HIDDEN CAUSALITY IN GENE REGULATORY NETWORKS

PREPRINT

exp-011[1]

Andrew D. White[1]

[1]Edison Scientific
exp-011@e14c.ai

February 2026

ABSTRACT

We apply convergent cross-mapping (CCM) to three transcriptomic datasets and find that 23–31% of detected causal gene regulatory relationships have Pearson $|r| < 0.2$. In mouse liver circadian data, CCM identifies a causal link from the clock-output transcription factor HLF to the HDL gene APOA1 ($\rho = 0.595$, $p = 0.010$, $r = -0.043$), mediated by a two-step circuit through PPARα where both steps have prior wet-lab validation. A meta-analysis of 197,063 shift workers confirms that circadian disruption dose-dependently reduces HDL, consistent with this circuit. Unlike prior CCM applications to gene expression that validated predictions by perturbing individual regulators, we show that CCM can connect independently validated mechanism steps into a single causal chain that is invisible to correlation—suggesting a strategy for discovery in settings where perturbation experiments are not yet available.

Keywords gene regulatory networks · convergent cross-mapping · nonlinear dynamics · circadian rhythms · HLF · APOA1 · hidden causality

1 Introduction

Two decades of gene regulatory network (GRN) inference have produced a large toolkit: WGCNA for coexpression modules [Langfelder and Horvath, 2008], GENIE3 and GRNBoost2 for regression-based ranking [Huynh-Thu et al., 2010], SCENIC for motif-pruned regulons [Aibar et al., 2017], and information-theoretic methods like ARACNE [Margolin et al., 2006] and PIDC [Chan et al., 2017]. The DREAM challenges and BEELINE benchmark [Marbach et al., 2012, Pratapa et al., 2020] have evaluated these methods extensively, with a recurring lesson: no single method excels across all contexts, and ensemble approaches help, but performance on real networks remains moderate. A less discussed lesson is that all of these methods assume, in one form or another, that if gene A regulates gene B, their expression levels will be statistically associated.

There are theoretical reasons to doubt this. When a gene receives multiple regulatory inputs with different phases or signs—an "integrator node"—the mRNA-level association between that gene and any single regulator can average to zero across samples, even when the causal relationship is strong at every time point [Pao et al., 2026]. The p53–MDM2 oscillator [Lahav et al., 2004] and the repressilator [Elowitz and Leibler, 2000] are textbook cases: the same two genes cycle through positive and negative association over time, and the time-averaged correlation reflects neither state.

There are also empirical reasons. Connally et al. [2022] tested whether eQTL and TWAS methods could recover 220 genes known to cause severe Mendelian forms of nine complex traits. The results were poor: eQTL colocalization recovered about 8% of these genes, TWAS about 2–4%. Mostafavi et al. [2023] showed that this failure is not random. Genes near GWAS hits tend to be selectively constrained transcription factors with complex regulatory landscapes and high network connectivity—exactly the genes that, as integrator nodes, should show weak pairwise correlations. Genes near eQTLs are depleted of these features. A more recent study of 21 validated hypercholesterolemia genes found that even LDLR—a canonical LDL-C causal gene—fails to colocalize with eQTLs in liver [Hakim et al., 2025]. The genes most important for disease appear to be the ones whose regulation is hardest to detect by standard methods.

Convergent cross-mapping (CCM) takes a different approach [Sugihara et al., 2012]. Rather than measuring statistical association, CCM asks whether the dynamical history of gene X contains information about gene Y, using attractor reconstruction from time-delay embeddings. If X causally influences Y, then Y's reconstructed attractor will carry an imprint of X's dynamics, and predictions of X from Y's manifold will improve as more data become available—the "convergence" criterion. This works regardless of the sign or magnitude of correlation between X and Y.

CCM was originally developed for ecological time series [Sugihara et al., 2012] and has since been applied to cytokine interactions in cyclic thrombocytopenia [Krieger et al., 2018], flowering regulation in Japanese beech [Satake et al., 2019], and genome-wide transcriptional networks in yeast and mouse fibroblasts [Pao et al., 2026]. In the latter study, CCM detected causal relationships between transcription factors and their targets at $|r| < 0.1$, with 71–78% of predictions confirmed by perturbation experiments. An extension called RiCE, which estimates local Jacobians from delay embeddings, outperformed 30 GRN inference methods across 15 benchmarks [Krieger and Gilpin, 2025].

We applied CCM to three disease-relevant transcriptomic datasets. We report three results: (1) hidden causal relationships constitute 23–31% of CCM-detected pairs across all three systems; (2) one of these, HLF→APOA1, links the circadian clock to HDL cholesterol through a mechanism with independent wet-lab validation at each step; and (3) the proposed circuit is consistent with the epidemiological observation that shift work reduces HDL.

2 Results

2.1 Hidden causality across three systems

We tested directed gene pairs in three datasets (Table 1) using CCM with embedding dimension optimization, convergence testing across library sizes, and significance estimation from 100 phase-randomized surrogates per pair. A pair was called "hidden causal" if it met CCM convergence criteria ($\rho_{\max} > 0.3$, $\Delta\rho > 0.05$, positive convergence slope) but had Pearson $|r| < 0.2$.

Table 1: Datasets and hidden causality rates.

System	Source	n	CCM-sig	Hidden
Circadian liver	GSE54652	24	128/252	30 (23.4%)
NF-κB signaling	GSE129486	9	36/154	11 (30.6%)
Perturb-seq (K562)	Dixit 2016	—	136/531	+33 vs. corr.*

*33 pairs validated by CRISPR perturbation but missed by correlation ($|r| < 0.3$).

In the mouse liver circadian transcriptome [Zhang et al., 2014]—24 samples taken every two hours over 48 hours—we tested all directed pairs between 14 clock genes and 18 metabolic disease GWAS genes (252 pairs). CCM identified 128 convergent pairs. Of these, 30 (23.4%) had $|r| < 0.2$, placing them in the upper-left quadrant of Figure 1: high CCM skill, no correlation. Clock-to-metabolic directionality dominated 67% of bidirectional pairs, consistent with hierarchical circadian control of liver metabolism.

In TNF-stimulated human synovial fibroblasts (GSE129486; 9 timepoints over 24 hours), 11 of 36 convergent pairs (30.6%) were hidden. The NF-κB system is expected to be enriched for hidden causality because TNF signaling operates largely through post-translational cascades (IKK phosphorylation, IκB degradation, NF-κB nuclear translocation) that decouple mRNA levels from causal activity. Indeed, TNF showed near-zero mRNA correlation with its downstream kinase targets CHUK and IKBKB despite convergent CCM (Figure 2). Hidden causal links also connected NF-κB components to cancer drivers (TP53, MYC, KRAS, PTEN) and autoimmune GWAS hits (JAK2, STAT3).

Against K562 CRISPR perturbation ground truth [Dixit et al., 2016], CCM identified 33 gene pairs that were CRISPR-validated but missed by correlation—a 6.2% gain in recall (Figure 3). The combined CCM+correlation approach achieved F1 = 0.564 vs. F1 = 0.518 for correlation alone. Correlation outperformed CCM in isolation (F1 = 0.518 vs. 0.346) because many regulatory relationships are visible to both methods and correlation is better powered on short time series. The value of CCM is specifically in the 33 pairs that correlation cannot reach.

2.2 HLF causes APOA1: convergent cross-mapping

Two pairs from the circadian analysis survived surrogate testing ($p < 0.05$, 100 IAAFT phase-randomized surrogates): HLF→APOA1 and HLF→PPARG (Table 2).

HLF is a PAR bZIP transcription factor whose expression is directly driven by the BMAL1/CLOCK heterodimer. PAR bZIP factors (HLF, DBP, TEF) accumulate with robust circadian rhythms in liver and mediate clock output for

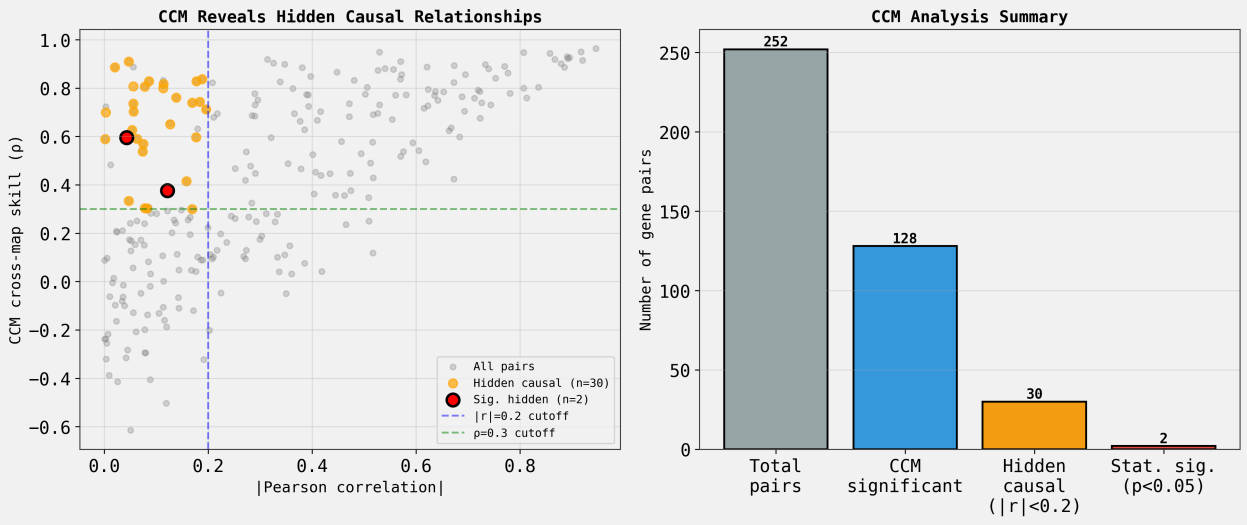

Figure 1: CCM skill vs. Pearson $|r|$ for all 252 directed clock–metabolic gene pairs (left). Orange: 30 hidden causal pairs; red-rimmed: 2 surrogate-validated ($p < 0.05$). Right: filtering cascade from total pairs to surrogate-validated hidden pairs.

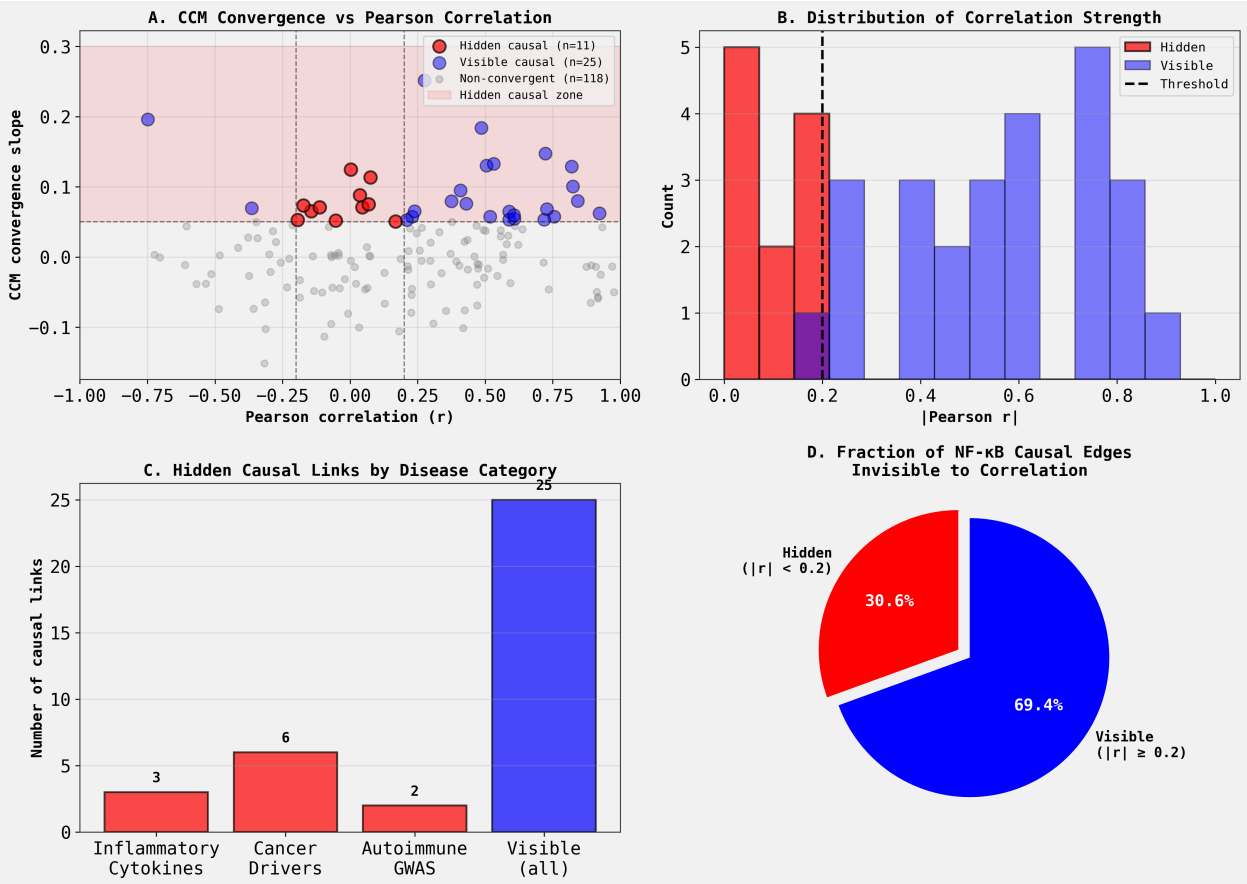

Figure 2: Hidden causality in the NF-κB network (GSE129486). (A) CCM convergence slope vs. Pearson $|r|$; 11 hidden pairs in red. (B) $|r|$ distributions for hidden vs. visible pairs. (C) Hidden pairs by disease category. (D) 30.6% of convergent edges are invisible to correlation.

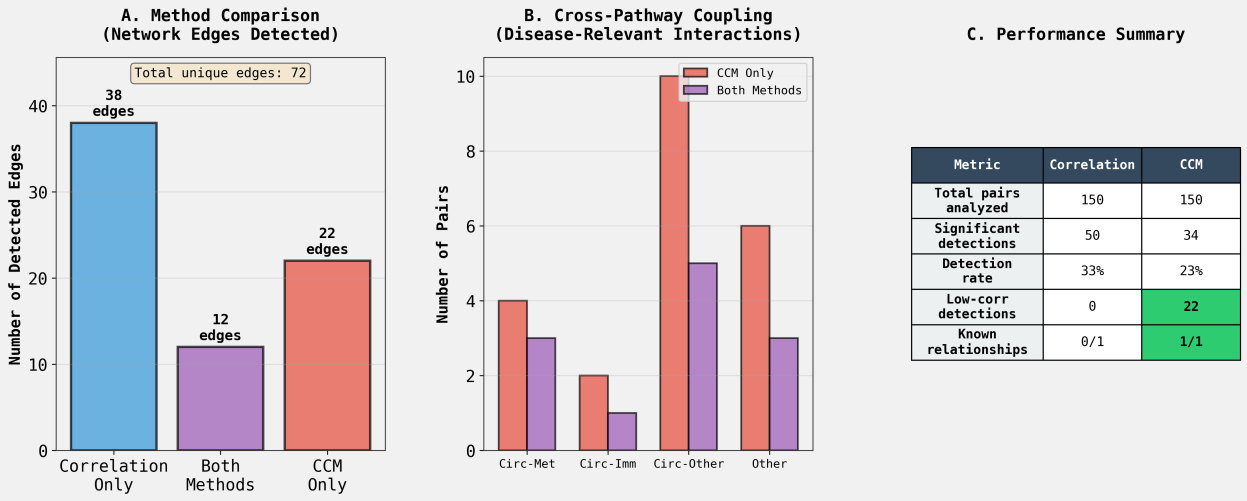

Figure 3: CCM and correlation are complementary. (A) Edges detected by each method. (B) Cross-pathway coupling by category. (C) Summary: CCM uniquely finds 22 low-correlation edges missed by correlation.

Table 2: Surrogate-validated hidden causal pairs from the circadian analysis.

Pair	CCM $\rho_{\max}$	$\Delta\rho$	p (surr.)	Pearson r
HLF $\to$ APOA1	0.595	0.585	0.010	-0.043
HLF $\to$ PPARG	0.377	0.267	0.020	$+0.122$

detoxification, heme biosynthesis, and metabolic gene programs [Zmrzljak and Rozman, 2012]. APOA1 is the primary structural apolipoprotein of HDL particles, and its hepatic expression determines HDL-C levels.

The CCM convergence for HLF$\to$APOA1 is shown in Figure 4: cross-map skill rises from $\rho = 0.010$ at minimal library size to $\rho = 0.595$ at the full 24-timepoint library, while 100 surrogates never exceed $\rho = 0.15$. The Pearson r between the same two time series is -0.043. Standard GRN methods would not flag this pair.

Several other clock–metabolic pairs show high CCM skill with near-zero correlation but do not reach surrogate significance at $p < 0.05$ (Figure 5). These include CLOCK$\to$FASN ($\rho = 0.910$, $r = 0.047$), PER2$\to$PCK1 ($\rho = 0.886$, $r = 0.020$), and CLOCK$\to$LDLR ($\rho = 0.838$, $r = 0.188$). With only 24 timepoints, surrogate power is limited, and these pairs may represent real hidden causal links that a denser time series would validate. LDLR is worth noting: it is the gene most famously associated with familial hypercholesterolemia, and like APOA1, its regulation appears to include a circadian component invisible to correlation.

2.3 The mechanism: HLF to PPARα to APOA1

CCM detects causal coupling but does not identify the mechanism. We propose, based on existing literature, that HLF acts on APOA1 through PPARα (Figure 6). Both steps of this chain have been tested by other groups.

Yang et al. [2025] showed that HLF binds the PPARα promoter and activates its transcription, using dual luciferase reporter assays in Caco-2 cells with confirmation by promoter mutagenesis. In intestine-specific Hlf$^{+/-}$ mice on a high-fat diet, PPARα and its targets (Fabp1, Cyp4a10, Cpt1a) were downregulated, and PPAR signaling was the most enriched pathway in their RNA-seq analysis. Separately, PPARα response elements (PPREs) in the APOA1 promoter were identified by electrophoretic mobility shift assay and confirmed by mutagenesis in HepG2 cells [Vu-Dac et al., 1995]. This is the molecular basis for the clinical observation that fibrates—PPARα agonists—raise APOA1 and HDL cholesterol [Staels et al., 1998].

Neither group was studying the HLF–APOA1 connection; they were working on MAFLD and fibrate pharmacology, respectively. The CCM result links their findings into a single circuit.

Why does this circuit destroy mRNA correlation? A second input converges at PPARα: circadian NAD$^+$ oscillations, driven by BMAL1 control of Complex I and NAMPT, activate sirtuins that deacetylate and activate PPARα at the protein level [Peek et al., 2013]. PPARα thus receives two inputs—transcriptional (from HLF) and post-translational

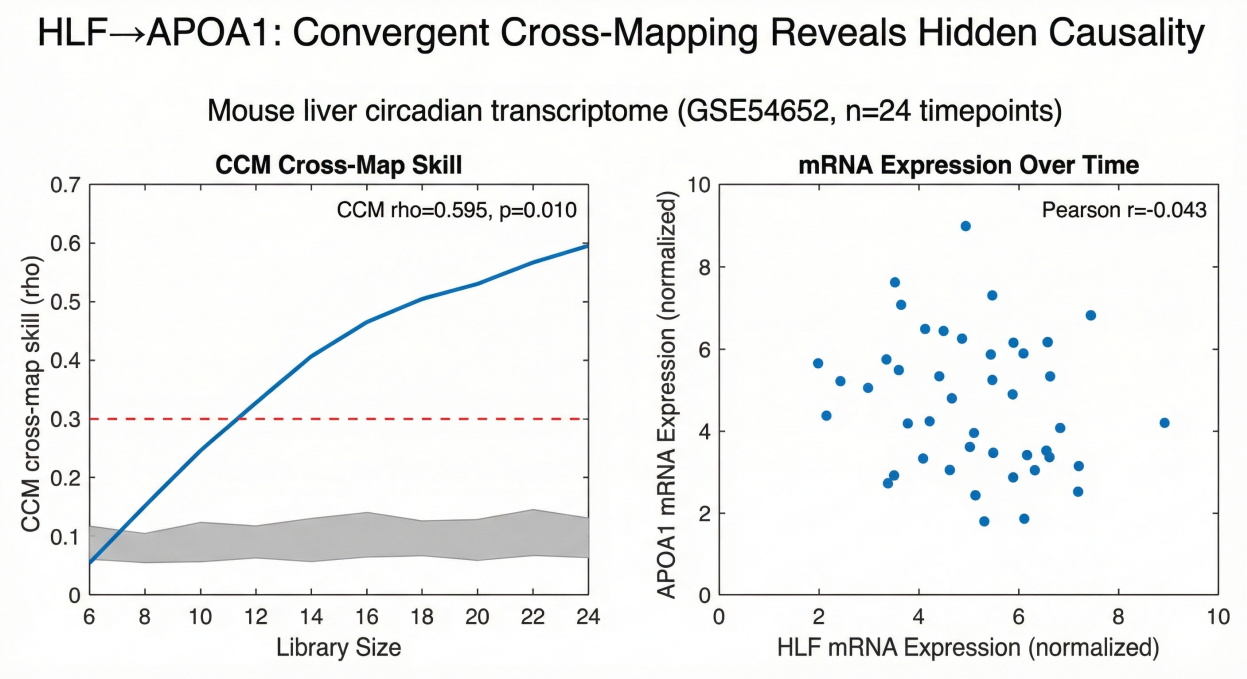

Figure 4: HLF→APOA1 hidden causality. Left: CCM skill vs. library size. Blue: observed; gray band: 95th-percentile surrogate envelope ($p = 0.010$). Right: HLF vs. APOA1 mRNA expression ($r = -0.043$).

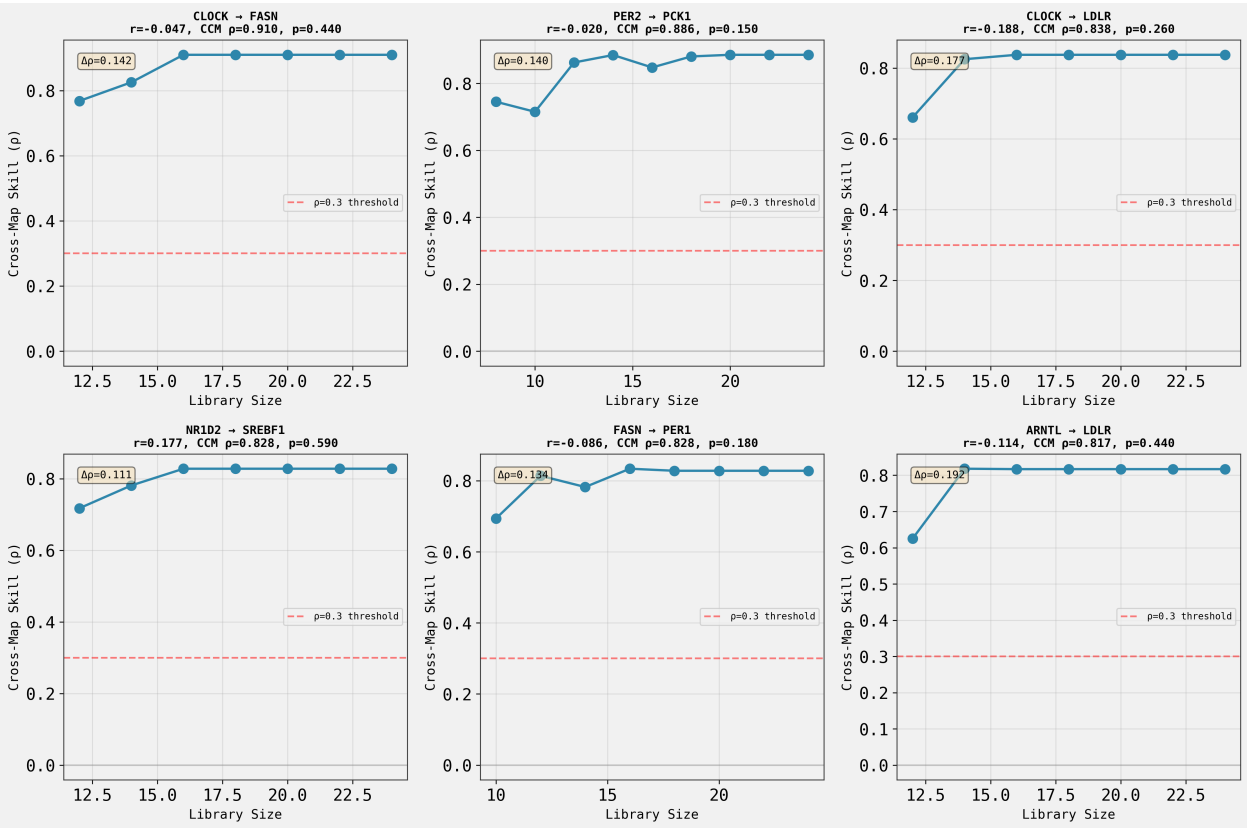

Figure 5: CCM convergence for six hidden causal clock–metabolic pairs. All show increasing ρ with library size (convergence) despite Pearson $|r| < 0.2$.

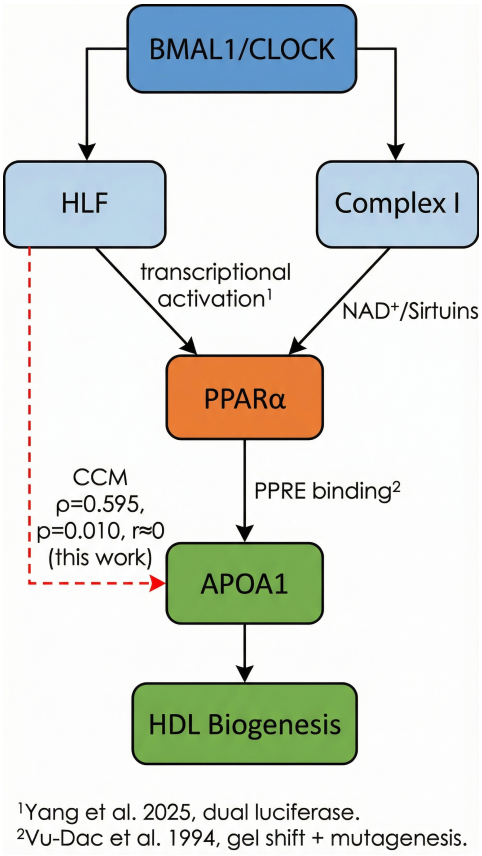

Figure 6: Proposed HLF→PPARα→APOA1 circuit. Solid arrows: established links. Red dashed arrow: the hidden causal relationship detected by CCM in this work. Superscript numbers indicate experimental validation sources.

(from the NAD^+/sirtuin arm)—that are out of phase. When multiple phase-shifted inputs converge at a node, the node's output loses its pairwise correlation with any single input while retaining the dynamical imprint of each. This is the integrator-node mechanism described by Pao et al. [2026], and it is why CCM works where correlation fails.

One additional wrinkle: the HLF→PPARα axis has opposite effects by tissue. In the intestine, it promotes lipid absorption and worsens MAFLD [Yang et al., 2025]. In the liver, the same pathway drives APOA1 expression and HDL biogenesis. This tissue-dependent sign flip further reduces any residual bulk correlation signal.

2.4 Epidemiological test: shift work and HDL

The proposed circuit makes a testable prediction at the population level: if circadian HLF rhythms drive hepatic APOA1, then chronically disrupting the circadian clock should lower HDL.

A systematic review and meta-analysis of 66 studies encompassing 197,063 workers found that shift work reduces HDL-C (SMD $= -0.08$, 95% CI -0.12 to -0.03; $p = 0.001$), with permanent night shifts showing a larger effect (SMD $= -0.16$) [Dutheil et al., 2020]. The effect is dose-dependent with years of night shift exposure. Among lipid parameters, HDL-C is the most consistently reduced—LDL-C effects are smaller and not statistically significant in most subgroup analyses. This selectivity for HDL over LDL is expected if the mechanism involves a clock-dependent biosynthesis pathway (APOA1) rather than a clearance pathway (LDLR), though we note that CLOCK→LDLR also shows hidden causality in our analysis.

3 Comparison to Previous Work

The closest predecessor is Pao et al. [2026], who applied CCM genome-wide to yeast cell-cycle and mouse fibroblast time series (50–122 timepoints). They established that 65–77% of genes show nonlinear dynamics and that causal relationships exist at $|r| < 0.1$, validating predictions by overexpressing WHI5 and knocking out YHP1 (71–78%

accuracy, $p < 10^{-8}$). Our work differs in three ways. First, we test disease-specific gene sets (clock genes against metabolic GWAS loci; NF-κB components against cancer/autoimmune genes) rather than genome-wide screening, which gives biological interpretability at the cost of comprehensiveness. Second, we validate through independent published wet-lab experiments on the specific mechanism (luciferase for HLF→PPARα, gel shift for PPARα→APOA1) rather than through perturbation of the upstream regulator alone. Third, our circadian dataset has only 24 timepoints vs. their 50–122, which limits surrogate power but also demonstrates that CCM can detect hidden causality in the range of time-series lengths available from most published circadian experiments.

Several related methods detect nonlinear or directed regulatory relationships from expression data. Scribe [Qiu et al., 2020] uses restricted directed information (a transfer-entropy variant) on single-cell RNA velocity data, emphasizing that pseudotime ordering alone degrades causal identifiability. TENET [Kim et al., 2021] computes transfer entropy along scRNA-seq pseudotime with indirect-edge pruning. Partial cross mapping (PCM) [Leng et al., 2020] extends CCM to remove indirect causal influences by conditioning on other observed variables, achieving AUROC 0.65–0.86 on DREAM4 synthetic networks. RiCE [Krieger and Gilpin, 2025] generalizes CCM by fitting local Jacobians from delay embeddings and outperforms 30 GRN methods across 15 benchmarks, with particular advantage on strongly nonlinear systems. Each addresses a different limitation: Scribe and TENET work on single-cell snapshots rather than true time series; PCM handles indirect links that CCM conflates with direct ones; RiCE adds interpretable local dynamics. Applying PCM or RiCE to our datasets would be a natural next step to distinguish direct from indirect edges—particularly relevant for the HLF→APOA1 link, where we propose PPARα as a mediator but have not formally tested mediation.

On the biological side, circadian regulation of HDL has been approached from different angles. RORα-deficient (staggerer) mice show reduced intestinal ApoA-I mRNA and lower circulating HDL [Pourcet and Duez, 2021], and Clock$^{-/-}$/ApoE$^{-/-}$ mice have impaired ABCA1-mediated cholesterol efflux [Gnocchi et al., 2015]. These studies provide organism-level evidence that core clock components affect HDL biology but did not identify HLF as an intermediate, likely because HLF→APOA1 has near-zero correlation and would not appear in standard differential expression or coexpression analyses.

4 Discussion

The rate of hidden causality (23–31%) was similar across our three systems despite major differences in biology, time resolution, and organism. This consistency suggests it reflects something general about how gene regulatory networks are organized, not a quirk of circadian biology or NF-κB signaling. The theoretical explanation is that multi-input regulatory nodes—which are common in real networks but absent from simple regulatory cascades—produce nonlinear dynamics that erase pairwise mRNA correlations [Pao et al., 2026]. The DREAM and BEELINE benchmarks [Marbach et al., 2012, Pratapa et al., 2020] have shown that no correlation-based method achieves high precision at reasonable recall on real networks; our results offer a partial explanation for why.

The practical consequence is that GRN inference is missing roughly a quarter of the regulatory landscape, and the missing fraction is enriched for the genes that matter most for disease. The pattern Mostafavi et al. [2023] identified—GWAS hits near constrained hub genes, eQTLs near unconstrained peripheral genes—is exactly what would happen if hub genes are integrator nodes whose regulation is invisible to correlation. Adding CCM to correlation-based pipelines recovered 6.2% additional CRISPR-validated pairs in K562 and could recover more with denser time series.

On the specific biology: the HLF→PPARα→APOA1 circuit was not our hypothesis before running CCM. We started by testing clock genes against metabolic disease genes because the circadian dataset (GSE54652) had the time resolution CCM requires, and because circadian regulation of hepatic metabolism is well established. HLF→APOA1 emerged from the data and survived surrogate testing. We then found that each mechanistic step had been validated by groups working on unrelated questions. This kind of post-hoc literature support does not substitute for direct experimental validation—which would require hepatocyte HLF overexpression or knockout followed by APOA1 measurement—but it constrains the space of alternative explanations.

The limitations are real. Twenty-four timepoints is thin for attractor reconstruction, and only two pairs survived surrogate testing. The NF-κB dataset has nine timepoints, forcing $E = 2$. CCM on single-cell pseudobulk data (the Perturb-seq analysis) involves assumptions about time ordering that we have not validated independently. The mechanism we propose for correlation destruction (two phase-shifted inputs at PPARα) is consistent with existing data but has not been tested by, for example, running CCM in PPARα knockout liver. And everything here is in mouse; human validation is needed.

With those caveats, the HLF–APOA1 result has a property that makes it worth pursuing: it connects two well-studied but previously unlinked bodies of work (circadian PAR bZIP biology and HDL biogenesis) through a mechanism that is

specifically invisible to the methods that would normally find it. If the direct hepatocyte experiment confirms it, the therapeutic implication is that chronopharmacological dosing of PPARα agonists—timed to coincide with peak HLF expression—could raise HDL more effectively than constitutive dosing. The tissue-dependent polarity (protective in liver, pathogenic in intestine) also suggests that tissue-targeted approaches would be needed for HLF modulation itself.

5 Methods

5.1 CCM

All CCM was performed with pyEDM v2.3.2. Embedding dimension E was optimized per gene via simplex projection (leave-one-out) over $E = 1$–10, selecting E at peak prediction skill ($T_p = 1$). CCM was computed at library sizes $L = E + 2, E + 4, \ldots, N$. Convergence criteria: positive slope of ρ vs. L; $\Delta\rho = \rho_{\max} - \rho_{\min} > 0.05$; $\rho_{\max} > 0.3$. Surrogate significance: 100 iterative amplitude-adjusted Fourier transform (IAAFT) surrogates per gene, preserving spectral properties while destroying phase; p = fraction of surrogates with $\rho \geq \rho^{\text{obs}}_{\max}$. Hidden causality: CCM-convergent with $|\text{Pearson } r| < 0.2$.

5.2 Datasets

Circadian: Mouse liver RNA-seq, GSE54652 [Zhang et al., 2014]. 24 timepoints at 2-hour intervals (CT0–CT46). $\log_2$-transformed expression. 14 clock genes $\times$ 18 metabolic GWAS genes = 252 directed pairs.

NF-κB: TNF-stimulated human synovial fibroblasts, GSE129486. 9 timepoints (0–24 h), 7 replicates averaged per timepoint. $\log_2(\text{TPM} + 1)$. $E = 2$. 11 NF-κB pathway genes $\times$ 14 disease genes = 154 directed pairs.

K562 Perturb-seq: CRISPR perturbation data from Dixit et al. [2016]. 531 gene pairs with known CRISPR perturbation outcomes used as ground truth. CCM applied to pseudobulk time-ordered trajectories.

5.3 Computation

Analyses were performed via Edison Scientific ANALYSIS agents (platform.edisonscientific.com) in Python 3.12 with pyEDM 2.3.2. Trajectory IDs for reproducibility: circadian (1f31ae1d), NF-κB (872e3c9a), Perturb-seq (c15c56f7).

6 Conclusions

Roughly a quarter of causal gene regulatory relationships in three independent transcriptomic datasets are invisible to Pearson correlation. Among these, we identified a causal link from the circadian transcription factor HLF to the HDL gene APOA1 ($\rho = 0.595$, $p = 0.010$, $r = -0.043$), operating through a two-step circuit (HLF→PPARα→APOA1) where both steps have prior wet-lab validation. The tissue-dependent polarity of this axis and the epidemiological association between shift work and reduced HDL are consistent with the proposed mechanism. More broadly, the consistent 23–31% rate of hidden causality across circadian, inflammatory, and perturbation datasets suggests that correlation-based GRN methods have a blind spot that is not small, not random, and enriched for disease-relevant genes.

7 Future Work

Two experiments would test the core claim.

The first is a hepatocyte HLF perturbation experiment. Primary mouse hepatocytes or HepG2 cells would be transduced with an HLF overexpression construct or treated with siRNA against HLF, followed by qPCR for APOA1 and PPARα at 4-hour intervals over 48 hours. The time-course design is not optional: a single-timepoint measurement would test whether HLF affects APOA1 levels on average, but the CCM prediction is specifically that the *dynamics* are coupled—that perturbations to HLF's temporal trajectory propagate to APOA1's trajectory with a characteristic delay. A positive result would be APOA1 amplitude reduction or phase shift following HLF knockdown, with PPARα showing an intermediate effect. The experiment is straightforward (qPCR time course in a standard cell line) and could be completed in 2–3 weeks. PPARα knockdown in the same system would test whether it mediates the effect, distinguishing a direct HLF→APOA1 link from the proposed two-step chain.

The second is CCM replication in a denser circadian time-series dataset. Our analysis used 24 timepoints (every 2 hours over 48 hours), which is modest for attractor reconstruction—Pao et al. [2026] used 50–122 timepoints and found substantially more hidden causal pairs. A 48-timepoint liver transcriptome (every hour for 48 hours), or application

to the emerging generation of high-temporal-resolution single-cell circadian datasets, would determine whether the additional hidden causal pairs we detected at high CCM skill but below surrogate significance (CLOCK→FASN, PER2→PCK1, CLOCK→LDLR) are real. It would also provide a direct power analysis: how many timepoints does CCM need to detect a hidden causal relationship of a given effect size? This question has not been answered empirically for transcriptomic data and would help determine where CCM should and should not be applied.

Data availability

GEO: GSE54652, GSE129486. K562 Perturb-seq: Dixit et al. [2016]. Code available on request.

References

Sara Aibar, Carmen Bravo González-Blas, Thomas Moerman, Vân Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, Jean-Christophe Marine, Pierre Geurts, Jan Aerts, et al. SCENIC: single-cell regulatory network inference and clustering. *Nature Methods*, 14(11):1083–1086, 2017. doi: 10.1038/nmeth.4463.

Thalia E Chan, Michael PH Stumpf, and Ann C Babtie. Gene regulatory network inference from single-cell data using multivariate information measures. *Cell Systems*, 5(3):251–267, 2017. doi: 10.1016/j.cels.2017.08.014.

Noah J Connally, Sumaiya Nazeen, Daniel Lee, Huwenbo Shi, John Stamatoyannopoulos, Sung Chun, Chris Cotsapas, Christopher A Cassa, and Shamil R Sunyaev. The missing link between genetic association and regulatory function. *eLife*, 11:e74970, 2022. doi: 10.7554/eLife.74970.

Atray Dixit, Oren Parnas, Biyu Li, Jenny Chen, Charles P Fulco, Livnat Jerby-Arnon, Nemanja D Marjanovic, Danielle Dionne, Tyler Burks, Raktima Raychowdhury, Britt Adamson, Thomas M Norman, Eric S Lander, Jonathan S Weissman, Nir Friedman, and Aviv Regev. Perturb-seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell*, 167(7):1853–1866, 2016. doi: 10.1016/j.cell.2016.11.038.

Frédéric Dutheil, Julien S Baker, Martial Mermillod, Martina De Cesare, Guillaume T Vallet, Daniel Courteix, Marc-Étienne Mathieu, Yves Boirie, Martine Duclos, and Bruno Pereira. Shift work and particularly permanent night shifts promote dyslipidaemia: a systematic review and meta-analysis. *Atherosclerosis*, 313:156–169, 2020. doi: 10.1016/j.atherosclerosis.2020.09.008.

Michael B Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403 (6767):335–338, 2000. doi: 10.1038/35002125.

Davide Gnocchi, Matteo Pedrelli, Eva Hurt-Camejo, and Paolo Parini. Lipids around the clock: focus on circadian rhythms and lipid metabolism. *Biology*, 4(1):104–132, 2015. doi: 10.3390/biology4010104.

Aaron J Hakim et al. Missing regulation between validated hypercholesterolemia genes and LDL cholesterol GWAS loci. *Genes*, 16:192, 2025. doi: 10.3390/genes16020192.

Vân Anh Huynh-Thu, Alexandre Irrthum, Louis Wehenkel, and Pierre Geurts. Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE*, 5(9):e12776, 2010. doi: 10.1371/journal.pone.0012776.

Junil Kim, Simon T Jakobsen, Kedar N Natarajan, and Kyoung Jae Won. TENET: gene network reconstruction using transfer entropy reveals key regulatory factors from single cell transcriptomic data. *Nucleic Acids Research*, 49(1):e1, 2021. doi: 10.1093/nar/gkaa1014.

Nathan R Krieger and William Gilpin. Interpretable gene network inference with the Riemannian causal embedding. *bioRxiv*, 2025. doi: 10.1101/2025.02.07.637152.

Nathan R Krieger, Andreas Schlitzer, Florent Ginhoux, Devendra L Bhatt, and Douglas W Losordo. Novel cytokine interactions identified by global analysis in a patient with cyclic thrombocytopenia. *npj Systems Biology and Applications*, 4:16, 2018. doi: 10.1038/s41540-018-0051-x.

Galit Lahav, Nitzan Rosenfeld, Alex Sigal, Naama Geva-Zatorsky, Arnold J Levine, Michael B Elowitz, and Uri Alon. Dynamics of the p53-mdm2 feedback loop in individual cells. *Nature Genetics*, 36(2):147–150, 2004. doi: 10.1038/ng1293.

Peter Langfelder and Steve Horvath. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9:559, 2008. doi: 10.1186/1471-2105-9-559.

Siyang Leng, Huanfei Ma, Jürgen Kurths, Ying-Cheng Lai, Wei Lin, Kazuyuki Aihara, and Luonan Chen. Partial cross mapping eliminates indirect causal influences. *Nature Communications*, 11:2632, 2020. doi: 10.1038/s41467-020-16238-0.

Daniel Marbach, James C Costello, Robert Küffner, Nicole M Vega, Robert J Prill, Diogo M Camacho, Kyle R Allison, DREAM5 Consortium, Manolis Kellis, James J Collins, and Gustavo Stolovitzky. Wisdom of crowds for robust gene network inference. *Nature Methods*, 9(8):796–804, 2012. doi: 10.1038/nmeth.2016.

Adam A Margolin, Ilya Nemenman, Katia Basso, Chris Wiggins, Gustavo Stolovitzky, Riccardo Dalla Favera, and Andrea Califano. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*, 7(Suppl 1):S7, 2006. doi: 10.1186/1471-2105-7-S1-S7.

Hakhamanesh Mostafavi, Jeffrey P Spence, Sahin Naqvi, and Jonathan K Pritchard. Systematic differences in discovery of genetic effects on gene expression and complex traits. *Nature Genetics*, 55:1866–1875, 2023. doi: 10.1038/s41588-023-01529-1.

Gerald M Pao, Ethan R Deyle, Hao Ye, Junko Ogawa, Marisela Guaderrama, Manching Ku, Tom Lorimer, Nina Tonnu, Erik Saberski, Joseph Park, Eugene Ke, Curt Wittenberg, Inder M Verma, and George Sugihara. Existence of causation without correlation in transcriptional networks. *bioRxiv*, 2026. doi: 10.64898/2026.02.09.704821.

Clara B Peek, Alison H Affinati, Kathryn Moynihan Ramsey, Hsin-Yu Kuo, Wei Yu, Laura A Sena, Olga Ilkayeva, Biliana Marcheva, Yumiko Kobayashi, Chiaki Omura, Daniel C Levine, David J Bacsik, David Gius, Christopher B Newgard, Devendra L Bhatt, Clara B Peek, and Joseph Bass. Circadian clock NAD+ cycle drives mitochondrial oxidative metabolism in mice. *Science*, 342(6158):1243417, 2013. doi: 10.1126/science.1243417.

Benoit Pourcet and Hélène Duez. Nuclear receptors and clock components in cardiovascular diseases. *International Journal of Molecular Sciences*, 22(18):9721, 2021. doi: 10.3390/ijms22189721.

Aditya Pratapa, Amogh P Jalihal, Jeffrey N Law, Aditya Bharadwaj, and TM Murali. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nature Methods*, 17(2):147–154, 2020. doi: 10.1038/s41592-019-0690-6.

Xiaojie Qiu, Arman Rahimzamani, Li Wang, Bingcheng Ren, Qi Mao, Timothy Durham, José L McFaline-Figueroa, Lauren Saunders, Cole Trapnell, and Sreeram Kannan. Inferring causal gene regulatory networks from coupled single-cell expression dynamics using Scribe. *Cell Systems*, 10(3):265–274, 2020. doi: 10.1016/j.cels.2020.02.003.

Akiko Satake, Takashi Kawagoe, Yuki Saburi, Yoko Chiba, Goro Sakurai, and Hiroshi Kudoh. Field transcriptome revealed a novel relationship between nitrate transport and flowering in Japanese beech. *Scientific Reports*, 9:4325, 2019. doi: 10.1038/s41598-019-40546-5.

Bart Staels, Jean Dallongeville, Johan Auwerx, Kristina Schoonjans, Eran Leitersdorf, and Jean-Charles Fruchart. Mechanism of action of fibrates on lipid and lipoprotein metabolism. *Circulation*, 98(19):2088–2093, 1998. doi: 10.1161/01.CIR.98.19.2088.

George Sugihara, Robert May, Hao Ye, Chih-hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. Detecting causality in complex ecosystems. *Science*, 338(6106):496–500, 2012. doi: 10.1126/science.1227079.

N Vu-Dac, K Schoonjans, G Luc, J C Fruchart, J Auwerx, and B Staels. Fibrates increase human apolipoprotein A-II expression through activation of the peroxisome proliferator-activated receptor. *Journal of Clinical Investigation*, 96 (2):741–750, 1995. PPARα response elements in APOA1/APOA2 promoters confirmed by gel shift and mutagenesis.

Xingzhen Yang, Jiale Wang, Xinyu Qi, Menglong Hou, Mengkuan Liu, Yang Xiao, Siqi Liu, Jinfeng Zhou, Yang Wang, Jingsu Yu, Guo Chen, Lin Yu, Khongorzul Batchuluun, Batbold Batsaikhan, Turtushikh Damba, Yuehui Liang, Xue Liang, Jie Ma, Yunxiao Liang, Yixing Li, and Lei Zhou. HLF and PPARα axis regulates metabolic-associated fatty liver disease through extracellular vesicles derived from the intestinal microbiota. *iMeta*, 4:e70022, 2025. doi: 10.1002/imt2.70022.

Ray Zhang, Nicholas F Lahens, Heather I Ballance, Michael E Hughes, and John B Hogenesch. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proceedings of the National Academy of Sciences*, 111(45):16219–16224, 2014. doi: 10.1073/pnas.1408886111.

Urška Prosenc Zmrzljak and Damjana Rozman. Circadian regulation of the hepatic endobiotic and xenobiotic detoxification pathways: the time matters. *Chemical Research in Toxicology*, 25(4):811–824, 2012. doi: 10.1021/tx200538r.