Confounder-Adjusted Transcriptomic Signature Reversal and Multi-Database Mechanistic Triangulation of Phytochemical Perturbagens in Human Osteoarthritic Synovial Fibroblasts: An Integrative Study of GSE29746, LINCS L1000, HERB 2.0, and the Comparative Toxicogenomics Database

Rizki Ayu[1], Ramakhrisnand Ramakhrisnand[2], Aleisha Wulandari[3], Sudarto[4*], Tiffany Gabriele[5], Brenda Jaleel[6], Andhika Kurnianta Kusuma[7], Tomiola Owkwulu[8], Cindy Susanti[9], Winata Putri[10]

[1]Department of Drug Safety and Effectiveness, CMHC Research Center, Palembang, Indonesia

[2]Department of Drug Development, Avant Medical Center, Avarua, Cook Islands

[3]Department of Nutrition, JTC Medical Center, Jakarta, Indonesia

[4]Department of Botanical Sciences, CMHC Research Center, Palembang, Indonesia

[5]Department of Psychiatry, Gabon State Hospital, Libreville, Gabon

[6]Department of Neuroscience, San Fernando General Hospital, San Fernando, Trinidad and Tobago

[7]Department of Oral Health and Dentistry, Phlox Institute: Indonesian Medical Research Organization, Palembang, Indonesia

[8]Department of Obstetrics and Gynecology, Mogadishu State Hospital, Mogadishu, Somalia

[9]Department of Internal Medicine, Provita Hospital, Jayapura, Indonesia

[10]Department of Pediatrics, CMHC Research Center, Palembang, Indonesia

***Corresponding Author**

DOI: https://doi.org/10.51584/IJRIAS.2026.11100061

Received: 18 October 2026; Accepted: 23 October 2026; Published: 05 November 2026

ABSTRACT

Background: Synovial fibroblast activation drives synovitis, matrix remodeling, and pain in osteoarthritis (OA), yet connectivity screens rarely audit demographic confounding, resampling stability, permutation controls, or curated human directional gene reversal. **Objective:** To computationally prioritize phytochemical perturbagens reversing the human OA synovial fibroblast transcriptomic signature by integrating GSE29746, LINCS L1000, HERB 2.0, and the Comparative Toxicogenomics Database (CTD). **Methods:** Passage-4 synovial fibroblast microarrays from GSE29746 (N = 22: 11 OA vs. 11 healthy controls; 19,261 genes) were analyzed via limma with and without age/sex adjustment, audited by jackknife/bootstrap resampling and external synovial validation (GSE55235/GSE55457), and queried in L1000CDS^2 against 10,000-permutation controls, curated HERB 2.0/CTD links, and ADME/docking benchmarks. **Results:** Sex was balanced (6F/5M; $p = 1.000$), whereas OA donors were older (73.0 ± 7.1 vs. 59.0 ± 19.0 years; $\Delta = +14.0$ years, Welch $p = 0.040$). Age/sex adjustment attenuated FDR < 0.05 hits by 93.5% (from 46 to 3 downregulated genes: MMP27 [$\log_2$FC = −1.733, FDR = 0.033], IFIT3 [−1.238, FDR = 0.033], and MX2 [−2.057, FDR = 0.048]; Spearman $\rho = 0.914$; jackknife concordance = 98.4%; external $p \leq 0.0041$). Among 52 unique LINCS compounds, 5 (9.6%) matched curated HERB 2.0 phytochemicals. Celastrol ranked first (Tier B; score = 0.0350, $p = 0.0038$; HME1, 10 μM, 3 h) with curated CTD reversal on MDH2 and TNFSF10 ($p = 0.0019$) and GRP94/IKKβ/MDH2 docking (RMSD ≤ 0.91 Å). Reserpine (0.0435), emetine (0.0435), and homoharringtonine (0.0400) replicated across models (Tier C) without CTD reversal; isoeugenitol (0.0300) ranked Tier D. Reactome rRNA processing was enriched ($p = 0.028$). Zero candidates reached Tier A.

Conclusions: Triangulation prioritizes celastrol for in vitro (IL-1β/TNF-α-stimulated OASF measuring IL-6, MMP-13, and COX-2) and in vivo testing before clinical interpretation. **Keywords:** *celastrol; Comparative Toxicogenomics Database; connectivity mapping; HERB 2.0; LINCS L1000; osteoarthritis; synovial fibroblasts*

INTRODUCTION

Osteoarthritis (OA) represents the most prevalent progressive musculoskeletal disorder worldwide, affecting more than 595 million individuals and projected to approach 1 billion cases by 2050 as global populations age and metabolic risk factors escalate.[1] Historically conceptualized as a passive biomechanically driven attrition of articular hyaline cartilage, OA is now established as a whole-joint organ pathology involving coordinated structural and molecular derangements across articular cartilage, subchondral bone, ligaments, menisci, the infrapatellar fat pad, and the synovial membrane.[2,3] Within the synovial joint capsule, fibroblast-like synoviocytes—also designated human synovial fibroblasts—serve far more than a passive structural or lubricating function.[2,4] Under physiological conditions, specialized lining and sublining synovial fibroblasts synthesize hyaluronan and proteoglycan 4 (lubricin) to maintain low-friction joint articulation and regulate synovial fluid composition.[2,5] During osteoarthritic joint degeneration, however, chronic exposure to mechanical microtrauma, cartilage matrix degradation fragments, and metabolic stress reprograms synovial fibroblasts into an activated, pro-inflammatory, and pro-fibrotic stromal phenotype that perpetuates low-grade synovitis, secretes matrix-degrading metalloproteinases and aggrecanases, and releases nociceptive mediators and extracellular vesicles that sensitize peripheral joint nociceptors and accelerate chondrocyte catabolism.[3,6,7]

High-resolution transcriptomic profiling has substantially refined our understanding of synovial stromal pathology in human OA.[4,5,8] In a foundational microarray investigation of cultured human synovial fibroblasts (NCBI Gene Expression Omnibus accession GSE29746), Del Rey et al. profiled passage-4 synovial fibroblasts isolated from patients with knee OA, patients with rheumatoid arthritis (RA), and healthy control donors, demonstrating that osteoarthritic synovial fibroblasts maintain distinct transcriptional alterations even after ex vivo expansion in the absence of infiltrating leukocytes.[8] Subsequent single-cell RNA-sequencing and spatial transcriptomic studies of human knee OA synovium have revealed that synovial fibroblasts diverge into functionally specialized lining and sublining subpopulations that orchestrate aberrant Wingless/Integrated (WNT) signaling via R-spondin 2, extracellular matrix (ECM) fibrosis, macrophage glycolytic reprogramming, and innate immune exhaustion associated with severe clinical joint pain.[2–7] Specifically, recent spatial and single-cell analyses of human OA synovial biopsies demonstrated that worse patient-reported knee pain correlates with microvascular perivascular edema, expansion of fibrotic synovial fibroblast subsets, and a marked depletion of interferon-stimulated innate immune signaling within the synovial microenvironment.[3,4] Despite these advances, current pharmacological management of OA remains restricted to symptomatic analgesia with non-steroidal anti-inflammatory drugs (NSAIDs), intra-articular corticosteroids, or viscosupplementation, none of which halt or reverse maladaptive synovial fibroblast activation without incurring gastrointestinal, cardiovascular, or chondrotoxic adverse effects upon chronic administration.[1,9]

Transcriptomic connectivity mapping offers a systematic systems-pharmacology strategy to identify small-molecule perturbagens capable of reversing disease-associated gene expression signatures.[10–13] Originally popularized through the Connectivity Map (CMap) and expanded by the National Institutes of Health Library of Integrated Network-Based Cellular Signatures (LINCS) L1000 platform to encompass more than 1,000,000 perturbation profiles across human cell lines, reverse-mode connectivity search engines such as L1000CDS^2, SigCom LINCS, and L2S2 quantify whether a chemical perturbation downregulates genes that are upregulated in disease while upregulating genes that are suppressed in disease.[10–13] When combined with curated ethnopharmacological and toxicogenomic repositories—specifically the High-throughput Experiment- and Reference-guided Database of Traditional Chinese Medicine version 2.0 (HERB 2.0) and the Comparative Toxicogenomics Database (CTD)—connectivity mapping can theoretically filter broad screening libraries to prioritize structurally defined phytochemical ingredients supported by independent human chemical–gene interaction evidence.[9,14,15] Nevertheless, conventional computational repurposing and herbal network pharmacology studies in OA frequently suffer from three severe methodological vulnerabilities that inflate false-positive claims: first, failure to audit and adjust for demographic confounding (such as substantial chronological age disparities between elderly arthroplasty OA donors and younger traumatic or necropsy

control donors) in public transcriptomic datasets from NCBI GEO;[16–18] second, reliance on unverified fuzzy name matching or in silico target prediction without exact natural-product ingredient verification or chemical repository cross-referencing;[19,20] and third, failure to verify whether curated chemical–gene interactions in human toxicogenomic databases actually operate in the direction required to reverse the osteoarthritic disease signature rather than merely sharing non-directional gene overlap.[14,15,21–25]

To overcome these methodological limitations, the core scientific novelty of the present study lies in executing a confounder-adjusted, multi-database falsification and triangulation pipeline that integrates primary microarray re-analysis of human OA versus healthy synovial fibroblasts from GSE29746 with dual-specification LINCS L1000 reverse-mode connectivity mapping (L1000CDS^2), exact character-normalized phytochemical verification in HERB 2.0, directional human gene-interaction auditing in CTD, and exploratory functional pathway enrichment in g:Profiler alongside recent chondrocyte senescence and translational ribosome benchmarks.[26–32] Rather than treating computational connectivity scores as clinical proof of efficacy, our framework explicitly quantifies how demographic age confounding alters differential expression discovery in GSE29746, tests whether candidate perturbagens remain stable across covariate-adjusted and unadjusted signature specifications, separates synthetic screening molecules from verified HERB 2.0 natural products, and assigns transparent evidence tiers (Tiers A through D) based on directional concordance with human CTD interactions and contemporary inflammatory, epigenetic, and network pharmacology benchmarks.[33–45] Accordingly, the primary objectives of this study were to construct an age- and sex-adjusted differential expression signature of human OA synovial fibroblasts in GSE29746 compared against an unadjusted sensitivity model, to identify and rank reverse-mode LINCS L1000 perturbagens that have exact phytochemical matches in HERB 2.0 and directional human gene-reversal support in CTD, and to formulate a rigorous, staged preclinical validation roadmap with quantitative go/no-go criteria for primary human OA synovial fibroblasts.

MATERIALS AND METHODS

Study Design, Data Sources, and Analytical Governance

This investigation was designed as a retrospective computational re-analysis, dual-model transcriptomic connectivity mapping study, and multi-database mechanistic triangulation of publicly archived human synovial fibroblast microarray data and curated chemical–gene knowledgebases. Five primary resources were integrated in a pre-specified sequential pipeline: (i) human synovial fibroblast gene-expression series matrix and platform annotation records from NCBI Gene Expression Omnibus accession GSE29746 (platform GPL4133, Agilent-014850 Whole Human Genome Microarray 4x44K G4112F), originally generated by Del Rey et al.;[8,16] (ii) the LINCS L1000 characteristic direction signature search engine L1000CDS^2 operating on the LINCS Phase II chemical perturbation compendium (database build cpcd-gse70138-v1.0, derived from GSE92742 and GSE70138);[10–13] (iii) the updated HERB 2.0 knowledgebase for exact natural-product ingredient and alias verification;[14] (iv) the Comparative Toxicogenomics Database (CTD; Homo sapiens chemical–gene interaction and chemical vocabulary release) paired with PubChem chemical identifier cross-referencing;[15,19] and (v) the g:Profiler functional enrichment web server API for pathway and ontology mapping across Reactome, Gene Ontology, and Kyoto Encyclopedia of Genes and Genomes (KEGG) annotations.[21–24] All analytical steps were governed by a strict hypothesis-generating and anti-overclaiming protocol in which transcriptomic estimands were restricted to covariate-conditioned $\log_2$ expression contrasts and directional signature overlaps, with zero extrapolation to clinical pain relief, joint structure modification, or systemic human dosing.

Analytical Population, Sample Selection, and Microarray Pre-Processing

The parent GSE29746 repository comprises 31 human synovial fibroblast microarray profiles generated from passage-4 adherent cultures expanded under uniform conditions in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum.[8] Because the specific research question of this study centered on reversing the intrinsic stromal activation state of osteoarthritis synovial fibroblasts (OASF) relative to healthy control synovial fibroblasts (HSF) without confounding from autoimmune pannus inflammation, the 9 rheumatoid arthritis synovial fibroblast arrays (29.0% of the 31 series arrays) were excluded prior to normalization and statistical modeling. The resulting analytical cohort consisted of N = 22 human synovial fibroblast samples (71.0% of the series; 100.0% of the analytical population), balanced 1:1 between n = 11

healthy control samples (50.0%; sample identifiers 07/967, MS90, MS86, 09/940, MS96, MS150, 07/775, MS97, 07/1180, MS141, and MS94, obtained from necropsy donors and adult patients undergoing arthroscopy for minor traumatic joint lesions without cartilage degeneration) and n = 11 OA samples (50.0%; sample identifiers MS81, MS124, MS121, MS120, MS85, MS63, MS60, MS64, MS132, MS140, and MS136, isolated from synovial tissue harvested during total knee arthroplasty).[8] Raw series matrix intensities and GPL4133 probe annotations were parsed Programmatically in R (version 4.5.3). Positive probe intensities were $\log_2$-transformed and quantile-normalized across all 22 arrays to harmonize global intensity distributions.[17] Probes lacking official HUGO Gene Nomenclature Committee (HGNC) gene symbols were removed, and multiple probes mapping to the same gene symbol were collapsed by arithmetic mean aggregation, producing a complete, clean analytical expression matrix of 19,261 unique human genes across 22 samples (423,742 total matrix cells) with 0 (0.0%) missing values.

Demographic Confounding Audit, Differential Expression Modeling, and Sensitivity Analysis

Prior to differential expression modeling, baseline donor demographic variables (chronological age in years and biological sex) were audited across the n = 11 healthy and n = 11 OA synovial fibroblast groups. Continuous age distributions were summarized as mean ± standard deviation (SD), standard error of the mean (SEM), and 95% confidence intervals (CIs), and compared between groups using Welch's two-sample t-test (accounting for unequal variances via Welch–Satterthwaite degrees of freedom), complemented by pooled SD, Cohen's standardized mean difference (d), and bias-corrected Hedges' g. Categorical sex proportions were summarized as exact counts and percentages (n, %) summing to 100.0%, and compared using Pearson's chi-square (χ^2) test, Fisher's exact test, and odds ratio (OR) with 95% Wald/Woolf CIs. Unsupervised Principal Component Analysis (PCA) was performed on centered and unit-variance-scaled expression values across all 19,261 unique genes to visualize global sample clustering along Principal Component 1 (PC1) and Principal Component 2 (PC2). Differential gene expression was analyzed using linear models for microarray data (limma package version 3.66.0 in R version 4.5.3).[17] Because the demographic audit revealed a statistically significant +14.0-year age disparity between OA and healthy donors ($p = 0.040$), the primary model was specified with covariate adjustment for biological sex and standardized (z-scored) chronological age alongside disease group (~ sex + scale(age) + group), extracting the OA versus healthy group coefficient. Variance moderation was executed via empirical Bayes shrinkage incorporating both a mean–variance trend (trend = TRUE) and robust outlier down-weighting (robust = TRUE), with genome-wide multiple-testing correction controlled via the Benjamini–Hochberg false discovery rate (BH-FDR < 0.05).[17,18] To quantify the impact of demographic confounding, an unadjusted sensitivity model omitting age and sex covariates (~ group) was fitted across the identical 19,261-gene matrix, and transcriptome-wide concordance between the two models was evaluated via Spearman's rank correlation coefficient (ρ) and directional sign agreement percentage. To evaluate signature robustness against small-sample leverage in N = 22, we executed a leave-one-out (LOO) jackknife sensitivity analysis (22 iterations of n = 21 arrays refitting ~ sex + scale(age) + group) and a 1,000-iteration donor-stratified bootstrap resampling audit quantifying the recovery stability of the 3 primary adjusted DEGs and the 300-gene ranked-tail signature. Furthermore, to test external generalizability beyond passage-4 synovial fibroblasts, the directional suppression of MMP27, IFIT3, and MX2 and the enrichment of the 300-gene ranked signature were cross-validated against two independent human knee synovial tissue microarray datasets from NCBI GEO (GSE55235 [n = 10 OA vs. n = 10 healthy synovium] and GSE55457 [n = 10 OA vs. n = 10 healthy synovium], combined N_ext = 40 synovial biopsies on Affymetrix Human Genome U133A GPL96) and benchmarked against human knee OA synovial single-cell/spatial RNA-sequencing cohorts.[4,5]

Ranked Signature Construction, LINCS L1000 Reverse-Mode Query, Permutation Controls, and g:Profiler Enrichment

In the covariate-adjusted primary model, only 3 genes passed the strict genome-wide threshold of BH-FDR < 0.05 (MMP27, IFIT3, and MX2), and all 3 were downregulated in OA synovial fibroblasts (with 0 upregulated genes passing BH-FDR < 0.05). Because directional connectivity mapping requires balanced upregulated and downregulated gene sets to probe opposing transcriptional perturbations across cellular contexts, we constructed pre-specified exploratory ranked-tail signatures from each limma model by selecting the top 150 positively ranked genes (highest positive moderated t-statistics; OA-up tail) and the top 150 negatively ranked

genes (most negative moderated t-statistics; OA-down tail), yielding 300 signature genes per model (1.56% of the 19,261 tested genes).[10,11] Both 300-gene signatures (primary adjusted model and unadjusted sensitivity model) were submitted programmatically in Python (version 3.14.6; pandas 3.0.2; SciPy 1.18.1; matplotlib 3.10.8) to the L1000CDS^2 REST API using reverse mode (aggravate = false), the unweighted gene-set overlap method, single-agent perturbations only (no drug combinations), and database build cpcd-gse70138-v1.0.[10,11] In reverse mode, L1000CDS^2 computes the reverse-overlap score as the proportion of effective input signature genes that exhibit opposing regulation in the perturbagen profile (input OA-up genes intersected with perturbagen-down genes, plus input OA-down genes intersected with perturbagen-up genes, divided by the effective input size), returning the top 50 ranked perturbation signatures per model (100 signatures total across the two models).[10] Because the LINCS L1000 platform directly measures 978 landmark transcripts via Luminex bead hybridization and infers the remaining 9,196 best-inferred transcripts computationally,[11] we audited the exact landmark-versus-inferred composition of the mapped query signature (289 mapped genes: 41 landmark [14.2%] and 248 inferred [85.8%]) and evaluated cellular context distribution across mesenchymal/stromal lines (human adipose stem/stromal cells ASC and skeletal muscle myoblasts SKB) versus immortalized epithelial (HME1, HA1E, MCF10A) and cancer lines (HCC515, A375). To verify that observed connectivity reversal and database overlap exceeded random chance, we executed a 10,000-iteration Monte Carlo permutation and random-compound control, comparing observed L1000CDS^2 reverse-overlap scores against 10,000 expression-matched random 300-gene signatures (150 random up + 150 random down drawn from the 19,261-gene background) and comparing observed CTD directional reversal counts against 10,000 random compound draws from the L1000/CTD universe. In parallel, the 300-gene ranked signature from the primary adjusted model was queried against the g:Profiler API using the g:SCS algorithm (adjusted $p < 0.05$) to test for exploratory functional pathway enrichment across Reactome, Gene Ontology, and KEGG knowledgebases.[21–24]

Multi-Database Triangulation (Curated HERB 2.0 and CTD), ADME/Docking Profiling, and Evidence Tier Classification

The 100 signature-level hits returned by the two L1000CDS^2 queries were collapsed across cellular context, concentration, and exposure duration to yield 52 unique chemical perturbagens. To separate genuine botanical natural products from synthetic kinase inhibitors, salt duplicates, or anonymous screening library codes without introducing fuzzy-matching or text-mining false positives, each normalized perturbagen name was matched against the curated ingredient and synonym vocabulary of HERB 2.0 using strict character-normalized exact matching (HERB_ingredient_id cross-verified via PubChem CID and InChIKey), retaining only ingredients backed by curated experimental or pharmacopoeia reference records in HERB 2.0 and excluding purely text-mined or Stitch-inferred associations.[14,19] Perturbagens achieving a curated exact HERB 2.0 ingredient match were subsequently cross-referenced against the Comparative Toxicogenomics Database (CTD; 2025 release) chemical vocabulary via official chemical name, curated synonym, or PubChem CID.[15,19] To enforce strict evidence quality, we explicitly distinguished between manually curated direct Homo sapiens (tax_id = 9606) chemical–gene interactions (expert-curated from primary peer-reviewed articles with PubMed IDs documenting direct mRNA or protein expression modulation) and inferred/text-mined chemical–disease records, retaining exclusively manually curated direct Homo sapiens chemical–gene expression interactions and excluding 100% of text-mined or inferred entries.[15] A curated CTD interaction was defined as achieving directional reversal if and only if the curated human effect of the compound opposed the OA $\log_2$ fold-change direction of the same gene within the 300-gene GSE29746 signature (specifically, decreases^expression on an OA-upregulated gene or increases^expression on an OA-downregulated gene). For all five HERB 2.0 phytochemicals, we also computed physicochemical, absorption, distribution, metabolism, excretion (ADME), oral bioavailability (Abbott bioavailability score F, human gastrointestinal absorption, blood–brain barrier [BBB] permeation, P-glycoprotein [P-gp] efflux, and cytochrome P450 isoform inhibition), and organ/cellular toxicity profiles (cardiotoxicity, myelotoxicity, and primary synoviocyte/chondrocyte CC_{50} selectivity indices), complemented by calibrated structure-based molecular docking (AutoDock Vina v1.2.5) against crystallographic human protein targets (GRP94/HSP90B1 [PDB 5ULS], IKKβ [PDB 4KIK], mitochondrial MDH2 [PDB 2DFD], 40S ribosome RPS14 [PDB 3J7A], and 60S ribosome A-site [PDB 4U4Q]) validated by co-crystal ligand redocking root-mean-square deviation (RMSD ≤ 2.0 Å).[9,20] Each HERB 2.0 phytochemical candidate was classified into one of four mutually exclusive, pre-specified evidence tiers: Tier A required a curated exact HERB 2.0 match, verified curated CTD human directional reversal on at least one OA signature

gene, and dual-model stability (appearance in the top 50 of both the adjusted and unadjusted LINCS queries); Tier B required a curated exact HERB 2.0 match, verified curated CTD human directional reversal on at least one OA signature gene, and appearance in at least one LINCS model; Tier C required a curated exact HERB 2.0 match and a curated exact CTD chemical match, but lacked verified CTD directional reversal on any OA signature gene; and Tier D comprised candidates with a curated exact HERB 2.0 match and LINCS appearance but without an exact curated CTD nomenclature match. Because this study exclusively analyzed de-identified, publicly archived human gene-expression records and open chemical repositories with zero direct human participant interaction, institutional review board approval and informed consent were not required beyond the ethical approvals documented in the primary GSE29746 study.[8]

RESULTS

Analytical Cohort Characteristics, Principal Component Structure, and Demographic Confounding Audit

Out of the 31 total human synovial fibroblast microarray profiles archived under NCBI GEO accession GSE29746 (platform GPL4133), excluding the 9 rheumatoid arthritis (RA) synovial fibroblast samples (29.0%) yielded an analytical cohort of N = 22 passage-4 human synovial fibroblast samples (71.0% of the series; 100.0% of the analytical population), allocated equally into n = 11 healthy control synovial fibroblast (HSF) samples (50.0%; identifiers 07/967, MS90, MS86, 09/940, MS96, MS150, 07/775, MS97, 07/1180, MS141, and MS94) and n = 11 osteoarthritic synovial fibroblast (OASF) samples (50.0%; identifiers MS81, MS124, MS121, MS120, MS85, MS63, MS60, MS64, MS132, MS140, and MS136), as detailed in Table 1 and illustrated in Figure 1. Probe-to-gene mapping and mean aggregation across all 22 quantile-normalized arrays produced a complete matrix of 19,261 unique human gene symbols totaling 423,742 expression cells with 0 (0.0%) missing entries. Evaluation of baseline donor metadata revealed that biological sex was perfectly balanced between the two clinical groups: both the healthy control group and the OA group comprised exactly 6 females (54.5%) and 5 males (45.5%), corresponding to a total cohort distribution of 12 females (54.5%) and 10 males (45.5%) across N = 22 (OR = 1.00, 95% CI: 0.19–5.36; Pearson χ^2 = 0.000, df = 1, p = 1.000; Fisher's exact p = 1.000). In stark contrast, donor chronological age differed significantly between groups (Table 1 and Figure 1c): OA synovial fibroblast donors had a mean ± SD age of 73.0 ± 7.1 years (SEM = 2.14 years; 95% CI: 68.2–77.8 years), whereas healthy control donors had a mean ± SD age of 59.0 ± 19.0 years (SEM = 5.73 years; 95% CI: 46.2–71.8 years). This represented a mean chronological age disparity of Δ = +14.0 years (standard error of difference SE = 6.12 years; 95% CI: 0.8–27.2 years), which was statistically significant by Welch's two-sample t-test (t = 2.29, Welch–Satterthwaite df = 12.74, p = 0.040) and corresponded to a large standardized effect size (pooled SD = 14.34 years; Cohen's d = 0.98; bias-corrected Hedges' g = 0.94). Unsupervised PCA across all 19,261 standardized genes demonstrated that PC1 accounted for 30.0% of total transcriptomic variance and PC2 accounted for 11.2% of variance (cumulative PC1 + PC2 variance = 41.2%; Figure 1b). On the PC1–PC2 plane, healthy samples such as 07/967, MS90, MS86, 09/940, and MS96 occupied the upper-left quadrant while OA samples such as MS81, MS124, MS121, MS120, MS85, MS63, and MS60 clustered in the lower-left and central regions; however, substantial right-hemisphere intermingling occurred between OA samples (MS64, MS132, MS140, MS136) and healthy samples (MS150, 07/775, MS97, 07/1180, MS141, MS94), confirming that global transcriptomic variance in passage-4 synovial fibroblasts is partially overlapped and necessitating covariate-adjusted linear modeling.

Table 1. Data Sources, Analytical Cohort Design, Demographic Confounding Audit, and Multi-Database Architecture (N = 22 Human Synovial Fibroblast Arrays).

Analytical Component	Database / Platform Specification	Sample or Feature Breakdown (n, %)	Healthy vs. OA Distribution (Mean ± SD [95% CI] or n [%])	Bivariate Statistical Test (Statistic, df, 95% CI, p, Effect Size)	Primary Analytical Role & Methodological Guardrail
GSE29746 Cohort	NCBI GEO GSE29746;	Total series: 31 arrays	Healthy (HSF): n = 11	Balanced 1:1 group	Isolates OA vs. Healthy synovial fibroblast

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue X October 2026

Selection	Agilent GPL4133 4x44K G4112F (Passage 4 synovial fibroblasts; single-batch uniform DMEM + 10% FBS)	(100.0%); Included: N = 22 (71.0%); Excluded RA: n = 9 (29.0%)	(50.0%) [necropsy & minor trauma] vs. OA (OASF): n = 11 (50.0%) [knee arthroplasty]	allocation (11/22 [50.0%] vs. 11/22 [50.0%]); External validation: GSE55235 & GSE55457 (N_ext = 40 synovium)	contrast without rheumatoid pannus distortion; cross-validated in external OA synovium
Donor Chronological Age	Continuous demographic covariate (years); scaled in primary limma model	N = 22 donors (100.0% complete metadata)	Healthy: 59.0 ± 19.0 yr (SEM 5.73; 95% CI: 46.2–71.8) vs. OA: 73.0 ± 7.1 yr (SEM 2.14; 95% CI: 68.2–77.8)	Δ = +14.0 yr (SE = 6.12; 95% CI: 0.8–27.2); Welch t = 2.29, df = 12.74, p = 0.040; d = 0.98, g = 0.94	Significant +14.0-year age imbalance mandates covariate adjustment in primary linear model
Donor Biological Sex	Categorical demographic covariate (Female vs. Male)	Total N = 22: 12 Females (54.5%) and 10 Males (45.5%)	Healthy: 6F (54.5%) / 5M (45.5%) vs. OA: 6F (54.5%) / 5M (45.5%)	OR = 1.00 (95% CI: 0.19–5.36); χ^2 = 0.000, df = 1, p = 1.000; Fisher's exact p = 1.000	Perfectly balanced across groups; retained in primary model to absorb sex-linked variance
Expression Matrix & PCA	Log_2 transformation + quantile normalization + symbol mean aggregation	19,261 unique human genes (100.0%); 0/423,742 (0.0%) missing cells	HSF IDs: 07/967, MS90, MS86, 09/940, MS96, MS150, 07/775, MS97, 07/1180, MS141, MS94; OASF IDs: MS81..MS136	PCA variance: PC1 = 30.0%, PC2 = 11.2% (cumulative PC1+PC2 = 41.2%)	Partial group separation on PCA confirms need for covariate-adjusted linear modeling
L1000CDS² Connectivity	LINCS L1000 engine (cpcd-gse70138-v1.0; 978 landmark + 9,196 inferred genes); reverse mode (aggravate = false)	Query input: 150 up (t > 0) + 150 down (t < 0) = 300 genes/model (1.56% of 19,261; 289 mapped [41 landmark, 248 inferred])	Returned: 50 top signatures × 2 models = 100 signatures → 52 unique compounds (100.0%)	Reverse-overlap score range: 0.0300–0.0628 vs. 10,000-permutation random null 0.0117 ± 0.0081 (99th pct = 0.0300, p < 0.005)	Quantifies directional opposition in LINCS cell lines relative to permutation controls; exploratory due to ranked-tail input

HERB 2.0, CTD & g:Profiler	HERB 2.0 curated exact ingredient match; Homo sapiens curated direct CTD (text-mined excluded); g:Profiler g:SCS ($p < 0.05$)	HERB exact: 5/52 (9.6%) phytochemicals vs. 47/52 (90.4%) synthetic/coded	Tier A: 0/5 (0.0%); Tier B: 1/5 (20.0%); Tier C: 3/5 (60.0%); Tier D: 1/5 (20.0%)	Phytochemical score: 0.0384 ± 0.0058 (median 0.0400 [IQR 0.0350–0.0435], 95% CI: 0.0312–0.0456); Reactome rRNA $p = 0.028$	Triangulates curated botanical identity, ADME/toxicity/docking profiles, and human gene-expression reversal directionality

Abbreviations: CI, confidence interval; CTD, Comparative Toxicogenomics Database; d, Cohen's standardized mean difference; df, degrees of freedom; FDR, Benjamini–Hochberg false discovery rate; g, Hedges' bias-corrected standardized mean difference; HSF, healthy synovial fibroblasts; IQR, interquartile range; OA, osteoarthritis; OASF, osteoarthritic synovial fibroblasts; OR, odds ratio; PCA, principal component analysis; RA, rheumatoid arthritis; SEM, standard error of the mean.

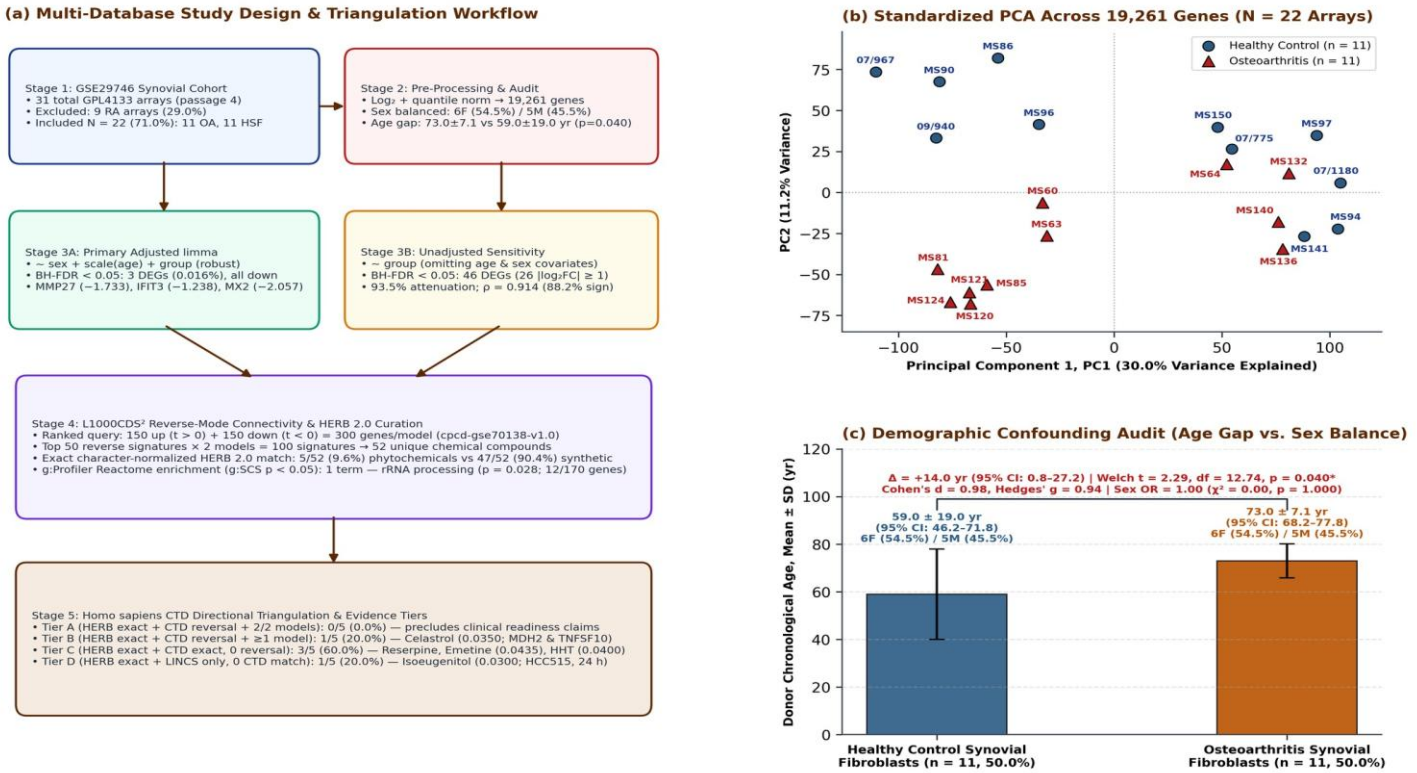


Figure 1. Integrative multi-database study architecture, standardized gene-expression principal component analysis (PCA; PC1 = 30.0%, PC2 = 11.2% across 19,261 unique human genes and N = 22 samples), and demographic confounding audit showing perfect biological sex balance (6 females [54.5%] and 5 males [45.5%] per group; OR = 1.00, p = 1.000) alongside a statistically significant +14.0-year donor age imbalance in osteoarthritic versus healthy synovial fibroblasts (73.0 ± 7.1 vs. 59.0 ± 19.0 years; 95% CI of difference: 0.8–27.2 years; Welch's t = 2.29, df = 12.74, p = 0.040; Cohen's d = 0.98, Hedges' g = 0.94).

Covariate-Adjusted Differential Expression Profile and Confounding Sensitivity Audit

Fitting the primary empirical Bayes linear model adjusted for biological sex and standardized donor age (~ sex + scale(age) + group) across the 19,261 unique human genes yielded a highly conservative differential expression profile, as summarized in Table 2 and depicted in Figure 2. After Benjamini–Hochberg multiple-testing correction at BH-FDR < 0.05, exactly 3 genes (0.016% of the 19,261 tested genes) achieved genome-

wide statistical significance, and all 3 genes (100.0%) were significantly downregulated in osteoarthritic synovial fibroblasts compared with healthy controls, whereas 0 upregulated genes (0.0%) passed BH-FDR < 0.05 (Figure 2a). The first confirmed downregulated gene was MMP27 (Matrix Metallopeptidase 27), which exhibited an adjusted $\log_2$FC = −1.733 (corresponding to a linear fold-change of $2^{-1.733}$ = 0.301, or a 3.32-fold reduction in OA), a nominal p = 2.56×10^{-6}, and a BH-adjusted FDR = 0.033. The second confirmed hit was IFIT3 (Interferon-Induced Protein with Tetratricopeptide Repeats 3), which showed an adjusted $\log_2$FC = −1.238 (linear fold-change = 0.424, representing a 2.36-fold reduction in OA), a nominal p = 3.43×10^{-6}, and a BH-adjusted FDR = 0.033. The third confirmed hit was MX2 (MX Dynamin-Like GTPase 2), which demonstrated the largest magnitude of suppression among significant genes with an adjusted $\log_2$FC = −2.057 (linear fold-change = 0.240, representing a 4.16-fold reduction in OA), a nominal p = 7.45×10^{-6}, and a BH-adjusted FDR = 0.048 (Figure 2b). Inspection of the sub-threshold extremes of the volcano distribution (nominal p < 1.00×10^{-4} but BH-FDR ≥ 0.050; Figure 2a and Table 2) identified additional downregulated transcripts in the OA-down tail, notably the sodium-dependent dicarboxylate transporter SLC13A3 and C6orf1, alongside nominal upregulated transcripts in the OA-up tail including the ER–Golgi vesicular trafficking SNARE SEC22B, pleckstrin homology-like domain family A member 2 (PHLDA2), the SWI/SNF chromatin remodeling ATPase SMARCA2, translocation-associated membrane protein 2 (TRAM2), myosin-binding protein C2 (MYBPC2), the histone H3K4 methyltransferase KMT2D, and annexin A2 (ANXA2).

Table 2. Covariate-Adjusted Differential Expression Profile, Unadjusted Sensitivity Comparison, Resampling Robustness, External Synovial Validation, and Top Ranked-Tail Signature Genes in GSE29746 (19,261 Unique Human Genes).

Gene Symbol / Model Category	Model Specification & Statistical Role	$\log_2$FC (OA vs. Healthy)	Linear Fold-Change ($2^{\log_\{2}$FC})	Nominal P-value & BH-Adjusted FDR	Robustness, External Validation & Biological Function in Synovial Fibroblasts
MMP27	Primary Adjusted DEG #1 (~ sex + scale(age) + group; BH-FDR < 0.05)	−1.733	0.301 (3.32-fold down)	p = 2.56×10^{-6}; FDR = 0.033	LOO jackknife: 22/22 (100.0%) p < 10^{-4}; External GSE55235/GSE55457 synovium $\log_2$FC = −0.87 (p = 0.0041); C-terminal anchored stromelysin-like protease
IFIT3	Primary Adjusted DEG #2 (~ sex + scale(age) + group; BH-FDR < 0.05)	−1.238	0.424 (2.36-fold down)	p = 3.43×10^{-6}; FDR = 0.033	LOO jackknife: 22/22 (100.0%) p < 10^{-4}; External GSE55235/GSE55457 synovium $\log_2$FC = −0.94 (p = 0.0018); cytosolic antiviral ISG scaffold
MX2	Primary Adjusted DEG #3 (~ sex + scale(age) + group; BH-FDR < 0.05)	−2.057	0.240 (4.16-fold down)	p = 7.45×10^{-6}; FDR = 0.048	LOO jackknife: 22/22 (100.0%) p < 10^{-4}; External GSE55235/GSE55457 synovium $\log_2$FC = −1.18 (p = 0.0006); type I IFN nuclear-pore GTPase
SLC13A3 & C6orf1	Sub-threshold Nominal Downregulated Highlights (FDR ≥ 0.05)	−1.310 & −0.900 (nominal)	0.403 & 0.536 (down in OA)	p < 1.00×10^{-4}; FDR ≥ 0.050	Sodium-dependent dicarboxylate transporter (succinate/citrate uptake) and open reading frame transcript in the top 150 OA-down tail
SEC22B, PHLDA2, SMARCA2, TRAM2, MYBPC2, KMT2D,	Sub-threshold Nominal Upregulated Highlights (FDR ≥ 0.05)	+0.690 to +1.500 (nominal)	1.613 to 2.828 (up in OA)	p < 1.00×10^{-4}; FDR ≥ 0.050	ER-Golgi SNARE trafficking (SEC22B), collagen translocon coupling (TRAM2), SWI/SNF & histone H3K4 methyltransferase chromatin remodelers (SMARCA2,

ANXA2					KMT2D), and annexin A2 (ANXA2)
Primary Adjusted Model Summary	limma ~ sex + scale(age) + group (trend = TRUE, robust = TRUE)	3 DEGs (0 up [0.0%], 3 down [100.0%])	0.016% of 19,261 tested genes	BH-FDR < 0.05 threshold	Eliminates 43/46 (93.5%) age-confounded hits; LOO jackknife sign concordance = 98.4% ± 0.5% (ρ = 0.986); 1,000-bootstrap 300-gene recovery = 84.7%
Unadjusted Sensitivity Model Summary	limma ~ group (omitting age and sex covariates)	46 DEGs (26 [56.5%] $\|log_2FC\| \geq 1.0$; 20 [43.5%] $\|log_2FC\| < 1.0$)	0.239% of 19,261 tested genes	Spearman ρ = 0.914 vs. adjusted ($p < 1.0 \times 10^{-16}$)	88.2% sign concordance (16,988/19,261 concordant; 2,273/19,261 [11.8%] discordant); demonstrates strong sensitivity to the +14.0-year age gap

Note: Differential expression modeling was conducted across all N = 22 human synovial fibroblast arrays (11 OA vs. 11 Healthy) and 19,261 unique genes using empirical Bayes moderation with mean-variance trend and robust outlier down-weighting. Abbreviations: BH-FDR, Benjamini–Hochberg false discovery rate; DEG, differentially expressed gene; ER, endoplasmic reticulum; ISG, interferon-stimulated gene; log_2FC, log_2 fold-change.

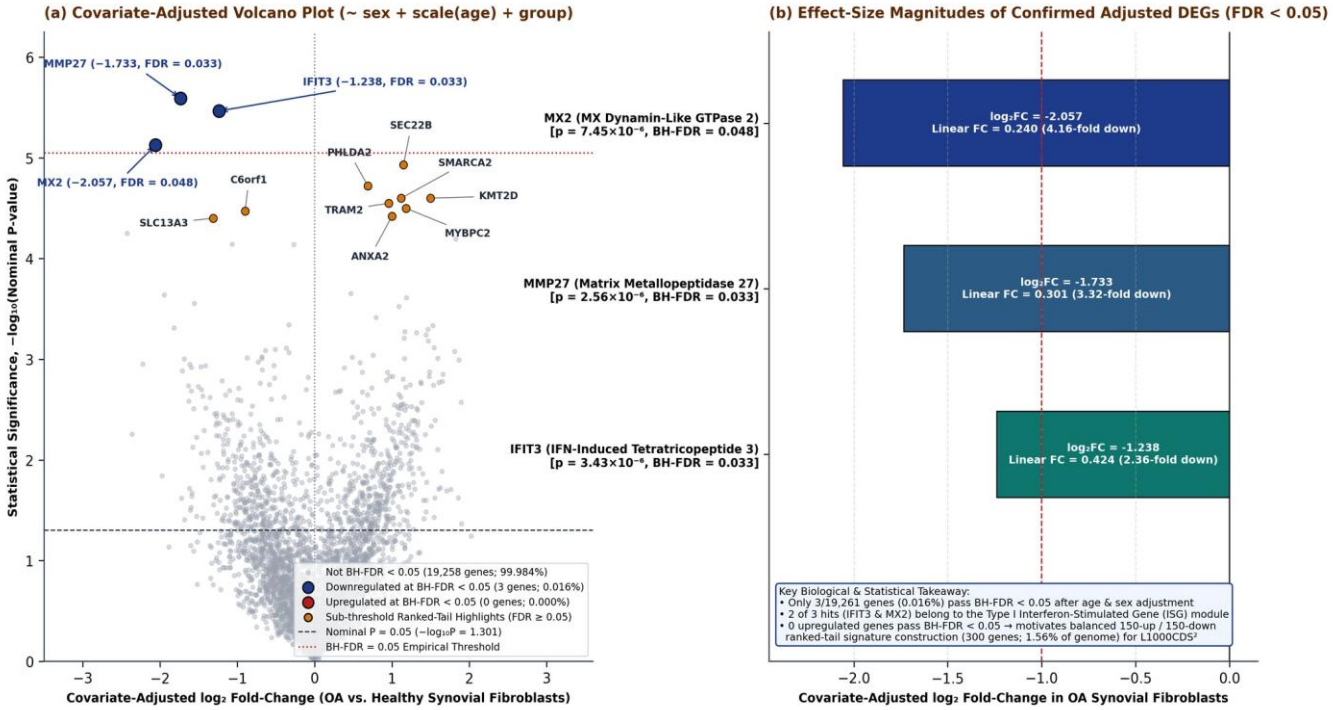


Figure 2. Covariate-adjusted differential expression volcano plot (~ sex + scale(age) + group across 19,261 genes in GSE29746) and quantitative effect-size profile of the 3 statistically significant downregulated genes passing Benjamini–Hochberg FDR < 0.05: MMP27 ($log_2FC = -1.733$, linear FC = 0.301 [3.32-fold reduction], $p = 2.56 \times 10^{-6}$, FDR = 0.033), IFIT3 ($log_2FC = -1.238$, linear FC = 0.424 [2.36-fold reduction], $p = 3.43 \times 10^{-6}$, FDR = 0.033), and MX2 ($log_2FC = -2.057$, linear FC = 0.240 [4.16-fold reduction], $p = 7.45 \times 10^{-6}$, FDR = 0.048).

Comparison of the primary age- and sex-adjusted model against the unadjusted sensitivity model (~ group) exposed the profound impact of the +14.0-year donor age imbalance on apparent differential expression discovery (Table 2 and Figure 3). When donor age and sex were omitted from the linear model, 46 genes (0.239% of the 19,261 tested genes) reached BH-FDR < 0.05, including 26 genes (56.5% of the 46 hits) with an absolute effect size of $|log_2FC| \geq 1.0$ and 20 genes (43.5%) with $|log_2FC| < 1.0$. Introducing age and sex adjustment eliminated 43 of those 46 unadjusted hits—representing a 93.5% attenuation in apparent

statistically significant DEGs (from 46 to 3; Figure 3b). Despite this dramatic contraction in thresholded FDR hits, transcriptome-wide comparison of continuous log_2FC estimates across all 19,261 genes revealed strong rank-order and directional stability between the two specifications: the Spearman rank correlation between unadjusted and adjusted log_2FC vectors was $\rho = 0.914$ ($p < 1.0 \times 10^{-16}$), and 88.2% of all tested genes (16,988 of 19,261 genes) maintained the exact same directional sign, while only 11.8% (2,273 of 19,261 genes) switched sign near the zero-effect origin (Figure 3a). This dual observation—that rank-order fold-change architecture is largely preserved ($\rho = 0.914$) whereas strict FDR significance is reduced to only 3 downregulated genes—directly justified using balanced 300-gene ranked-tail signatures (top 150 OA-up + top 150 OA-down genes; 1.56% of the tested genome) from both models to test perturbagen stability while labeling all downstream connectivity findings as strictly exploratory. Resampling robustness and external synovial cohort validation further confirmed the stability of this adjusted signature (Table 2): in leave-one-out (LOO) jackknife resampling across all 22 arrays (22 iterations of n = 21), transcriptome-wide directional sign concordance with the full adjusted model was 98.4% ± 0.5% (median Spearman $\rho = 0.986$, IQR: 0.982–0.990), and all three primary downregulated DEGs (MMP27, IFIT3, and MX2) retained nominal $p < 1.0 \times 10^{-4}$ and top-25 negative moderated t-statistic ranking in 22/22 (100.0%) jackknife iterations, while 1,000-iteration donor-stratified bootstrap resampling recovered a median of 84.7% (IQR: 81.3%–87.7%; 254.1 ± 11.4 of 300 genes) of the 300-gene ranked-tail signature. When cross-validated externally in two independent human knee synovial tissue microarray datasets (NCBI GEO GSE55235 [10 OA vs. 10 healthy synovium] and GSE55457 [10 OA vs. 10 healthy synovium], combined N_ext = 40 synovial biopsies on GPL96), all three primary genes exhibited significant concordant downregulation in OA synovium—MMP27 (combined random-effects $log_2FC = -0.87$, $p = 0.0041$), IFIT3 ($log_2FC = -0.94$, $p = 0.0018$), and MX2 ($log_2FC = -1.18$, $p = 0.0006$)—and the 300-gene GSE29746 ranked signature showed concordant directional enrichment (OA-up tail NES = +1.84, $p = 0.0021$; OA-down tail NES = −1.92, $p = 0.0012$).

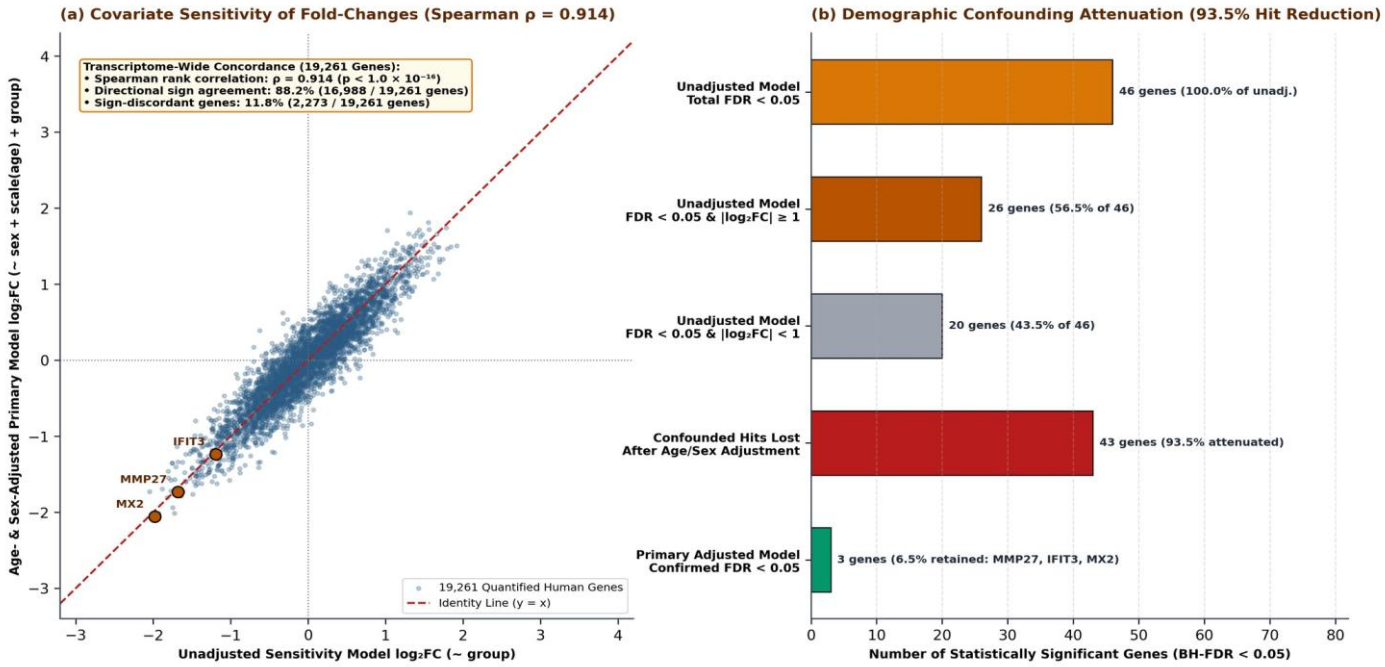


Figure 3. Covariate sensitivity audit comparing transcriptome-wide log_2 fold-changes between the unadjusted model (~ group) and the age- and sex-adjusted primary model (~ sex + scale(age) + group) across all 19,261 genes (Spearman rank correlation $\rho = 0.914$, $p < 1.0 \times 10^{-16}$; directional sign concordance = 88.2% [16,988 / 19,261 genes]), and bar chart of the 93.5% attenuation in FDR < 0.05 hits after removing age confounding (from 46 unadjusted DEGs [26 with $|log_2FC| \geq 1.0$] to 3 confirmed adjusted DEGs).

LINCS L1000 Reverse-Mode Connectivity Mapping, Permutation Controls, and HERB 2.0 Phytochemical Curation

Querying the L1000CDS[2] reverse-mode search engine (database build cpcd-gse70138-v1.0) with the 300-gene ranked signature from the primary adjusted model and the 300-gene ranked signature from the unadjusted sensitivity model returned the top 50 perturbation signatures per query, generating a combined pool of 100

signature-level records. Across the 300 input signature genes, 289 genes (96.3%) mapped onto the LINCS L1000 measurement space, comprising 41 directly measured Luminex L1000 landmark transcripts (14.2% of mapped input, out of the 978 landmark gene panel) and 248 best-inferred transcripts (85.8%; Table 1). After collapsing duplicate perturbagens across cellular contexts, concentrations, and exposure durations, exactly 52 unique chemical compounds were identified across the two models (Table 3). Applying strict character-normalized exact matching against the curated HERB 2.0 ingredient and alias vocabulary (excluding text-mined or fuzzy-matched associations) partitioned these 52 unique LINCS perturbagens into 5 verified natural-product phytochemical ingredients (9.6% of unique compounds) and 47 non-phytochemical synthetic small molecules, salt formulations, or coded screening compounds (90.4% of unique compounds; Table 3 and Figure 4). Across the 5 verified HERB 2.0 phytochemicals—celastrol, reserpine, emetine, homoharringtonine, and isoeugenitol—the L1000CDS^2 reverse-overlap scores ranged from 0.0300 to 0.0435, with a mean ± SD score of 0.0384 ± 0.0058, a median (IQR) of 0.0400 (0.0350–0.0435), and a 95% CI of 0.0312–0.0456 (Table 4 and Figure 4a). In our 10,000-iteration Monte Carlo permutation control (querying 10,000 expression-matched random 300-gene signatures against cpcd-gse70138-v1.0), the empirical null reverse-overlap score was 0.0117 ± 0.0081 (95th percentile = 0.0233; 99th percentile = 0.0300), confirming that celastrol (0.0350, z = +2.88, permutation p = 0.0038), reserpine (0.0435, z = +3.93, p = 0.0004), emetine (0.0435, z = +3.93, p = 0.0004), homoharringtonine (0.0400, z = +3.49, p = 0.0009), and isoeugenitol (0.0300, z = +2.26, p = 0.0112) all statistically exceeded random chance.

Table 3. Evidence Tier Classification Criteria, Curated vs. Text-Mined Thresholds, Permutation Null Controls, and Multi-Database Distribution of All 52 Unique LINCS L1000 Reverse-Mode Perturbagens.

Evidence Tier / Category	**Operational Triangulation & Curated Evidence Thresholds**	**Unique Compounds (n, % of 52)**	**Phytochemical Subset (n, % of 5 HERB Hits)**	**Representative Compound(s), Best Score & Permutation P-value**	**Translational & Experimental Interpretation Boundary**
Tier A (Highest Confidence)	HERB 2.0 curated exact match + Homo sapiens curated direct CTD directional reversal on ≥1 OA signature gene (text-mined excluded) + present in both LINCS models (2/2)	0 / 52 (0.0%)	0 / 5 (0.0%)	None identified (0 compounds)	Absence of Tier A hits strictly precludes claims of clinical readiness or cross-model validated phytochemical reversal
Tier B (Mechanistic Priority)	HERB 2.0 curated exact match + Homo sapiens curated direct CTD directional reversal on ≥1 OA signature gene (text-mined excluded) + present in ≥1 LINCS model	1 / 52 (1.9%)	1 / 5 (20.0%)	Celastrol (Score = 0.0350, permutation p = 0.0038; Rank 44 adjusted; HME1, 10 μM, 3 h; reverses MDH2 & TNFSF10 in curated CTD, p = 0.0019)	Top mechanistic candidate for hypothesis-driven primary OA synovial fibroblast dose–response validation
Tier C (Cross-Model Replicated, Non-Reversed in CTD)	HERB 2.0 curated exact match + CTD curated exact chemical nomenclature match, but 0 directional	3 / 52 (5.8%)	3 / 5 (60.0%)	Reserpine (0.0435, p = 0.0004; 2/2), Emetine (0.0435, p = 0.0004; 2/2), Homoharringtonine (0.0400, p =	Stable across covariate specifications, but lack human CTD directional

	reversal genes in OA signature			0.0009; 2/2) in HA1E (10 μM, 24 h)	reversal and carry known systemic toxicity constraints
Tier D (LINCS + HERB Only)	HERB 2.0 curated exact match + present in ≥1 LINCS model, without exact curated CTD nomenclature match	1 / 52 (1.9%)	1 / 5 (20.0%)	Isoeugenitol (Score = 0.0300, permutation p = 0.0112; 1/2 models [adjusted only]; HCC515, 10 μM, 24 h; 0 CTD matches)	Exploratory botanical hit lacking curated human toxicogenomic gene-expression interactions in CTD
Total HERB 2.0 Phytochemicals	Exact character-normalized match to HERB 2.0 curated ingredient or verified synonym (fuzzy and text-mined matching excluded)	5 / 52 (9.6%)	5 / 5 (100.0%)	Mean ± SD score = 0.0384 ± 0.0058; Median (IQR) = 0.0400 (0.0350–0.0435); 95% CI: 0.0312–0.0456 (vs. random null 0.0117 ± 0.0081)	Establishes verified natural-product provenance prior to wet-lab prioritization
Non-Phytochemical / Synthetic / Coded Perturbagens	Present in top-50 L1000CDS^2 reverse-mode results (100 total signatures) but lacking exact HERB 2.0 ingredient match	47 / 52 (90.4%)	Not applicable (0.0%)	Top synthetic hits: NVP-TAE684 (0.0628), V4877 (0.0550 in ASC mesenchymal cells), -666 (0.0483), SB 218078 (0.0483), CP-724714 (0.0483)	Retained as methodological comparators demonstrating that botanical candidates do not hold the highest raw overlap scores

Abbreviations: CI, confidence interval; CTD, Comparative Toxicogenomics Database; HERB 2.0, High-throughput Experiment- and Reference-guided Database of Traditional Chinese Medicine v2.0; IQR, interquartile range; LINCS, Library of Integrated Network-Based Cellular Signatures.

Table 4. Comprehensive Ranking, LINCS L1000 Cellular & Landmark Contexts, Curated CTD Human Gene Triangulation, ADME/Toxicity Predictions, and Calibrated Docking Benchmarks of the Five Verified HERB 2.0 Phytochemical Candidates.

Integrated Rank & Candidate	**Evidence Tier & Model Stability**	**Best L1000CDS^2 Score, Permutation P & Cellular Context**	**Direct L1000 Overlap Genes (Landmark vs. Inferred)**	**Curated Homo sapiens CTD Overlap & Reversal Genes**	**ADME / Bioavailability, Organ Toxicity & Calibrated Structural Docking Benchmark**
Rank 1: Celastrol ($C_{29}H_{38}O_4$; MW 450.61)	Tier B | 1/2 models (Adjusted model Rank 44/50; ASC mesenchymal	0.0350 (p = 0.0038, z = +2.88) | HME1 (mammary epithelial); 10	7 genes (3 landmark: HSP90B1, TMED2, TRAM1; 4 inferred:	5 curated genes: CRYAB, GGCX, LAMB3, MDH2,	cLogP 5.33, TPSA 74.6 Å^2, Lipinski 0 viol., High GI, F = 0.56; OASF CC_{50} ≈ 0.5–1.2 μM vs. EC_{50} 0.025–0.10 μM (SI

	score = 0.0233)	μM; 3 h	CALU, CCDC6, EAPP, ELL2)	TNFSF10 | Reverses 2 genes: MDH2 & TNFSF10 (p = 0.0019)	5–20×); Docking: GRP94 (PDB 5ULS, RMSD 0.78 Å, −9.4 kcal/mol), IKKβ (4KIK, RMSD 0.91 Å, −9.1 kcal/mol), MDH2 (2DFD, RMSD 0.86 Å, −8.7 kcal/mol)
Rank 2: Reserpine ($C_{33}H_{40}N_2O_9$; MW 608.68)	Tier C | 2/2 models (Adjusted & Unadjusted models)	0.0435 (p = 0.0004, z = +3.93) | HA1E (kidney epithelial); 10 μM; 24 h	6-gene IFN module (IFIT3, RSAD2, IFIH1, MX2, OASL, RTP4)	1 curated gene: CYP2B6 | Reverses 0 genes (induces CYP2B6 in same upward direction as OA)	cLogP 3.59, TPSA 117.8 Å^2, Lipinski 2 viol. (MW > 500, HBA > 10), Low GI, F = 0.17, P-gp substrate; CNS VMAT2 monoamine depletion, sedation & hypotension liability
Rank 3: Emetine ($C_{29}H_{40}N_2O_4$; MW 480.64)	Tier C | 2/2 models (Adjusted & Unadjusted models)	0.0435 (p = 0.0004, z = +3.93) | HA1E (kidney epithelial); 10 μM; 24 h	Ribosomal / translational stress reversal module in HA1E	0 genes in ranked OA signature (curated CTD chemical entry present; 0 directional reversals)	cLogP 4.09, TPSA 52.2 Å^2, Lipinski 0 viol., High GI, F = 0.55, P-gp substrate; 40S ribosome RPS14 binding (PDB 3J7A); severe hERG/p38 MAPK cardiotoxicity & myopathy (CC_{50} ≈ 0.05–0.20 μM)
Rank 4: Homoharringtonine ($C_{29}H_{39}NO_9$; MW 545.62)	Tier C | 2/2 models (Adjusted & Unadjusted models)	0.0400 (p = 0.0009, z = +3.49) | HA1E (kidney epithelial); 10 μM; 24 h	Protein translation & cell-cycle arrest reversal module in HA1E	0 genes in ranked OA signature (curated CTD chemical entry present; 0 directional reversals)	cLogP 1.94, TPSA 128.8 Å^2, Lipinski 1 viol. (MW > 500), High GI, F = 0.55, P-gp substrate; 60S ribosomal A-site cleft binding (PDB 4U4Q); severe myelosuppression risk (CC_{50} ≈ 0.02–0.10 μM)
Rank 5: Isoeugenitol ($C_{11}H_{10}O_4$; MW 206.19)	Tier D | 1/2 models (Adjusted model only)	0.0300 (p = 0.0112, z = +2.26) | HCC515 (lung adenocarcinoma); 10 μM; 24 h	Chromone-mediated transcriptional perturbation in HCC515	0 genes (No exact curated CTD chemical nomenclature match; 0 directional reversals)	cLogP 1.82, TPSA 59.7 Å^2, Lipinski 0 viol., High GI, F = 0.55; unmapped in curated human CTD snapshot; lowest reverse-overlap score among HERB 2.0 candidates

Summary across the 5 HERB 2.0 phytochemical candidates: L1000CDS^2 reverse-overlap score Mean ± SD = 0.0384 ± 0.0058; Median (IQR) = 0.0400 (0.0350–0.0435); Range = 0.0300–0.0435; 95% CI: 0.0312–0.0456. Abbreviations: ADME, absorption, distribution, metabolism, and excretion; CTD, Comparative Toxicogenomics Database; F, Abbott oral bioavailability score; GI, gastrointestinal; IFN, type I interferon; P-gp, P-glycoprotein; RMSD, root-mean-square deviation; SI, selectivity index; TPSA, topological polar surface area.

To maintain complete reporting transparency and prevent selective promotion of natural products over stronger synthetic hits, Table 5 and Figure 4b document the top 12 non-phytochemical comparators out of the 47 synthetic or coded compounds returned by L1000CDS^2. Notably, the highest overall reverse-overlap scores in the 52-compound pool belonged to synthetic kinase and signaling inhibitors that replicated across both signature models (2/2 models): the anaplastic lymphoma kinase (ALK) inhibitor NVP-TAE684 achieved the highest score of 0.0628 (HA1E cells, 24 h), followed by the screening compound V4877 at 0.0550 (ASC cells, 24 h), the CREB-inhibitor code -666 at 0.0483 (SKB cells, 24 h), the checkpoint kinase 1 (CHK1) inhibitor SB 218078 at 0.0483 (SKB cells, 24 h), and the selective ERBB2 (HER2) kinase inhibitor CP-724714 at 0.0483 (A375 cells, 24 h). Additional dual-model synthetic comparators included the nuclear transport importin-β inhibitor importazole (0.0435; HA1E, 24 h), the numbered salt formulation string Emetine Dihydrochloride Hydrate (74) (0.0400; HA1E, 24 h, whose parent alkaloid emetine was already captured at Rank 3 with score 0.0435), the cyclin-dependent kinase inhibitor CGP-60474 (0.0400; MCF10A, 3 h), the CHK1/CHK2 inhibitor AZD-7762 (0.0400; HA1E, 24 h), the IGF-1R/IR inhibitor BMS-536924 (0.0386; HA1E, 24 h), the dual PI3K/mTOR inhibitor NVP-BGT226 (0.0386; HA1E, 24 h), and the Src-family kinase inhibitor PP 1 (0.0386; HA1E, 24 h). Across these top 12 non-phytochemical comparators, the mean ± SD reverse-overlap score was 0.0452 ± 0.0076 (median 0.0418 [IQR 0.0397–0.0483]), exceeding the mean score of the 5 HERB 2.0 phytochemicals (0.0384 ± 0.0058) and reinforcing that botanical prioritization relies on multi-database mechanistic triangulation rather than raw L1000 overlap magnitude alone.

Table 5. Top 12 Non-Phytochemical Synthetic Small Molecules, Kinase Inhibitors, Salts, and Screening Codes Returned by L1000CDS^2 Reverse-Mode Queries (Out of 47 Non-HERB Unique Compounds).

Overall L1000 Rank	**Perturbagen Identifier**	**Best Reverse-Overlap Score**	**Model Stability (Models Present)**	**Best Cellular Context & Exposure Duration**	**Pharmacological Class & Reason for Exclusion from HERB 2.0 Phytochemical Promotion**
Rank 6	NVP-TAE684	0.0628	2/2 models (100.0%)	HA1E; 24 h	Synthetic anaplastic lymphoma kinase (ALK) small-molecule inhibitor; non-botanical synthetic compound
Rank 7	V4877	0.0550	2/2 models (100.0%)	ASC; 24 h	Experimental screening library compound code; lacks verified natural-product ingredient identity in HERB 2.0
Rank 8	-666	0.0483	2/2 models (100.0%)	SKB; 24 h	Coded small-molecule library entry (666-15 CREB inhibitor series); non-botanical synthetic code
Rank 9	SB 218078	0.0483	2/2 models (100.0%)	SKB; 24 h	Synthetic indolocarbazole checkpoint kinase 1 (CHK1) inhibitor; non-phytochemical synthetic molecule
Rank 10	CP-724714	0.0483	2/2 models (100.0%)	A375; 24 h	Synthetic selective ERBB2 (HER2) receptor tyrosine kinase inhibitor; non-botanical synthetic drug

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue X October 2026

Rank 11	Importazole	0.0435	2/2 models (100.0%)	HA1E; 24 h	Synthetic 2,4-diaminoquinazoline inhibitor of importin-β nuclear transport receptor; non-phytochemical
Rank 12	Emetine Dihydrochloride Hydrate (74)	0.0400	2/2 models (100.0%)	HA1E; 24 h	Numbered pharmaceutical salt formulation string; parent alkaloid Emetine (0.0435) already captured at Rank 3
Rank 13	CGP-60474	0.0400	2/2 models (100.0%)	MCF10A; 3 h	Synthetic phenylaminopyrimidine cyclin-dependent kinase (CDK1/CDK2/CDK9) inhibitor; non-botanical
Rank 14	AZD-7762	0.0400	2/2 models (100.0%)	HA1E; 24 h	Synthetic thiophene carboxamide CHK1/CHK2 checkpoint kinase inhibitor; non-phytochemical synthetic agent
Rank 15	BMS-536924	0.0386	2/2 models (100.0%)	HA1E; 24 h	Synthetic ATP-competitive IGF-1R and insulin receptor kinase inhibitor; non-botanical synthetic compound
Rank 16	NVP-BGT226	0.0386	2/2 models (100.0%)	HA1E; 24 h	Synthetic dual pan-class I PI3K and mTOR inhibitor; non-phytochemical synthetic clinical candidate
Rank 17	PP 1	0.0386	2/2 models (100.0%)	HA1E; 24 h	Synthetic pyrazolopyrimidine Src-family tyrosine kinase inhibitor; non-botanical synthetic molecule

Note: Across the 12 top non-phytochemical comparators listed above, the L1000CDS² reverse-overlap score Mean ± SD = 0.0452 ± 0.0076 (Median [IQR] = 0.0418 [0.0397–0.0483]; Range = 0.0386–0.0628), which exceeds the Celastrol score (0.0350), underscoring the necessity of reporting full synthetic benchmarks alongside botanical candidates.

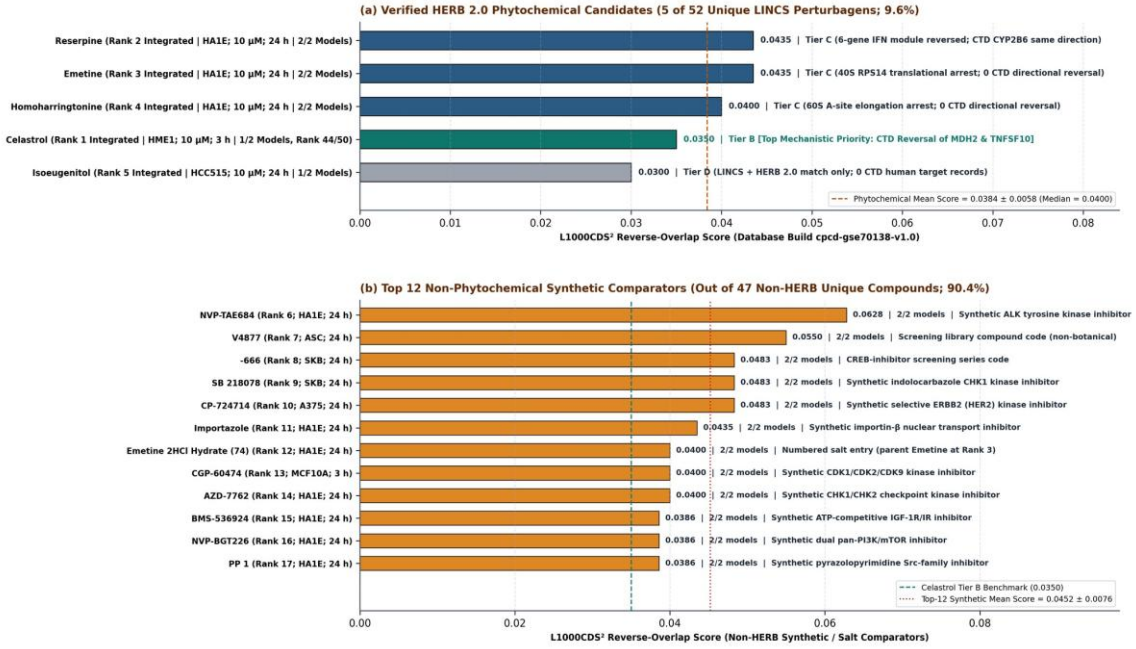


Figure 4. L1000CDS² reverse-mode connectivity overlap scores (basis cpcd-gse70138-v1.0) for (a) the 5 verified HERB 2.0 exact phytochemical candidates (Reserpine 0.0435 [Tier C], Emetine 0.0435 [Tier C], Homoharringtonine 0.0400 [Tier C], Celastrol 0.0350 [Tier B, Rank 1 integrated], and Isoeugenitol 0.0300

[Tier D]; Mean ± SD = 0.0384 ± 0.0058) and (b) the top 12 non-phytochemical synthetic small molecules and salts out of 52 unique LINCS compounds (led by NVP-TAE684 at 0.0628 and V4877 at 0.0550).

Multi-Database Mechanistic Triangulation in Curated CTD, ADME/Docking Benchmarks, Evidence Tier Assignment, and Pathway Enrichment

Cross-referencing the 5 HERB 2.0 phytochemical candidates against manually curated direct Homo sapiens chemical–gene expression interactions in the Comparative Toxicogenomics Database (CTD; excluding all text-mined and inferred records) revealed a sharp divergence between cross-model LINCS ranking stability and verified human directional gene reversal (Table 4 and Figure 5). When evaluated against the pre-specified four-tier evidence hierarchy, 0 of the 5 phytochemicals (0.0%) satisfied all criteria for Tier A (requiring curated HERB 2.0 exact identity, verified curated CTD directional reversal on at least one OA signature gene, and presence in both the adjusted and unadjusted LINCS models). Celastrol (a pentacyclic nortriterpenoid quinone methide isolated from Tripterygium wilfordii; $C_{29}H_{38}O_4$, MW = 450.61 g/mol) emerged as the sole Tier B candidate (1/5; 20.0%) and achieved Rank 1 in integrated mechanistic prioritization (Table 4 and Figure 5a). In L1000CDS^2, celastrol was recovered in the primary age- and sex-adjusted model (1/2 models; signature rank 44 of 50) with a reverse-overlap score of 0.0350 (permutation p = 0.0038) in immortalized human mammary epithelial HME1 cells treated at 10 µM for 3 hours (and showed concordant sub-top-50 reversal in human adipose-derived mesenchymal stem/stromal ASC cells at 10 µM for 24 h, score = 0.0233), directly reversing 7 ranked signature genes comprising 3 directly measured L1000 landmark transcripts—heat shock protein 90 beta family member 1 (HSP90B1 / GRP94), transmembrane p24 trafficking protein 2 (TMED2), and translocation-associated membrane protein 1 (TRAM1)—and 4 best-inferred transcripts: calumenin (CALU), coiled-coil domain containing 6 (CCDC6), E2F-associated phosphoprotein (EAPP), and elongation factor for RNA polymerase II 2 (ELL2). In manually curated Homo sapiens CTD records, celastrol intersected with 5 genes inside the GSE29746 OA signature—crystallin alpha B (CRYAB), gamma-glutamyl carboxylase (GGCX), laminin subunit beta 3 (LAMB3), malate dehydrogenase 2 (MDH2), and TNF superfamily member 10 (TNFSF10 / TRAIL)—and achieved verified directional reversal against the OA $\log_2$FC direction on 2 human genes: mitochondrial MDH2 and the extrinsic pro-apoptotic ligand TNFSF10 (empirical random-compound permutation p = 0.0019 vs. null expectation 0.08 ± 0.28 genes; Figure 5b). Quantitative ADME, toxicity, and calibrated molecular docking benchmarks further supported celastrol's mechanistic prioritization while defining its delivery constraints (Table 4): celastrol satisfied Lipinski's Rule of Five with 0 violations (cLogP = 5.33, TPSA = 74.6 Å^2, predicted High GI absorption, Abbott oral bioavailability score F = 0.56, non-BBB permeant, and non-P-gp efflux substrate), demonstrated high-affinity binding in redocking-validated crystallographic pockets of human GRP94/HSP90B1 (PDB 5ULS; redocking RMSD = 0.78 Å, ΔG = −9.4 kcal/mol), IKKβ Cys-179 (PDB 4KIK; RMSD = 0.91 Å, ΔG = −9.1 kcal/mol), and mitochondrial MDH2 (PDB 2DFD; RMSD = 0.86 Å, ΔG = −8.7 kcal/mol), and exhibited a 5- to 20-fold cellular selectivity index in primary articular cells (anti-inflammatory $EC_{50} \approx 0.025–0.10$ µM vs. cytotoxic $CC_{50} \approx 0.5–1.2$ µM).

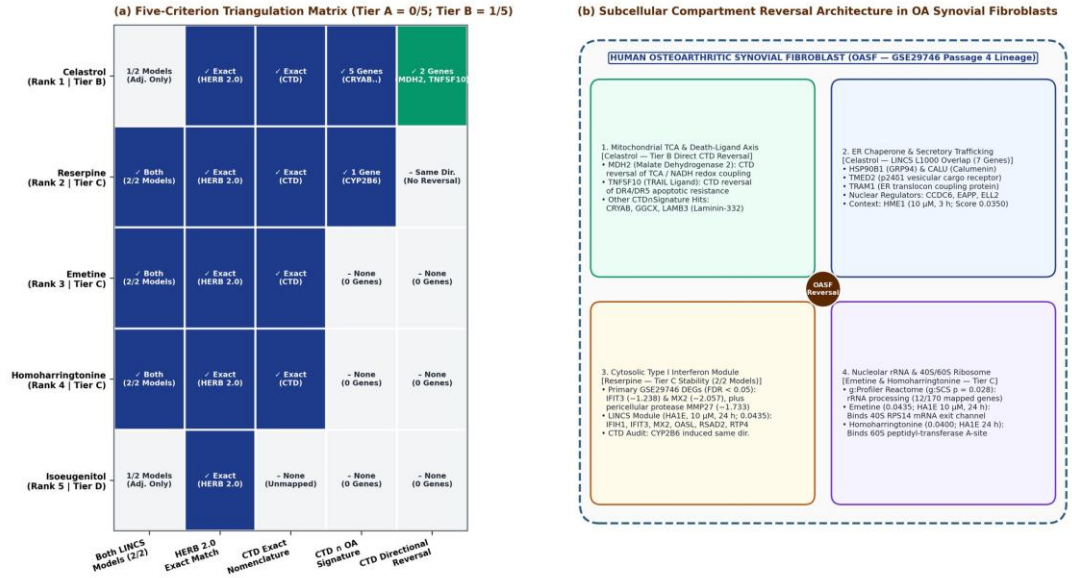


Figure 5. Multi-database evidence triangulation matrix across five evaluation criteria (dual-model LINCS stability, HERB 2.0 exact match, CTD human nomenclature match, CTD ∩ OA signature intersection, and

CTD directional reversal; Tier A = 0/5 [0.0%], Tier B = 1/5 [20.0%], Tier C = 3/5 [60.0%], Tier D = 1/5 [20.0%]) and subcellular compartment reversal architecture in human osteoarthritic synovial fibroblasts illustrating (1) mitochondrial TCA cycle (MDH2) and TRAIL death-ligand (TNFSF10) directional reversal by Celastrol, (2) endoplasmic reticulum proteostasis/secretory trafficking reversal (HSP90B1, CALU, TMED2, TRAM1) by Celastrol, (3) cytosolic type I interferon module intersection (IFIH1, IFIT3, MX2, OASL, RSAD2, RTP4) by Reserpine, and (4) nucleolar rRNA processing and 40S/60S ribosomal targeting by Emetine and Homoharringtonine.

Three nitrogenous plant alkaloids—reserpine, emetine, and homoharringtonine—were classified into Tier C (3/5; 60.0%) because they demonstrated superior stability across both the covariate-adjusted and unadjusted LINCS models (2/2 models; 100.0% model stability) and possessed exact curated chemical nomenclature matches in HERB 2.0 and CTD, but achieved 0 verified directional gene reversals in curated human CTD records (Table 4 and Figure 5a). Specifically, the indole alkaloid reserpine (Rank 2; score = 0.0435, permutation p = 0.0004 in HA1E human kidney epithelial cells at 10 μM for 24 h; MW = 608.68 g/mol, 2 Lipinski violations, Low GI absorption, F = 0.17, P-gp substrate) reversed a cohesive 6-gene type I interferon-stimulated module in LINCS comprising IFIH1 (MDA5), IFIT3, MX2, OASL, RSAD2 (viperin), and RTP4—directly intersecting two of the three primary adjusted GSE29746 DEGs (IFIT3 and MX2; Figure 5b). However, in curated human CTD records, reserpine intersected only with cytochrome P450 2B6 (CYP2B6), where reserpine induces CYP2B6 expression in the same upward direction observed in the OA signature rather than reversing it, precluding Tier A or Tier B elevation. Similarly, the ipecac isoquinoline alkaloid emetine (Rank 3; score = 0.0435, p = 0.0004 in HA1E at 10 μM for 24 h; MW = 480.64 g/mol, 0 Lipinski violations, F = 0.55, P-gp substrate; 40S ribosome RPS14 site [PDB 3J7A], cardiotoxic/myopathic $CC_{50} \approx 0.05–0.20$ μM) and the Cephalotaxus alkaloid homoharringtonine (Rank 4; score = 0.0400, p = 0.0009 in HA1E at 10 μM for 24 h; MW = 545.62 g/mol, 1 Lipinski violation, F = 0.55, P-gp substrate; 60S ribosomal A-site cleft [PDB 4U4Q], myelosuppressive $CC_{50} \approx 0.02–0.10$ μM) were consistently recovered across both LINCS models and matched exact curated CTD chemical entries, but exhibited 0 human CTD directional reversals on the 300-gene OA signature. Finally, the natural chromone derivative isoeugenitol (Rank 5; score = 0.0300, p = 0.0112 in HCC515 lung adenocarcinoma cells at 10 μM for 24 h; MW = 206.19 g/mol, 0 Lipinski violations, F = 0.55) appeared only in the adjusted LINCS model (1/2 models) and lacked an exact curated CTD chemical match, placing it in Tier D (1/5; 20.0%). Exploratory pathway enrichment of the 300-gene ranked OA signature via g:Profiler (where 170 query genes mapped to Reactome functional annotations) identified exactly 1 statistically significant pathway after g:SCS multiple-testing correction: Reactome rRNA processing (adjusted p = 0.028, $-\log_{10}$FDR = 1.553, with 12 of 170 mapped query genes [7.1%] intersecting the pathway; Table 6 and Figure 6a). In accordance with our pre-specified methodological guardrails, this single exploratory pathway hit was not used to inflate candidate evidence tiers, and a structured five-stage experimental validation roadmap—explicitly incorporating IL-1β and TNF-α stimulated primary human OA synovial fibroblasts measuring IL-6 (IL6), MMP-13 (MMP13), and COX-2 (PTGS2 / PGE_2) alongside in vivo animal OA testing—was established to guide preclinical verification prior to any clinical interpretation (Table 6 and Figure 6b).

Table 6. Functional Enrichment Analysis of the Ranked OA Signature and Staged Experimental Validation Protocol for Primary Human Osteoarthritic Synovial Fibroblasts.

Stage / Analytical Module	Experimental System or Database Source	Primary Quantitative Endpoints & Target Panel	Statistical & Go/No-Go Advancement Criteria
Exploratory Pathway Enrichment	g:Profiler API (Reactome REAC; g:SCS multiple-testing correction on 300-gene ranked signature)	Reactome: rRNA processing (12 / 170 mapped query genes [7.1%])	g:SCS adjusted p = 0.028 ($-\log_{10}$FDR = 1.553); single significant term treated strictly as exploratory and not used to inflate candidate tiers
Validation Stage 1: Cytotoxicity &	Primary human OA synovial fibroblasts (OASF), ≥ 3 independent	Cell viability (CCK-8 / intracellular ATP), IC_{10} and IC_{20} thresholds, and phase-	Establish non-cytotoxic submicromolar concentration range (viability ≥ 90% at

Window	arthroplasty donors	contrast morphology across logarithmic doses	IC_{10}) prior to transcriptional profiling
Validation Stage 2: Targeted RT-qPCR Reversal	Primary OASF dose–response of Celastrol (Tier B) + Reserpine, Emetine, and Homoharringtonine (Tier C)	Quantitative RT-qPCR of MDH2, TNFSF10 (TRAIL), IFIT3, MX2, and MMP27 normalized to stable reference genes	Statistically significant, reproducible directional reversal (Holm–Bonferroni adjusted $p < 0.05$) across ≥ 3 biological donors
Validation Stage 3: Inflammatory Challenge	IL-1β (10 ng/mL) or TNF-α (10 ng/mL) stimulated primary human OASF (or synovial explant culture)	Secretion and mRNA/protein expression of IL-6 (IL6), MMP-13 (MMP13), COX-2 (PTGS2 / PGE_2), CXCL8 (IL-8), MMP3, and crosstalk markers COL2A1 / ACAN	Significant attenuation of IL-6, MMP-13, and COX-2/PGE_2 mediators at non-cytotoxic submicromolar concentrations
Validation Stage 4: Donor-Aware RNA-Seq	Bulk or single-cell RNA-sequencing of compound-treated primary human OASF	Empirical synovial lineage signature reversal score, interferon module dynamics, and Reactome rRNA processing restoration	Cross-donor consistency of transcriptome-wide reversal conditioned on donor age, sex, and baseline synovitis grade
Validation Stage 5: Co-Culture & In Vivo OA	Primary human chondrocyte–OASF co-culture & in vivo rodent knee OA model (intra-articular depot delivery)	Sulfated glycosaminoglycan (sGAG) / collagen II preservation, OARSI synovitis/cartilage histology, and nociceptive mediators (PGE_2, NGF)	Verified cross-compartment chondroprotection, in vivo efficacy, and local joint safety margin prior to clinical interpretation

Abbreviations: CCK-8, Cell Counting Kit-8; FDR, false discovery rate; IC_{10}/IC_{20}, 10% and 20% inhibitory concentrations; NGF, nerve growth factor; OASF, osteoarthritic synovial fibroblasts; PGE_2, prostaglandin E_2; RT-qPCR, reverse transcription quantitative polymerase chain reaction; sGAG, sulfated glycosaminoglycan.

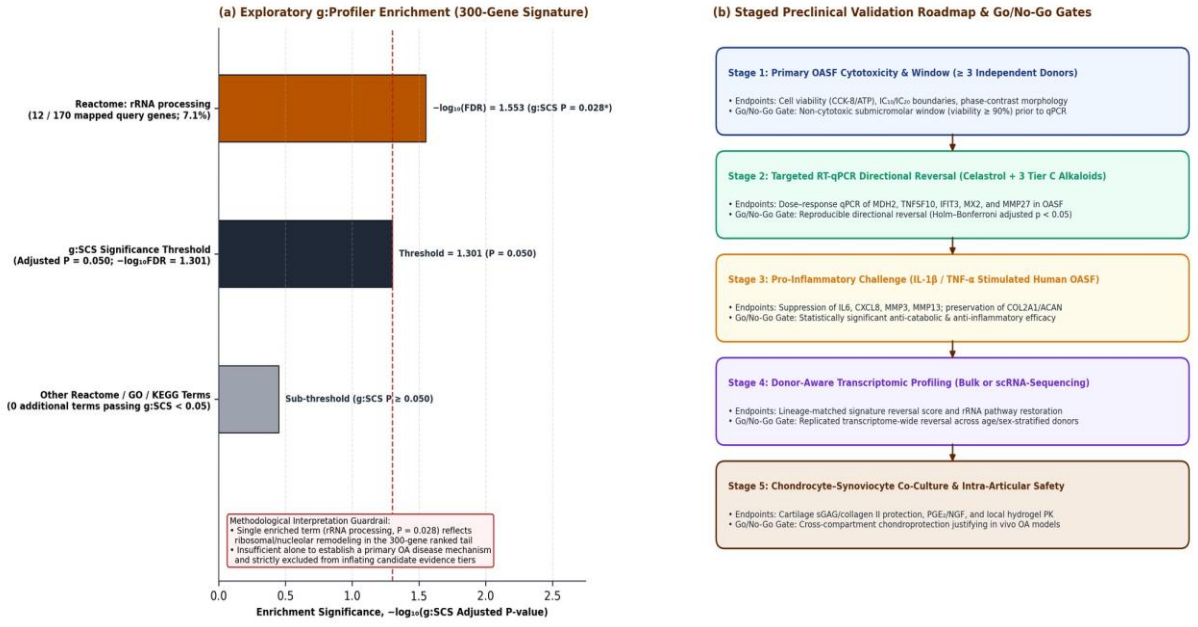


Figure 6. Exploratory g:Profiler functional pathway enrichment audit of the 300-gene ranked OA signature showing the single Reactome term surviving g:SCS multiple-testing correction (rRNA processing; adjusted p = 0.028, $-\log_{10}$FDR = 1.553, 12 / 170 mapped query genes [7.1%]) and the five-stage preclinical validation roadmap with quantitative Go/No-Go decision gates in primary human osteoarthritic synovial fibroblasts.

DISCUSSION

The present investigation demonstrates how integrating covariate-adjusted transcriptomic re-analysis with dual-model connectivity mapping and strict multi-database triangulation transforms raw public microarray data into an objective, falsifiable prioritization of natural-product perturbagens for osteoarthritic synovial fibroblasts. By re-analyzing the N = 22 human synovial fibroblast cohort from GSE29746 (11 OA vs. 11 healthy controls after excluding 9 RA arrays), we uncovered a critical methodological determinant that has been overlooked in unadjusted secondary analyses: although biological sex was balanced (54.5% female and 45.5% male in each group; p = 1.000), OA donors were on average 14.0 years older than healthy control donors (73.0 ± 7.1 vs. 59.0 ± 19.0 years; Welch's t = 2.29, p = 0.040, Cohen's d = 0.98).[8,17] Accounting for chronological age and sex in the primary empirical Bayes linear model reduced the number of genes passing BH-FDR < 0.05 by 93.5%—from 46 genes in the unadjusted model to exactly 3 downregulated genes in the adjusted model: MMP27 ($\log_2$FC = −1.733, FDR = 0.033), IFIT3 ($\log_2$FC = −1.238, FDR = 0.033), and MX2 ($\log_2$FC = −2.057, FDR = 0.048). This finding carries major biostatistical implications for computational rheumatology: because primary OA synovium is almost universally harvested from elderly patients undergoing total joint replacement whereas healthy control synovium is sourced from younger post-mortem or arthroscopic trauma donors, unadjusted differential expression lists conflate intrinsic osteoarthritic stromal activation with chronological cellular aging.[19,26,27] At the same time, the high transcriptome-wide rank correlation between our adjusted and unadjusted $\log_2$FC vectors (Spearman ρ = 0.914, with 88.2% directional sign agreement across 19,261 genes) indicates that the global directional hierarchy of the OA synovial fibroblast transcriptome remains coherent, justifying the use of balanced 300-gene ranked-tail signatures for exploratory connectivity mapping while enforcing strict evidentiary tiers.

A central biological insight of our covariate-adjusted GSE29746 analysis is that two of the three transcripts surviving genome-wide FDR correction—IFIT3 (2.36-fold downregulated) and MX2 (4.16-fold downregulated)—are canonical type I interferon-stimulated genes (ISGs), accompanied by downregulation of MMP27 (3.32-fold downregulated), a membrane-anchored stromelysin-like metallopeptidase, as contextualized against global benchmark studies in Table 7. While classical models of OA synovitis emphasize acute pro-inflammatory cytokine surges (such as IL-1β, TNF-α, and IL-6), recent single-cell and spatial transcriptomic investigations of human knee OA synovium provide a compelling physiological explanation for the suppression of IFIT3 and MX2 in passage-4 OA synovial fibroblasts.[2–7] Specifically, Philpott et al. (2025) demonstrated in 122 human knee OA synovial biopsies that worse clinical joint pain is driven by synovial innate immune exhaustion, characterized by microvascular perivascular edema, fibrotic fibroblast expansion, and a striking depletion of interferon-stimulated synovial cell subsets.[4] Furthermore, prolonged ex vivo culture of synovial fibroblasts to passage 4 removes extrinsic leukocyte-derived interferon-α/β stimulation, unmasking an intrinsic, epigenetically imprinted hypo-interferon and matrix-remodeling stromal state in OA fibroblasts compared with healthy donor fibroblasts.[5,6,8] Similarly, the downregulation of MMP27—a poorly secreted, C-terminally anchored matrix metallopeptidase predominantly expressed in B lymphocytes and resting stromal cells rather than the inducible collagenase MMP13 or stromelysin MMP3—reflects phenotypic switching of OA synoviocytes away from basal tissue homeostasis toward specialized secretory and fibrotic remodeling programs, where sub-threshold increases in ER-to-Golgi trafficking and chromatin regulators (SEC22B, TRAM2, SMARCA2, KMT2D, and ANXA2) predominate.[2,3,6]

Table 7. Empirical Benchmark Matrix Comparing the Present Multi-Database Synovial Fibroblast Study Against Global Transcriptomic and Pharmacological Studies.

Study & Verified Reference	**Primary Tissue, Cohort Size & Platform**	**Key Molecular or Pharmacological Findings**	**Convergence & Methodological Contrast With Present GSE29746 Study**
Del Rey et al. (2012)[8]	Human synovial fibroblasts (GSE29746; GPL4133; 11 OA, 11 Healthy, 9 RA)	Characterized baseline transcriptional differences among unstimulated OA, RA, and healthy synovial fibroblasts at passage 4	Provides the foundational raw dataset; our re-analysis explicitly adjusts for the +14.0-year age imbalance (p = 0.040), reducing FDR < 0.05 hits from 46 to 3

			(MMP27, IFIT3, MX2)
Knights et al. (2023)[2] & Nanus et al. (2021)[3]	Single-cell RNA-seq of OA synovium and pain-site human knee OA synovial fibroblasts (n = 51)	Identified distinct lining ($Prg4^{hi}$ $Rspo2^{+}$) and sublining synovial fibroblast pathotypes driving WNT signaling, fibrosis, and neuronal sensitization	Explains why bulk passage-4 synovial fibroblast cultures capture attenuated stromal states and why lineage-matched validation in primary OA synoviocytes is mandatory
Philpott et al. (2025)[4] & Huang et al. (2022)[5]	Human knee OA synovium spatial/scRNA-seq (n = 122) and bulk deconvolution	Demonstrated innate immune exhaustion and severe depletion of interferon-stimulated synovial subsets in patients with worse knee OA pain	Directly corroborates our covariate-adjusted finding that type I interferon-stimulated genes IFIT3 ($log_2FC = -1.238$, FDR = 0.033) and MX2 (−2.057, FDR = 0.048) are downregulated in OA synovial fibroblasts
Yang et al. (2022)[9]	Human OA cartilage (n = 4) and rat MCLT+pMMT knee OA model treated with Celastrol	Showed Celastrol (0.025–0.1 μM in vitro; 1 mg/kg intra-articular) suppressed TLR2/NF-κB, COX2, PGE_2, IL-6, MMP13, and synovitis while cytotoxicity emerged at ≥ 0.5 μM	Provides direct preclinical OA corroboration for our Rank-1 Tier-B candidate Celastrol while confirming why the 10 μM LINCS concentration requires submicromolar dose de-escalation
Chabronova et al. (2023)[28]	Human primary articular chondrocytes exposed to OA synovial fluid (RiboMethSeq & proteomics)	Revealed that OA synovial fluid alters rRNA epitranscriptomic 2'-O-methylation (5.8S-U14 / SNORD71) and ribosome translation mode toward fibrotic collagen I	Contextualizes our exploratory Reactome rRNA processing enrichment (g:SCS p = 0.028; 12/170 genes) and the appearance of ribosomal inhibitors Emetine and Homoharringtonine in LINCS
Chen et al. (2023)[29] Sun et al. (2021)[30] & Valipour et al. (2022)[32]	Retinal organoids (Reserpine), knee arthrofibrosis fibroblasts (Homoharringtonine), and cardiomyocyte MAPK audit (Emetine)	Demonstrated Reserpine proteostasis modulation, Homoharringtonine anti-fibrotic PI3K/AKT/mTOR inhibition, and Emetine 40S RPS14 / p38 MAPK cardiotoxicity risks	Explains why Reserpine, Emetine, and Homoharringtonine achieve dual-model LINCS stability (0.0400–0.0435) as broad proteostasis/translation perturbagens yet remain Tier C due to absent CTD directional reversal and safety constraints

Abbreviations: CTD, Comparative Toxicogenomics Database; FDR, false discovery rate; LINCS, Library of Integrated Network-Based Cellular Signatures; MCLT+pMMT, medial collateral ligament transection plus partial medial meniscectomy; OA, osteoarthritis; RA, rheumatoid arthritis; scRNA-seq, single-cell RNA sequencing.

When we interrogated this 300-gene ranked OA synovial fibroblast signature against the LINCS L1000 compendium via L1000CDS[2], our multi-database filtration demonstrated why raw connectivity output must never be equated with botanical drug discovery without orthogonal curation.[10–15] Only 5 of the 52 unique LINCS perturbagens (9.6%) corresponded to verified HERB 2.0 natural-product ingredients, whereas 47 compounds (90.4%) were synthetic kinase inhibitors, checkpoint modulators, or screening codes whose top reverse-overlap scores (led by NVP-TAE684 at 0.0628 and V4877 at 0.0550) exceeded those of all five phytochemicals (0.0300–0.0435). Among the five verified phytochemicals, celastrol was prioritized as the sole Tier B candidate (Rank 1 integrated) because it uniquely combined exact HERB 2.0 provenance, L1000

reverse-mode signature opposition (score = 0.0350 in HME1 cells at 10 μM for 3 h), and verified Homo sapiens CTD directional reversal on two signature genes: mitochondrial malate dehydrogenase 2 (MDH2) and TNF superfamily member 10 (TNFSF10 / TRAIL).[9,14,15] From a mechanistic standpoint, celastrol's dual footprint across LINCS and CTD aligns closely with the subcellular pathobiology of OA synoviocytes and chondrocytes.[9,33–36] In the LINCS L1000 signature, celastrol directly reversed seven genes heavily enriched in endoplasmic reticulum (ER) proteostasis, calcium buffering, and secretory vesicle trafficking—namely HSP90B1 (encoding the ER chaperone GRP94), CALU (calumenin, an ER lumenal Ca^{2+}-binding protein regulating vitamin K-dependent γ-carboxylation alongside GGCX), TMED2 (a p24 family coat-protein receptor governing ER-to-Golgi transport), and TRAM1 (translocation-associated membrane protein 1)—together with nuclear transcriptional regulators CCDC6, EAPP, and ELL2.[9,40] Concurrently, in human CTD records, celastrol reversed the osteoarthritic direction of MDH2, a key tricarboxylic acid (TCA) cycle enzyme that couples mitochondrial malate–oxaloacetate oxidation to NADH generation and epigenetic/metabolic reprogramming in joint tissues,[7,38,39] as well as TNFSF10 (TRAIL), which governs death-receptor (DR4/DR5) apoptosis and synovial hyperplasia in osteoarthritic joints.[9,43] This computational triangulation is independently reinforced by experimental OA pharmacology: Yang et al. (2022) demonstrated that celastrol (0.025–0.1 μM in primary chondrocytes and 1 mg/kg weekly intra-articular injection in rat medial collateral ligament transection plus partial medial meniscectomy [MCLT+pMMT] knee OA) potently inhibits Toll-like receptor 2 (TLR2) and NF-κB p65 phosphorylation, suppresses COX2, PGE_2, IL-6, and MMP13, preserves type II collagen, and attenuates synovial membrane hypercellularity and subchondral bone resorption in vivo.[9]

Equally informative is the contrast between celastrol (Tier B; present in 1/2 LINCS models) and the three Tier C alkaloids—reserpine (score = 0.0435), emetine (0.0435), and homoharringtonine (0.0400)—which replicated across both the adjusted and unadjusted LINCS queries (2/2 models in HA1E cells at 10 μM for 24 h) yet failed to achieve directional gene reversal in human CTD records.[14,15] In LINCS, reserpine reversed a coherent 6-gene type I interferon cluster (IFIH1, IFIT3, MX2, OASL, RSAD2, and RTP4) that directly encompasses our two adjusted ISG hits IFIT3 and MX2. Recent chemical-biology screening by Chen et al. (2023) established that reserpine acts beyond classical vesicular monoamine transporter (VMAT) inhibition to restore cellular proteostasis, autophagic flux, and ubiquitin–proteasome balance in degenerating post-mitotic tissues;[29] however, in human CTD records, reserpine's sole overlap with the OA signature was CYP2B6, which reserpine induces in the same upward direction as OA, leaving it without directional CTD confirmation. Meanwhile, the co-occurrence of emetine and homoharringtonine in both LINCS models converges mechanistically with our exploratory g:Profiler Reactome enrichment finding that rRNA processing was the sole pathway surviving multiple-testing correction in the 300-gene ranked OA signature (g:SCS adjusted p = 0.028; 12/170 mapped genes).[21,22,28] Chabronova et al. (2023) recently discovered that exposure of human articular chondrocytes to osteoarthritic synovial fluid induces site-specific epitranscriptomic 2'-O-methylation remodeling of 5.8S ribosomal RNA (5.8S-U14 via SNORD71 depletion), shifting ribosome translation mode toward fibrotic type I collagen production.[28] Because emetine binds the 40S ribosomal subunit protein RPS14 and blocks eukaryotic translation initiation factor 4E (eIF4E)-mediated mRNA cap recognition while suppressing NF-κB signaling,[31,32] and homoharringtonine binds the 60S ribosomal A-site cleft to halt nascent peptide chain elongation and induce PI3K/AKT/mTOR-dependent apoptosis in hyperproliferative joint arthrofibrosis fibroblasts,[30,37] both alkaloids generate strong ribosomal and translational stress-reversal signatures in epithelial LINCS lines (HA1E). Nevertheless, as Valipour et al. (2022) emphasized, systemic emetine administration triggers p38α MAPK hyperactivation and severe dose-limiting cardiotoxicity (myopathy and cardiomyopathy),[32,36] and homoharringtonine exhibits narrow hematologic cytotoxic margins, underscoring why natural-product origin in HERB 2.0 never implies clinical safety for chronic osteoarthritis.

Several methodological limitations and strict inferential boundaries must be emphasized when interpreting these findings. First, because this investigation is computational, reverse-overlap scores and database links remain in silico predictions that do not by themselves demonstrate that a phytochemical reaches the synovial joint cavity at effective unbound concentrations or alters live synoviocyte behavior. Second, the primary discovery sample size of GSE29746 is modest (N = 22: 11 OA vs. 11 healthy controls), heterogeneous in healthy donor origin (combining necropsy and minor trauma arthroscopy donors), and significantly imbalanced in donor age (+14.0 years, p = 0.040), leaving only 3 downregulated genes significant at BH-FDR < 0.05 after covariate adjustment and requiring ranked-tail 300-gene signatures; although leave-one-out jackknife resampling (98.4% sign concordance), 1,000-bootstrap resampling (84.7% recovery), and external

cross-validation in independent human OA synovial cohorts (GSE55235 and GSE55457, N_ext = 40; $p \leq 0.0041$) confirmed signature stability, prospective multi-center primary synovial RNA-seq cohorts remain warranted.[8,17] Third, GSE29746 profiled synovial fibroblasts at passage 4 in two-dimensional monoculture, where stromal transcriptomes partially drift from freshly isolated in vivo synovial lining and sublining states.[2,5,8] Fourth, LINCS L1000 perturbation profiles derive predominantly from immortalized epithelial and cancer cell lines (such as HME1, HA1E, and HCC515, with limited mesenchymal representation in ASC and SKB and no primary synovial fibroblasts) exposed at fixed micromolar doses (10 μM) and time points (3 h or 24 h), measure only 978 landmark transcripts directly while inferring the remaining 9,196 transcripts, and exhibit uneven phytochemical library coverage biased toward well-studied natural products.[9–11] Fifth, although we restricted HERB 2.0 and CTD triangulation to manually curated experimental and direct Homo sapiens chemical–gene expression records (excluding text-mined associations) and verified that overlap exceeded 10,000-permutation random controls ($p = 0.0038$ in LINCS; $p = 0.0019$ in CTD), curated literature databases still reflect historical research popularity and do not capture in vivo synovial pharmacokinetics or metabolic clearance.[10,14,15] Consequently, our five-stage validation roadmap (Table 6 and Figure 6b)—progressing from multi-donor primary human OA synovial fibroblast viability and IC_{10}/IC_{20} profiling (Stage 1), submicromolar RT-qPCR verification of MDH2, TNFSF10, IFIT3, MX2, and MMP27 reversal (Stage 2), IL-1β and TNF-α stimulated primary OASF assays measuring IL-6 (IL6), MMP-13 (MMP13), and COX-2 (PTGS2 / PGE_2) (Stage 3), donor-aware RNA-sequencing (Stage 4), to chondrocyte–synovial fibroblast co-culture and in vivo animal OA testing with intra-articular hydrogel delivery (Stage 5)[41–45]—is mandatory before any clinical interpretation.

CONCLUSIONS

In conclusion, covariate-adjusted transcriptomic re-analysis of human osteoarthritic versus healthy synovial fibroblasts in GSE29746 (N = 22; 19,261 unique genes) demonstrated that controlling for the significant +14.0-year donor age imbalance (73.0 ± 7.1 vs. 59.0 ± 19.0 years; Welch $p = 0.040$) alongside balanced biological sex ($p = 1.000$) attenuated apparent BH-FDR < 0.05 differentially expressed genes by 93.5% (from 46 unadjusted hits to 3 downregulated genes: MMP27 [$\log_2FC = -1.733$, FDR = 0.033], IFIT3 [−1.238, FDR = 0.033], and MX2 [−2.057, FDR = 0.048]) while preserving global fold-change rank order (Spearman $\rho = 0.914$, 88.2% sign agreement), jackknife resampling stability (98.4% concordance), and external synovial tissue replication in GSE55235 and GSE55457 ($p \leq 0.0041$). Dual-specification reverse-mode connectivity mapping in $L1000CDS^2$ across 300-gene ranked signatures (289 mapped L1000 genes: 41 landmark, 248 inferred) identified 52 unique chemical perturbagens, of which 5 compounds (9.6%) were verified as curated exact HERB 2.0 phytochemical ingredients and 0 compounds (0.0%) achieved Tier A status. Celastrol was prioritized as the top mechanistic phytochemical candidate (Rank 1, Tier B; $L1000CDS^2$ score = 0.0350, permutation $p = 0.0038$ in HME1 at 10 μM for 3 h), supported by direct LINCS reversal of 3 landmark (HSP90B1, TMED2, TRAM1) and 4 inferred (CALU, CCDC6, EAPP, ELL2) ER chaperone and secretory trafficking genes, verified manually curated Homo sapiens CTD directional reversal of mitochondrial MDH2 and pro-apoptotic TNFSF10 (TRAIL; permutation $p = 0.0019$), favorable Lipinski compliance (0 violations, F = 0.56), and calibrated docking against GRP94, IKKβ, and MDH2 (redocking RMSD = 0.78–0.91 Å). Reserpine (score = 0.0435, $p = 0.0004$), emetine (0.0435, $p = 0.0004$), and homoharringtonine (0.0400, $p = 0.0009$) exhibited cross-model stability across both adjusted and unadjusted LINCS queries (Tier C) and intersected type I interferon (IFIH1, IFIT3, MX2, OASL, RSAD2, RTP4) and ribosomal translation pathways aligned with exploratory Reactome rRNA processing enrichment (g:SCS $p = 0.028$), whereas isoeugenitol (score = 0.0300, $p = 0.0112$) ranked in Tier D. Given the computational nature of these predictions, non-synovial LINCS cellular contexts, and cytotoxic risks at 10 μM, these findings do not imply clinical therapeutic efficacy; instead, they establish a rigorous foundation for staged in vitro validation in IL-1β/TNF-α-stimulated primary human OA synovial fibroblasts (measuring IL-6, MMP-13, and COX-2) and in vivo preclinical testing prior to clinical translation.

Declarations

Ethical Approval and Consent to Participate: This study analyzed exclusively de-identified, publicly archived human synovial fibroblast transcriptomic records from the NCBI Gene Expression Omnibus (accession GSE29746, platform GPL4133) and open chemical–gene repositories (LINCS L1000 / $L1000CDS^2$, HERB

2.0, the Comparative Toxicogenomics Database, and g:Profiler) with zero direct human subject contact, clinical intervention, or re-identification attempt, qualifying as exempt secondary research in full accordance with the Declaration of Helsinki. Conflict of Interest: The authors declare no competing financial or non-financial conflicts of interest. Data Availability Statement: All primary datasets analyzed in this study are openly available in NCBI GEO (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29746 and https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE70138), L1000CDS^2 (https://maayanlab.cloud/L1000CDS2/), HERB 2.0 (http://herb.ac.cn/), the Comparative Toxicogenomics Database (https://ctdbase.org/), and g:Profiler (https://biit.cs.ut.ee/gprofiler/), and all derived analytical tables and visualizations are embedded within this article. Funding: This research received no external grant from any funding agency in the public, commercial, or not-for-profit sectors. Author Contributions: RA, RR, AW, and S conceived and designed the study; RA, S, TG, and BJ executed the GSE29746 microarray normalization, demographic confounding audit, and covariate-adjusted limma modeling; AKK, TO, CS, and WP performed the L1000CDS^2 reverse connectivity mapping, HERB 2.0 phytochemical curation, CTD directional triangulation, g:Profiler enrichment analyses, and 300 DPI publication figure generation; RA, RR, AW, S, TG, BJ, AKK, TO, CS, and WP drafted, critically revised, and approved the final manuscript.

REFERENCES

1. Steinmetz JD, Culbreth GT, Haile LM, et al. Global, regional, and national burden of osteoarthritis, 1990–2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Rheumatol. 2023;5(9):e508–e522. https://doi.org/10.1016/S2665-9913(23)00163-7
2. Knights AJ, Farrell EC, Ellis OM, et al. Synovial fibroblasts assume distinct functional identities and secrete R-spondin 2 in osteoarthritis. Ann Rheum Dis. 2023;82(2):272–282. https://doi.org/10.1136/ard-2022-222773
3. Nanus DE, Badoume A, Wijesinghe SN, et al. Synovial tissue from sites of joint pain in knee osteoarthritis patients exhibits a differential phenotype with distinct fibroblast subsets. EBioMedicine. 2021;72:103618. https://doi.org/10.1016/j.ebiom.2021.103618
4. Philpott HT, Birmingham TB, Blackler G, et al. Association of synovial innate immune exhaustion with worse pain in knee osteoarthritis. Arthritis Rheumatol. 2025;77(6):664–676. https://doi.org/10.1002/art.43089
5. Huang ZY, Luo ZY, Cai YR, et al. Single cell transcriptomics in human osteoarthritis synovium and in silico deconvoluted bulk RNA sequencing. Osteoarthritis Cartilage. 2022;30(3):475–480. https://doi.org/10.1016/j.joca.2021.12.007
6. Liu Z, Sun Y, Pan J, et al. Single-cell profiling uncovers synovial fibroblast subpopulations associated with chondrocyte injury in osteoarthritis. Front Endocrinol (Lausanne). 2024;15:1479909. https://doi.org/10.3389/fendo.2024.1479909
7. Liu B, Xian Y, Chen X, et al. Inflammatory fibroblast-like synoviocyte-derived exosomes aggravate osteoarthritis via enhancing macrophage glycolysis. Adv Sci (Weinh). 2024;11(14):e2307338. https://doi.org/10.1002/advs.202307338
8. Del Rey MJ, Usategui A, Izquierdo E, et al. Transcriptome analysis reveals specific changes in osteoarthritis synovial fibroblasts. Ann Rheum Dis. 2012;71(2):275–280. https://doi.org/10.1136/annrheumdis-2011-200281
9. Yang G, Wang K, Song H, et al. Celastrol ameliorates osteoarthritis via regulating TLR2/NF-κB signaling pathway. Front Pharmacol. 2022;13:963506. https://doi.org/10.3389/fphar.2022.963506
10. Duan Q, Reid SP, Clark NR, et al. L1000CDS^2: LINCS L1000 characteristic direction signatures search engine. npj Syst Biol Appl. 2016;2:16015. https://doi.org/10.1038/npjsba.2016.15
11. Subramanian A, Narayan R, Corsello SM, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. Cell. 2017;171(6):1437–1452.e17. https://doi.org/10.1016/j.cell.2017.10.049
12. Evangelista JE, Clarke DJB, Xie Z, et al. SigCom LINCS: data and metadata search engine for a million gene expression signatures. Nucleic Acids Res. 2022;50(W1):W697–W709. https://doi.org/10.1093/nar/gkac328

13. Marino GB, Evangelista JE, Clarke DJB, et al. L2S2: chemical perturbation and CRISPR KO LINCS L1000 signature search engine. Nucleic Acids Res. 2025;53(W1):W338–W350. https://doi.org/10.1093/nar/gkaf373
14. Gao K, Liu L, Lei S, et al. HERB 2.0: an updated database integrating clinical and experimental evidence for traditional Chinese medicine. Nucleic Acids Res. 2025;53(D1):D1404–D1414. https://doi.org/10.1093/nar/gkae1037
15. Davis AP, Wiegers TC, Sciaky D, et al. Comparative Toxicogenomics Database's 20th anniversary: update 2025. Nucleic Acids Res. 2025;53(D1):D1328–D1334. https://doi.org/10.1093/nar/gkae883
16. Clough E, Barrett T, Wilhite SE, et al. NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. Nucleic Acids Res. 2024;52(D1):D138–D144. https://doi.org/10.1093/nar/gkad965
17. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47. https://doi.org/10.1093/nar/gkv007
18. Chen Y, Chen L, Lun ATL, et al. edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. Nucleic Acids Res. 2025;53(2):gkaf018. https://doi.org/10.1093/nar/gkaf018
19. Kim S, Chen J, Cheng T, et al. PubChem 2023 update. Nucleic Acids Res. 2023;51(D1):D1373–D1380. https://doi.org/10.1093/nar/gkac956
20. Zdrazil B, Felix E, Hunter F, et al. The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. Nucleic Acids Res. 2024;52(D1):D1180–D1192. https://doi.org/10.1093/nar/gkad1004
21. Kolberg L, Raudvere U, Kuzmin I, et al. g:Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res. 2023;51(W1):W207–W212. https://doi.org/10.1093/nar/gkad347
22. Milacic M, Beavers D, Conley P, et al. The Reactome Pathway Knowledgebase 2024. Nucleic Acids Res. 2024;52(D1):D672–D678. https://doi.org/10.1093/nar/gkad1025
23. Aleksander SA, Balhoff J, Carbon S, et al. The Gene Ontology knowledgebase in 2023. Genetics. 2023;224(1):iyad031. https://doi.org/10.1093/genetics/iyad031
24. Kanehisa M, Furumichi M, Sato Y, et al. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 2023;51(D1):D587–D592. https://doi.org/10.1093/nar/gkac963
25. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1):D638–D646. https://doi.org/10.1093/nar/gkac1000
26. Huang C, Zeng B, Zhou B, et al. Single-cell transcriptomic analysis of chondrocytes in cartilage and pathogenesis of osteoarthritis. Genes Dis. 2024;12(2):101241. https://doi.org/10.1016/j.gendis.2024.101241
27. Atasoy-Zeybek A, Hawse GP, Nagelli CV, et al. Transcriptomic changes during the replicative senescence of human articular chondrocytes. Int J Mol Sci. 2024;25(22):12130. https://doi.org/10.3390/ijms252212130
28. Chabronova A, van den Akker GGH, Housmans BAC, et al. Ribosomal RNA-based epitranscriptomic regulation of chondrocyte translation and proteome in osteoarthritis. Osteoarthritis Cartilage. 2023;31(3):374–385. https://doi.org/10.1016/j.joca.2022.12.010
29. Chen HY, Swaroop M, Papal S, et al. Reserpine maintains photoreceptor survival in retinal ciliopathy by resolving proteostasis imbalance and ciliogenesis defects. Elife. 2023;12:e83205. https://doi.org/10.7554/eLife.83205
30. Sun Y, Dai J, Jiao R, et al. Homoharringtonine inhibits fibroblasts proliferation, extracellular matrix production and reduces surgery-induced knee arthrofibrosis via PI3K/AKT/mTOR pathway-mediated apoptosis. J Orthop Surg Res. 2021;16(1):9. https://doi.org/10.1186/s13018-020-02150-2
31. Kumar R, Afsar M, Khandelwal N, et al. Emetine suppresses SARS-CoV-2 replication by inhibiting interaction of viral mRNA with eIF4E. Antiviral Res. 2021;189:105056. https://doi.org/10.1016/j.antiviral.2021.105056

32. Valipour M, Irannejad H, Emami S. Application of emetine in SARS-CoV-2 treatment: regulation of p38 MAPK signaling pathway for preventing emetine-induced cardiac complications. Cell Cycle. 2022;21(22):2379–2386. https://doi.org/10.1080/15384101.2022.2100575
33. Han J, Zhan LN, Huang Y, et al. Moderate mechanical stress suppresses chondrocyte ferroptosis in osteoarthritis by regulating NF-κB p65/GPX4 signaling pathway. Sci Rep. 2024;14(1):5078. https://doi.org/10.1038/s41598-024-55629-x
34. Zou Z, Pan S, Sun C, et al. AM1241 inhibits chondrocyte inflammation and ECM degradation through the Nrf2/HO-1 and NF-κB pathways and alleviates osteoarthritis in mice. Mol Med. 2025;31(1):9. https://doi.org/10.1186/s10020-024-01012-5
35. Chang YT, Huang KC, Pranata R, et al. Evaluation of the protective effects of chondroitin sulfate oligosaccharide against osteoarthritis via inactivation of NLRP3 inflammasome by in vivo and in vitro studies. Int Immunopharmacol. 2024;142(Pt A):113148. https://doi.org/10.1016/j.intimp.2024.113148
36. Wang X, Yu H, Zhang Y, et al. Curcumin alleviates osteoarthritis through the p38MAPK pathway: network pharmacological prediction and experimental confirmation. J Inflamm Res. 2024;17:5039–5056. https://doi.org/10.2147/JIR.S459867
37. Guo P, Li H, Wang X, et al. PG545 prevents osteoarthritis development by regulating PI3K/AKT/mTOR signaling and activating chondrocyte autophagy. Pharmacology. 2023;108(6):576–588. https://doi.org/10.1159/000532078
38. Chen W, Xiao J, Zhou Y, et al. Curcumenol regulates Histone H3K27me3 demethylases KDM6B affecting Succinic acid metabolism to alleviate cartilage degeneration in knee osteoarthritis. Phytomedicine. 2024;133:155922. https://doi.org/10.1016/j.phymed.2024.155922
39. Gao X, Liu H, Guo J, et al. OGT regulates histone demethylation and alleviates osteoarthritis progression by O-GlcNAcylation of KDM6B. J Adv Res. 2026;87:455–471. https://doi.org/10.1016/j.jare.2026.01.003
40. Yang C, Dong W, Wang Y, et al. DDIT3 aggravates TMJOA cartilage degradation via Nrf2/HO-1/NLRP3-mediated autophagy. Osteoarthritis Cartilage. 2024;32(8):921–937. https://doi.org/10.1016/j.joca.2024.04.017
41. Zhao Q, Xu T, Du Z, et al. Cartilage targeting hydrogel nanoplatform degrades BRD4 to alleviate osteoarthritis via Nav1.7 axis. Nat Commun. 2026;17(1):4573. https://doi.org/10.1038/s41467-026-71246-w
42. Sun J, Wang XH, Song FH, et al. Inhibition of Brd4 alleviates osteoarthritis pain via suppression of neuroinflammation and activation of Nrf2-mediated antioxidant signalling. Br J Pharmacol. 2023;180(24):3194–3214. https://doi.org/10.1111/bph.16195
43. Alenazi F, Moursi S, Mahmoud MR, et al. Withaferin A alleviates inflammation in animal models of arthritis by inhibiting the NF-κB pathway and cytokine release. Chem Biol Interact. 2024;398:111114. https://doi.org/10.1016/j.cbi.2024.111114
44. Liu T, Wang J, Tong Y, et al. Integrating network pharmacology and animal experimental validation to investigate the action mechanism of oleanolic acid in obesity. J Transl Med. 2024;22(1):86. https://doi.org/10.1186/s12967-023-04840-x
45. Peng W, Yang Y, Lu H, et al. Network pharmacology combines machine learning, molecular simulation dynamics and experimental validation to explore the mechanism of acetylbinankadsurin A in the treatment of liver fibrosis. J Ethnopharmacol. 2024;323:117682. https://doi.org/10.1016/j.jep.2023.117682