227, pg. 2622–2631, 2012, https://doi.org/10.1002/jcp.23001.

29 J. C. Cheng, R. M. Castelein, W. C. Chu, A. J. Danielsson, M. B. Dobbs, T. B. Grivas, C. A. Gurnett, K. D. Luk, A. Moreau, P. O. Newton, I. A. Stokes, S. L. Weinstein, R. G. Burwell Adolescent idiopathic scoliosis. Nat Rev Dis Primers. Vol. 1, pg. 15030, 2015, https://doi.org/10.1038/nrdp.2015.30.

30 M. M. Newton Ede, S. W. Jones Adolescent idiopathic scoliosis: evidence for intrinsic factors driving aetiology and progression. Int Orthop. Vol. 40, pg. 2075–2080, 2016, https://doi.org/10.1007/s00264-016-3132-4.

31 Z. T. Liang, J. Li, H. Tang, J. K. Li, M. J. Li, C. F. Guo, H. Q. Zhang Critical role of NLRP3 in causing paravertebral muscle injury in adolescent idiopathic scoliosis. Clin Transl Med. Vol. 14, pg. e1528, 2024, https://doi.org/10.1002/ctm2.1528.

32 A. F. Schmidt, C. Finan, M. Gordillo-Marañ ón, F. W. Asselbergs, D. F. Freitag, R. S. Patel, B. Tyl, S. Chopade, R. Faraway, M. Zwierzyna, A. D. Hingorani Genetic drug target validation using Mendelian randomisation. Nat Commun. Vol. 11, pg. 3255, 2020, https://doi.org/10.1038/s41467-020-16969-0.

33 S. Zheng, H. Zhou, B. Gao, Y. Li, Z. Liao, T. Zhou, C. Lian, Z. Wu, D. Su, T. Wang, P. Su, C. Xu Estrogen promotes the onset and development of idiopathic scoliosis via disproportionate endochondral ossification of the anterior and posterior column in a bipedal rat model. Exp Mol Med. Vol. 50, pg. 1–11, 2018, https://doi.org/10.1038/s12276-018-0161-7.

34 J. Spychala, E. Lazarowski, A. Ostapkowicz, L. H. Ayscue, A. Jin, B. S. Mitchell Role of estrogen receptor in the regulation of ecto-5'-nucleotidase and adenosine in breast cancer. Clin Cancer Res. Vol. 10, pg. 708–717, 2004, https://doi.org/10.1158/1078-0432.CCR-0811-03.

35 A. Senizza, G. Rocchetti, J. I. Mosele, V. Patrone, M. L. Callegari, L. Morelli, L. Lucini Lignans and gut microbiota: an interplay revealing potential health implications. Molecules. Vol. 25, pg. 5709, 2020, https://doi.org/10.3390/molecules25235709.

36 J. Li, C. Liu, Y. Xu, C. Ling, Z. Tang, A. Kiram, Z. Hu, Z. Zhu, Y. Qiu, Z. Liu Gut microbiota alterations in adolescent idiopathic scoliosis are associated with aberrant bone homeostasis. Orthop Surg. Vol. 16, pg. 965–975, 2024, https://doi.org/10.1111/os.14019.

37 Y. M. Lee, Y. Kim Is curcumin intake really effective for chronic inflammatory metabolic disease? A review of meta-analyses of randomized controlled trials. Nutrients. Vol. 16, pg. 1728, 2024, https://doi.org/10.3390/nu16111728.

38 Y. Takahashi, I. Kou, A. Takahashi, T. A. Johnson, K. Kono, N. Kawakami, K. Uno, M. Ito, S. Minami, H. Yanagida, H. Taneichi, T. Tsuji, T. Suzuki, H. Sudo, T. Kotani, K. Watanabe, K. Chiba, N. Hosono, N. Kamatani, T. Tsunoda, Y. Toyama, M. Kubo, M. Matsumoto, S. Ikegawa A genome-wide association study identifies common variants near LBX1 associated with adolescent idiopathic scoliosis. Nat Genet. Vol. 43, pg. 1237–1240, 2011, https://doi.org/10.1038/ng.974.

39 I. Kou, Y. Takahashi, T. A. Johnson, A. Takahashi, L. Guo, J. Dai, X. Qiu, S. Sharma, A. Takimoto, Y. Ogura, H. Jiang, H. Yan, K. Kono, N. Kawakami, K. Uno, M. Ito, S. Minami, H. Yanagida, H. Taneichi, N. Hosono, T. Tsuji, T. Suzuki, H. Sudo, T. Kotani, I. Yonezawa, D. Londono, D. Gordon, J. A. Herring, K. Watanabe, K. Chiba, N. Kamatani, Q. Jiang, Y. Hiraki, M. Kubo, Y. Toyama, T. Tsunoda, C. A. Wise, Y. Qiu, C. Shukunami, M. Matsumoto, S. Ikegawa Genetic variants in GPR126 are associated with adolescent idiopathic scoliosis. Nat Genet. Vol. 45, pg. 676–679, 2013, https://doi.org/10.1038/ng.2639.

40 N. Makki, J. Zhao, Z. Liu, W. L. Eckalbar, A. Ushiki, A. M. Khanshour, J. Wu, J. Rios, R. S. Gray, C. A. Wise, N. Ahituv Genomic characterization of the adolescent idiopathic scoliosis-associated transcriptome and regulome. Hum Mol Genet. Vol. 29, pg. 3606–3615, 2020, https://doi.org/10.1093/hmg/ddaa242.