Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Copper-stress-associated cell-state remodeling and auditable candidate prioritization in AML12 hepatocytes

Jinwei Li

1791605967@qq.com

Huazhong University of Science and Technology

Article

Keywords: copper stress, AML12, MYH9 template, cytoskeletal remodeling, candidate ranking, DESeq2, limma-voom, composition bias

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11262191/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Copper-stress-associated cell-state remodeling and auditable candidate prioritization in AML12 hepatocytes

Author (first and corresponding): Jinwei Li

Affiliation: Huazhong University of Science and Technology

Email: 1791605967@qq.com

Article type: Hypothesis-generating bioinformatics reanalysis

Main question: Using DESeq2 on uniformly reprocessed integer counts as the primary inference framework, which Cu/ES-associated gene-expression changes in AML12 stay directionally concordant across count-based methods, and which structural, copper-handling, metabolic and stress-response genes merit experimental prioritization when MYH9 is treated only as an external template rather than a competing candidate?

Main contribution: A computationally auditable candidate-priority framework, not validation of a cuproptosis execution mechanism.

Abstract

Background: Copper imbalance is relevant to liver injury, Wilson disease, metabolic liver disease and copper-ionophore cancer therapy. Cuproptosis is a literature-defined copper-dependent death concept, but its downstream structural events, candidate boundary and cross-cell transferability remain unresolved. A 2026 Cell study proposed copper-induced MYH9 polymerization and actin disruption as an external mechanism template. That template can guide candidate ranking, but it cannot substitute for independent validation in AML12 cells.

Methods: I reanalyzed GSE300382, a mouse AML12 experiment containing control, CuCl2 plus elesclomol (Cu/ES), Cu/ES plus taurocholic acid (TCAu) and TCAu alone, with three biological replicates per group. Associated SRA reads were uniformly processed against GRCm39 to generate integer counts; DESeq2 was the primary differential-expression and factorial model, and edgeR and limma-voom were used for confirmation on the same 12 samples. The public FPKM matrix was analyzed with limma only as a historical comparison with the original public matrix, not as the primary inference framework. A prespecified universe of 77 structural, copper-handling, metabolic and stress-related genes was assessed in template-referenced, leave-template-out and template-reduced class-adjusted models; MYH9 was used only as a fixed external template and was not included in candidate ranking. HPA was used only as human expression context with zero weight. PXD037278 was removed because its raw-file-to-condition mapping could not be verified.

Results: DESeq2 identified broad Cu/ES-associated changes in both copper-containing contrasts (12,414 and 12,677 genes at FDR<0.05 and |log2FC|>=1). edgeR identified 12,470 and 12,676 genes, and limma-voom identified 13,432 and 13,515 genes for the same contrasts. TCAu alone versus control produced a much smaller

response (353, 343 and 379 genes, respectively). Count-based results were directionally concordant for 15 of 17 prespecified key genes; Myh9 and Mt2 were discordant, Atp7a was concordant but not significant under limma-voom, and FDX1 was negative and significant under all three count methods. The factorial interaction was method-sensitive and was retained only as a secondary technical sensitivity. Copper-containing conditions showed increased proteostasis and redox-defense signatures and reduced mitochondrial, cell-cycle, adhesion, cytokinesis and actin-myosin transcript programs. Count-level composition changed: genes with count>0 increased, the count>=10 set remained nearly stable, and the ribosomal read fraction rose; FPKM-based detected-gene and ribosomal metrics are reported separately. Candidate rankings were sensitive to template and class priors. MYH9 was used only as an external template and was excluded from the candidate ranking; its separate sensitivity term moved from rank 2 to rank 48 when template-linked evidence was removed, which describes model behavior rather than a biological rank. CD2AP, RDX, PDHA1, FDX1, LIAS and PDHB retained higher model-based priority within the prespecified universe in some models, not data-supported mechanisms. ATP7A-MYH9 was a direction-agnostic transcript association, and MYH9-ANLN remained an external interaction hypothesis.

Structural context: Ten AlphaFold3/ColabFold v2 input specifications were generated but not executed, and no predicted coordinates or copper-conditioned results exist. Three homologous or orthologous templates completed copper-free, short-timescale molecular dynamics; these exploratory calculations do not provide binding, copper-coordination or AML12-complex evidence.

Conclusion: Cu/ES is associated with broad AML12 cell-state remodeling and supports template-dependent model-based prioritization across structural, copper-handling, metabolic and stress-response candidate layers. The main contribution is a transparent, computationally auditable candidate-priority framework in which each evidence domain can be removed and audited separately. A step-by-step evidence-chain check limits same-study support to the transcript and pathway levels and leaves transcription-factor and chromatin steps as testable hypotheses. No independent biological replication was achieved; the evidence boundary is stated once below.

Evidence boundary statement: This is a hypothesis-generating reanalysis of public transcriptomic data. Copper binding, MYH9 polymerization/acetylation, DLAT lipoylation/release, F-actin/G-actin fractionation, paired AML12 chromatin measurements, genetic perturbation and rescue were not performed. The study therefore supports copper-stress-associated cell-state changes, analytical concordance and candidate-priority relationships, but not cuproptosis execution, direct protein interaction or causal mechanism. Later sections refer to this statement rather than repeating the full boundary.

Terminology (MYH9-related names): Four objects are kept strictly separate. The **MYH9 external-template score** is a fixed prior from external literature used to compute similarity; it is not a candidate gene in this study. The **MYH9-as-a-candidate sensitivity term** is a separate diagnostic computed only to audit template dependence, and is reported outside the candidate ranking. The **template-derived MYH9 term** is the ranking object produced by the MYH9 prior in the full-template score; its position reflects how the score was built, not a biological rank.

The **candidate ranking excluding MYH9** is the ranking table for real candidate genes; MYH9 does not appear in that table or in its figures. The MYH9-as-a-candidate sensitivity term moves from rank 2 to rank 48 when template-linked evidence is removed; this describes model behavior and must never be read as a biological candidate rank for MYH9 in AML12.

Keywords: copper stress; AML12; MYH9 template; cytoskeletal remodeling; candidate ranking; DESeq2; limma-voom; composition bias

1. Introduction

Copper is required for redox chemistry, mitochondrial respiration and antioxidant defense, but excess labile copper can cause oxidative damage, proteotoxic stress and cell death. ATP7B deficiency in Wilson disease, chronic hepatic copper accumulation and copper-ionophore cancer therapy illustrate the medical relevance of copper imbalance. The same clinical context also exposes a major gap: cell-type sensitivity cannot currently be converted into a specific downstream structural mechanism.

The strongest cuproptosis literature links copper-dependent death to mitochondrial metabolism, protein lipoylation, Fe-S cluster instability and protein aggregation. The FDX1-LIAS-DLAT axis is widely discussed, but the term "executioner" remains poorly defined. The recent MYH9 polymerization study adds a cytoskeletal hypothesis; here MYH9 is fixed as an external template rather than treated as an independent finding from GSE300382.

This study asks two limited questions. First, which gene-expression changes in Cu/ES-treated AML12 hepatocytes remain directionally concordant across count models? Second, under an external MYH9-polymerization template, which structural, copper-handling, metabolic and stress-response candidates should be prioritized, and how sensitive is that ordering to template, class prior and composition assumptions?

2. Results

2.1 An auditable evidence structure and the main question

The primary question is whether Cu/ES-associated gene-expression changes remain directionally concordant at count level, and which structural, copper-handling, metabolic and stress-response genes merit experimental prioritization without treating the external MYH9 mechanism as a candidate. The analysis has four evidence layers. Layer 1 is the primary integer-count analysis using GRCm39, featureCounts and DESeq2, with edgeR and limma-voom as count-level confirmation; these count results are the main inference framework in this paper. Layer 2 is the historical FPKM-limma description, used only for comparability with the public matrix. Layer 3 is the candidate-priority framework, in which MYH9 is fixed as an external template and is not ranked as a candidate. Layer 4 is external background evidence.

This structure records template dependence, external context and model inference as separate evidence layers and makes each layer auditable.

Naming is fixed throughout. The **primary model** is DESeq2 on integer counts. **Count-based confirmation** is edgeR and limma-voom applied to the same 12 samples. The **historical comparator** is FPKM-limma on the public matrix. **Analytical concordance** means agreement among the three count-based methods; it is not independent biological replication. The historical cross-scale comparison is reported separately and is never used to define the primary result.

2.2 Raw-count analysis: primary inference framework

All 12 SRA samples were uniformly reprocessed. A total of 268,256,905 reads were generated; 265,651,384 mapped to GRCm39, giving an overall mapping rate of 99.03% (sample range 98.54%-99.38%). GRCm39 annotation yielded 78,028 gene symbols. DESeq2 and limma-voom each tested 19,275 filtered genes; edgeR tested 18,974 genes after filterByExpr. This count-level input set is not directly comparable with the 11,554 FPKM-filtered genes. On the 11,141 genes shared by the count filter and the FPKM filter, the three count methods changed 6,712 (DESeq2), 6,833 (edgeR) and 6,840 (limma-voom) genes for Cu/ES versus control; 6,467, 7,054 and 7,073 genes for Cu/ES+TCAu versus TCAu; 146, 124 and 133 genes for TCAu alone; and 950, 885 and 855 genes for the interaction. These common-set estimates are lower than the full-count values because the common set excludes genes that the FPKM filter removes, and they are distinct from the 7,135 and 7,477 copper-comparison genes reported by FPKM-limma. In the primary DESeq2 model, Cu/ES versus control and Cu/ES+TCAu versus TCAu changed 12,414 and 12,677 genes at FDR<0.05 and |log2FC|>=1. edgeR changed 12,470 and 12,676 genes, and limma-voom changed 13,432 and 13,515 genes under the same threshold. TCAu alone versus control changed 353 genes in DESeq2, 343 in edgeR and 379 in limma-voom. For the 17 prespecified key genes, 15 showed the same direction across the three count-based methods; Myh9 and Mt2 were directionally discordant. Fourteen key genes were significant in all three methods, and 16 were significant in at least one. Atp7a was directionally concordant but not significant in limma-voom. FDX1 was negative and significant in all three methods.

The historical FPKM-limma comparison is reported separately and never used to define the primary result. It shows broad directional agreement with the count-based analyses, but its statistical scale is not interchangeable with the primary model. The TCAu_at_CuES sensitivity contrast was method-dependent: DESeq2 was inversely correlated with historical FPKM-limma (Spearman=-0.548; shared-DEG direction concordance 2.45%), whereas edgeR and limma-voom reached 0.899 and 0.898, respectively. This contrast therefore remains a method-sensitivity result. DESeq2 is the primary model, with edgeR and limma-voom as count-based confirmation on the same 12 samples.

2.3 Cu/ES is associated with the dominant gene-expression state change

The primary count-level DESeq2 analysis compared Cu/ES versus control and Cu/ES+TCAu versus TCAu with their respective non-copper controls. These comparisons changed 12,414 and 12,677 genes at FDR<0.05 and |log2FC|>=1. edgeR changed 12,470 and 12,676 genes, and limma-voom changed 13,432 and 13,515 genes. TCAu alone versus control changed 353 genes in DESeq2, 343 in edgeR and 379 in limma-voom. Cu/ES was therefore the condition associated with the dominant transcriptome-scale response. The historical FPKM-limma matrix reported 7,135 and 7,477 genes for the corresponding copper comparisons and 24 genes for TCAu alone; those values are retained only for historical comparability and are not used for primary inference.

Copper-containing conditions increased heat-shock, chaperone, translation and ribosome-associated programs. Mitochondrial respiration, lipid catabolism, adhesion, actin-myosin and mitotic transcripts were reduced. This pattern is consistent with increased proteostatic burden and a shift in cell state, but it does not distinguish selective survival of proliferative cells, altered dead-cell fraction, library-composition effects and true program suppression.

The difference in the TCAu_at_CuES comparison between count-based models and FPKM-limma should be interpreted as a method-dependent interaction result, not as evidence for two incompatible biological effects. Count and FPKM inputs differ in gene filtering, normalization, mean-variance modelling and sensitivity to low-count genes; after composition adjustment and leave-one-sample-out resampling, the interaction was not stable enough for a discovery claim.

2.4 Composition bias is central to interpretation

In the public FPKM matrix, copper-containing conditions reduced the mean number of FPKM-detected genes from about 10,842 to about 8,510 and increased the ribosomal FPKM fraction from about 0.192 to about 0.287. In the count matrix the same contrast is different: the number of genes with count>0 increased from 25,454 to 29,325, while the number of genes with count>=10 stayed essentially constant (Control 16,066; Cu/ES 16,065; Cu/ES+TCAu 16,691). The apparent loss of detected genes is therefore a property of the FPKM>=1 threshold and of how RNA is distributed across genes, not a count-level disappearance of genes. The count-level ribosomal read fraction nevertheless rose in the same direction, from 0.0393 in control to 0.0681 in Cu/ES and 0.0624 in Cu/ES+TCAu, so the higher ribosomal share is visible in both representations. FPKM-detected genes, ribosomal FPKM fraction, cell-cycle score and oxidative-stress score are useful for historical comparability, but they are expression-fraction summaries rather than count-based quality-control metrics; the count-level metrics above are the ones most directly comparable to the primary analysis.

Two other composition metrics must be defined before they are compared. The cell-cycle score and the oxidative-stress score are column means of gene-level z-scores over the GO "cell cycle" and "response to oxidative stress" sets, respectively. The proteostasis/redox-defense module is a separate curated response set. The oxidative-stress z-score fell (mean change -0.903) while the proteostasis/redox-defense module rose (about +0.82); these two scores are built from different gene sets and are not interchangeable, so the study does not claim a single

"oxidative-stress level". Housekeeping dispersion was the per-sample standard deviation of log2(FPKM+1) across a fixed ten-gene housekeeping set (Actb, Gapdh, Rpl13a, Rps13, Tbp, Ppia, B2m, Hprt, Sdha, Ywhaz); it increased by 0.672, indicating that even reference transcripts were unevenly expressed between samples.

Together, these shifts indicate that cytoskeletal and mitotic transcript reduction may partly reflect surviving-cell selection and RNA composition rather than specific structural-gene inhibition.

The manuscript therefore uses "cell-state-associated reduction under copper-containing conditions" and avoids claims of direct cytoskeletal inhibition or MYH9-mediated collapse.

2.5 The Cu/ES-by-TCAu interaction is a secondary technical sensitivity

The interaction was defined as (Cu/ES+TCAu - TCAu) - (Cu/ES - Control), the difference in the TCAu-associated effect between copper-containing and non-copper-containing states. In the primary count analysis, DESeq2 reported 1,455 interaction genes at the prespecified threshold; edgeR and limma-voom reported 1,284 and 1,283. These counts alone are not treated as a biological discovery because the interaction was sensitive to model choice and composition. In the historical FPKM-limma analysis, 4,219 genes were reported at FDR<0.05, including 393 with absolute interaction log2FC>=1. After adjustment for detected-gene number and ribosomal fraction, only 67 interactions remained significant. Direction concordance was 96.42% (4,068/4,219), but the median leave-one-sample-out interaction correlation was -0.0127, showing that directional agreement is not equivalent to stable effect magnitude.

The interaction is therefore reported only as a composition-sensitive technical sensitivity and is moved to Extended Data. TCAu changes relative expression responses; it does not demonstrate inhibition of MYH9 polymerization, protection of ACTB, blockade of DLAT-dependent biology or altered copper accumulation. The main conclusions do not depend on this interaction result.

2.6 Fig25 correlations are driven by group separation and overlap

Raw whole-sample correlations included copper homeostasis versus actin rho=0.699, copper homeostasis versus cytokinesis rho=0.832, and protein refolding versus actin rho=-0.993. After treatment-group control, removal of overlapping genes and permutation testing, the corresponding values were 0.028, 0.427 and -0.336, none with permutation q<0.05.

Fig25 therefore cannot support copper-homeostasis-to-cytoskeleton coupling. It is retained in Extended Data as a descriptive group-associated pattern only.

2.7 Candidate-priority framework: the main contribution

The literature-informed candidate universe contained 77 genes, of which 63 passed expression filtering; the complete list, its category assignment and its inclusion source are given in Supplementary Table S4 before any rank results, together with the filter status of each gene. The universe was defined before the final ranking, but it

was built from the study hypothesis and the same MYH9-related literature that supplied the external template. It is therefore not independent of the prior and is better described as a literature-informed candidate universe than as an unbiased genome-wide selection. Category counts and the proportion of each category entering the top 10 and top 20 are reported so that the effect of the class prior can be inspected directly. This audit matters because the actomyosin/actin-polymerization class contains most candidates: it contributed 8 of the 10 and 17 of the 20 highest-ranked genes in the primary model, compared with 2 of 10 and 10 of 20 in the strict response model. The membrane/adhesion/cytoskeleton class contributed 5 of 10 and 5 of 20 in the strict model. These values show that class composition and class adjustment can change the apparent priority order. MYH9 was not part of the candidate ranking. Instead, the published MYH9 polymerization mechanism was used as a fixed external template to audit how dependent other candidates were on template-like evidence. The full-template score contained four non-interchangeable evidence domains: biological prior 32%, Cu/ES response 25%, MYH9/module coupling 25% and STRING proximity 18%. These weights were prespecified expert weights, not parameters learned from data and not optimized or cross-validated. They are treated as a modelling choice rather than a measured quantity. That model only asks which genes resemble the external MYH9 template; it is not a discovery-oriented candidate ranking.

To separate template-conditioned similarity from model-based priority within the prespecified candidate universe, I report one template-referenced model and three models that do not contain MYH9 as a competing candidate:

1. Template-referenced model: which genes resemble the external MYH9 template without adding MYH9 to the candidate list?
2. Leave-MYH9-template model: MYH9-associated template similarity, STRING proximity and literature prior are removed. The only rank change reported here belongs to the MYH9-as-a-candidate sensitivity term, which moves from rank 2 to rank 48. This indicates that the score depends on the template itself rather than on an independent finding, and it is not a biological candidate rank for MYH9.
3. Template-reduced, class-adjusted model: a necessary class adjustment is retained, but external-template influence is strongly reduced; the MYH9-as-a-candidate sensitivity term remains at rank 48. This is not a pure data-only model, and the value serves only to audit template dependence.
4. Strict response model: MYH9, class, STRING, HPA and PXD priors are removed. The top 10 are GCLM, DAG1, CAVIN3, HMOX1, FLOT2, CD2AP, KIFC1, PIK3R2, RALA and VIM.

Weight and model sensitivity were checked directly. For each candidate I report the minimum, maximum and median rank across models, the Monte-Carlo rank interval, and the top-10 and top-20 inclusion frequencies, together with equal-weight, template-zero, response-only, random-weight and leave-one-domain-out specifications. Stable candidates keep a narrow rank range and a high Monte-Carlo top-10/top-20 frequency; unstable candidates move widely. For example, VIM (rank interval 3-9) and MYO18A (1-22) keep high top-20 frequency, whereas RDX (2-24), CD2AP (1-21) and PDHA1 (5-19) have wider intervals and lower frequency. Rankings are compared by rank distance and rank stability rather than combined into a single causal score.

MYL12A, MYO18A, MYL6, TLN1 and ACTB are reported as candidates that resemble the template; this means similarity to the external MYH9 template, not shared biology. CD2AP, RDX, PDHA1, FDX1, LIAS and PDHB retain relative priority in at least one template-reduced model and are reported as candidates with higher model-based priority within the prespecified candidate universe; this means a higher rank in a specified model, not data-supported mechanistic evidence. The two groups have different evidence sources and must not be merged into one list of execution factors.

2.8 MYH9-related relationships are hypotheses

Within-treatment residual analysis prioritized an ATP7A-MYH9 association, but the cross-sectional correlation cannot identify direction or direct binding. STRING and literature evidence prioritized MYH9-ANLN as an external interaction hypothesis, not a validated AML12 interaction.

To keep the main text focused, co-expression, PPI and model-implied perturbation networks involving MYH9, ACTB, MYL12A, MYO18A, FDX1 and DLAT are moved to Extended Data. They are simplified design tools, do not reconstruct a biological regulatory network and do not establish causal links. Model-edge removal percentages and virtual-perturbation values are not reported as biological effect sizes.

2.9 External data are background evidence; model inference provides design clues only

GSE9539 is human HepG2 copper dose-time transcriptomics and supports general copper-stress transportability only. GSE5348 is a chronic Atp7b-deficient mouse liver model; no candidate survived BH correction, which means only that the candidates lacked a significant transcript change in that model. HPA is human expression context with zero scoring weight. PXD037278 was removed because sample mapping could not be verified.

External proteomes, untreated AML12 ATAC background and a non-hepatic Slc31a1 perturbation system are retained in Extended Data as cross-system background for actin, microtubule and mitochondrial programs. There is no paired AML12 Cu/ES or TCAu ATAC-seq, H3K9ac ChIP-seq or protein-polymerization experiment. Chromatin accessibility, acetylation, aggregation or direct interaction changes therefore cannot be claimed.

Model-implied single- and double-gene perturbations, model-edge removal and axis projection are supplemental design analyses. A strictly linear double-perturbation model is mathematically additive and cannot be interpreted as genetic epistasis. Model-edge removal percentages are not biological effect sizes, and projection recovery is not experimental rescue.

2.10 Same-system replication search and multi-time-point, multi-omic background evidence

I systematically searched GEO and other public resources for AML12 copper-treatment datasets independent of GSE300382, using Cu/ES, CuCl2, elesclomol, TCAu, time-course and dose-series terms. No independent, publicly accessible AML12 copper transcriptome, proteome or single-cell dataset was identified. There is

therefore no true same-system replication of this study; GSE300382 remains a discovery cohort and this gap cannot be replaced by cross-system correlation.

GSE279388 provides a multi-time-point, multi-tissue genetic copper-transporter context: 76 RNA-seq samples from liver and intestinal epithelium at 16 and 30 weeks, with KO, iKO, WT and iWT genotypes related to Atp7b pathway perturbation. I performed only prespecified within-age and within-tissue contrasts and did not combine time points into a single treatment axis. The effects were context-dependent. At 30 weeks, liver KO versus WT showed higher actin, actomyosin and microtubule modules but lower mitochondrial/lipoylation scores; liver iKO versus iWT moved in the opposite direction; at 16 weeks, intestinal epithelium showed higher actin and microtubule modules but lower copper-homeostasis scores. Myh9 and Actb also changed direction across contrasts. GSE279388 therefore supports age- and tissue-dependent heterogeneity in a genetic copper-transporter model, not a Cu/ES time course, and cannot validate direction in AML12.

Protein-layer evidence came from the cross-system copper-binding or copper-probe pull-down datasets PXD076672 and PXD076786. Among 45 evaluable proteins, MYH9 was detected in four records; TLN1, MYL6, ACTB and FLNA were detected in three each; and ANLN was detected in one. These findings indicate only that some candidates enter copper-associated protein layers in other cellular systems. They do not demonstrate abundance, localization, polymerization or direct interaction in AML12. At single-cell resolution, GSE254082 provides only human Wilson-disease liver cell-type and disease context, and HPA is a human expression reference. Neither substitutes for AML12 Cu/ES single-cell or proteomic validation. These layers are shown by evidence tier in ED21 and are explicitly labeled as background.

2.11 Node-by-node evidence-chain check from copper stress to cytoskeletal transcript remodeling

To ask whether the analysis forms a closed chain from copper stress to molecular/protein changes, pathway remodeling, transcription-factor and chromatin-accessibility changes and, finally, cytoskeletal transcript changes, I prespecified six steps (N1-N6) and classified each node and link with one internally defined evidence framework, Tier 0-7 (Table 1). Tiers 0-4 cover literature background, same-study transcript association, pathway or target-set association, structural or cross-system background, and exploratory template-level computation; Tiers 5-7 require physical interaction in relevant cells, interface-specific perturbation and causal perturbation with rescue. This is an internally defined framework for organizing claims, not a field-wide standard. The audit does not support a completed regulatory chain; it defines a partially assembled candidate chain with explicit evidence gaps.

N1 is Cu/ES exposure and is treated as an assigned experimental exposure, not a causal claim. N1-to-N2 reaches Tier 1 at the same-study transcript level because the three count-based methods were directionally concordant for 15 of 17 prespecified key candidates, and 14 of those 17 candidates were significant in all three methods. Myh9 and Mt2 changed direction between methods; Atp7a was directionally concordant but did not reach significance in limma-voom. The historical FPKM comparison remains a separate cross-scale check and does not define this tier. N2 contains same-study RNA evidence plus cross-system protein context. In the primary count model,

MYH9 had no stable direction: log2FC=+0.026 (q=0.709) in DESeq2, -0.104 (q=0.200) in edgeR and -0.984 (q=1.99e-09) in limma-voom, so only limma-voom reached significance and MYH9 is reported as directionally discordant rather than as a coherent decrease. ATP7A was log2FC=+1.104 (q=2.02e-05) in DESeq2, +0.970 (q=6.78e-04) in edgeR and +0.091 (q=0.749) in limma-voom, so ATP7A was directionally concordant but not significant under limma-voom. FDX1 was negative and significant in all three count methods (-2.203, -2.330 and -3.236; q<=2.75e-23). Cross-system MYH9 protein changed by log2FC=-0.53 (FDR=0.0084); this protein observation is not a paired AML12 measurement. N2-to-N3 reaches Tier 2: proteostasis/redox modules increased (+0.82, BH=2.2e-07), whereas OXPHOS/TCA/lipoylation decreased (-1.03, BH=1.26e-04), mitochondrial stress/mtUPR decreased (-1.01, BH=1.19e-04), the F-/G-actin transcriptional proxy decreased (-0.89, BH=2.2e-07; composition-adjusted -0.96), actomyosin contraction decreased (-0.69, BH=1.99e-04; composition-adjusted -0.32), and microtubule/spindle-associated transcripts decreased (-1.07, BH=1.75e-07; composition-adjusted -0.39).

N3-to-N4 also reaches only Tier 2. GTRD target-set enrichment identified directional enrichment for RTF1 (NES=2.09, padj=3.33e-07), GADD45A (NES=1.88, padj=0.0108) and NFE2 (NES=1.75, padj=4.83e-05) in the Cu/ES response. This is regulatory resemblance, not measured transcription-factor activity, binding or causality. N4-to-N5 is weaker: the available chromatin data are limited to baseline OmniATAC in untreated AML12 cells (GSE236123, Tier 3) and ATAC/H3K9ac background from a non-hepatic Slc31a1 perturbation system (GSE337266/GSE337269, Tier 3). No Cu/ES- or TCAu-paired AML12 ATAC-seq, H3K9ac ChIP-seq or CUT&RUN is available. N4-to-N5 is therefore supported only as external background and cannot show copper-stress-induced AML12 chromatin change. Most importantly, N5-to-N6 is broken: without paired same-system chromatin measurement, transcription-factor target-set enrichment cannot be interpreted as transcription-factor- and chromatin-driven reduction of cytoskeletal transcripts. The N6 F-/G-actin, microtubule and actomyosin results are transcriptional module statistics and do not replace F-actin/G-actin protein fractionation, imaging or protein-aggregation experiments.

The evidence ceiling is therefore same-study Tier 1-2, with partial Tier 3 only in cross-system protein and chromatin background. There is no Tier 5 physical-interaction evidence and no Tier 6-7 interface or causal evidence. The audit retains the copper-stress-to-transcript/pathway chain as an analytically concordant same-system association, retains the pathway-to-TF-target-set step as inferential support, and labels the TF/chromatin-to-cytoskeletal-transcript step as a hypothesis that requires paired Cu/ES ATAC-seq or CUT&RUN, TF perturbation and protein-function experiments.

2.12 Exploratory structural context: template-level computation, supplementary

I built a small structural material only to decide whether any MYH9-related interface is worth experimental testing, and it is kept as background rather than a main result. First, I generated 10 AlphaFold3/ColabFold v2 multimer input specifications in `reference/external_refs/af3_job_specs/`, each with a planned apo variant and a planned explicit-copper variant. Neither variant was executed, because no model weights, no

licensed server and insufficient GPU memory were available; every file is labelled `specification_only_not_executed`, and no predicted coordinates and no copper-conditioned structure exist. A planned copper-conditioned input would have to carry an explicit CU entity plus constraints on curated Cys/His/Met ligands, because AlphaFold3 does not place copper by coordination chemistry on its own. Second, because prediction could not be run, three experimentally resolved homolog or ortholog complexes were simulated in GBn2 implicit solvent over 0.30 ns each, with no copper ion: PDB 1S5G (myosin heavy chain with regulatory and essential light chains), PDB 5JLH (a human cytoplasmic actomyosin complex whose motor chain is the MYH14 orthologue, not MYH9) and PDB 3DAW (an actin-ADF-H complex used only as an unrelated actin-binding structural control).

Each trajectory is short and single-seed, so the values describe gross short-timescale behavior of the templates and cannot estimate whether those interfaces would be stable in cells, nor their affinity or copper dependence. The 5JLH template produced an anomalous radius-of-gyration estimate. Because the start and end values were similar and the anomaly is more consistent with template preparation or the choice of the Rg object than with a real conformational change, I report it only as a preparation and sampling caveat and do not interpret it as interface swelling or weakening. All numerical trajectory metrics are given in Extended Data ED25.

The structural layer therefore remains at Tier 3 for input specifications and cross-system background, and at most reaches Tier 4 for exploratory template-level computation. 1S5G and 3DAW did not disassemble on this short window, and 5JLH could not be scored reliably, which at most identifies interfaces that could be examined biochemically. MYH9-ATP7A, MYH9-ANLN and MYH9-MYO18A have no complex of record and no completed simulation; ATP7A-MYH9 is only a direction-agnostic transcript association and MYH9-ANLN only a database-supported interaction hypothesis. No result in this section is evidence of copper binding, polymerization, acetylation, direct interaction or death execution, and no system here contained copper.

3. Discussion

Seven conclusions are supported. First, Cu/ES is associated with broad transcriptome remodeling in AML12 cells, and count-level concordance is the main robustness evidence. Second, proteostasis and redox-defense programs increase. Third, mitochondrial, cell-cycle and cytoskeletal transcripts decrease, but this pattern is strongly affected by composition and survival selection. Fourth, TCAu interaction results are method-sensitive and are retained only as a secondary technical sensitivity. Fifth, candidate ranking depends strongly on the external MYH9 template; MYH9 is a template-dependent benchmark, not an independently discovered candidate. Sixth, step-by-step auditing supports a same-study copper-stress-to-transcript-and-pathway association and inferential TF-target enrichment, but the chromatin-to-cytoskeletal-transcript link remains unmeasured. Seventh, the data do not identify a death mode or validate a cuproptosis executioner.

The principal methodological contribution of this study is the separation of template-conditioned similarity from model-based ranking. MYH9 was used as an external reference mechanism and was excluded from the candidate

table. Removing MYH9-related prior terms changed the apparent position of the MYH9-as-a-candidate sensitivity diagnostic, showing that a high template-conditioned score cannot be interpreted as an independent discovery. Conversely, candidates that remained relatively prioritized after selected external evidence domains were removed should be understood only as model-based priorities within a prespecified candidate universe. These rankings identify experimental order, not molecular function that has been measured in AML12. The framework is reusable for other small-sample reanalyses in which a mechanism prior could otherwise be mistaken for a discovery.

The structural calculations should be read as exploratory template-level context only. AlphaFold3 was not executed, no copper-containing system was simulated, and the completed trajectories used homologous or orthologous complexes over a short, single-seed window. These analyses therefore do not establish the existence, affinity, copper dependence or intracellular relevance of any candidate complex in AML12. Their only appropriate use is to identify interfaces that might be considered for later biochemical testing, and they are kept in Extended Data rather than treated as a main result.

The evidence-chain check further limits interpretation. Cu/ES, differential expression, pathway modules and TF target-set enrichment can be linked as same-study or inferential associations, but a chromatin-accessibility-to-cytoskeletal-transcription mechanism has no paired AML12 Cu/ES chromatin measurement and cannot be treated as established. The regulatory axis can be upgraded only after paired Cu/ES ATAC-seq or CUT&RUN, TF perturbation, protein fractionation/insolubility assays and candidate-protein functional experiments are completed.

Biologically, the most defensible hypothesis is that AML12 sensitivity to Cu/ES may involve mitochondrial copper use, respiratory dependence, FDX1-LIAS-associated lipoylation, copper buffering and proteostasis reserve. Actin-myosin and adhesion changes may represent cell-state transition or a secondary structural event. The current transcriptome cannot distinguish these explanations. MYH9 remains a useful experimental anchor, but protein-level and functional work is required.

4. Limitations

1. The public matrix is FPKM-based and is used only for historical comparability. The primary inference framework is the count-level DESeq2 model, with edgeR and limma-voom confirmation; these analyses do not replace biological validation.
2. There are only three samples per group. Interaction, co-expression, network and ranking analyses are exploratory.
3. Copper-containing samples show major changes in detected-gene number, ribosomal fraction and cell state. Composition adjustment may remove both technical bias and treatment-mediated effects; adjusted genes should not be treated as absolute truth.

4. No independent AML12 copper-treatment dataset was found, so no same-system replication is available. GSE279388 contains 76 samples and spans 16 and 30 weeks but is a chronic genetic Atp7b-related model; candidate direction changes with tissue, age and genotype. PXD076672, PXD076786 and GSE254082 come from other cellular systems or human Wilson-disease liver and provide only protein-layer or single-cell background.
5. The Cu/ES-to-molecular/protein-to-pathway-to-TF/chromatin-to-cytoskeletal-transcript regulatory chain is not closed. The transcript and pathway layers are same-study associations, and the TF layer is target-set enrichment rather than measured TF activity. AML12 has no paired Cu/ES ATAC-seq, H3K9ac ChIP-seq or CUT&RUN; GSE236123 is an untreated background and GSE337266/GSE337269 come from a non-hepatic Slc31a1 perturbation system. TF enrichment therefore cannot be interpreted as copper-stress-driven TF and chromatin remodeling of cytoskeletal transcripts.
6. No new protein or functional experiment was performed. Copper binding, MYH9 polymerization/insolubility, MYH9 acetylation, DLAT lipoylation/release, F-actin/G-actin, subcellular localization, cellular copper, genetic perturbation and rescue remain untested.
7. "Candidate structural factor" remains a computational priority, not a validated functional role; F-/G-actin, actomyosin and microtubule modules are transcriptional proxies rather than protein fractionation or cytoskeletal imaging.
8. Clinical statements provide background only and do not support diagnosis, prognosis or treatment.

5. Conclusion

The raw-count primary reanalysis of GSE300382 shows that Cu/ES is associated with broad proteostatic, redox, mitochondrial, cell-cycle and cytoskeletal gene-expression changes in AML12 hepatocytes. DESeq2 is the primary inference framework, edgeR and limma-voom provide count-level confirmation, and the FPKM-limma analysis is retained only for historical comparability with the public matrix. The TCAu interaction is reported only as a secondary technical sensitivity. Detected depth, ribosomal fraction and cell state change substantially under copper-containing conditions; cytoskeletal and mitotic transcript reduction cannot be interpreted as specific structural-gene inhibition. The MYH9 polymerization mechanism is used only as an external template for conditional candidate ranking and is not part of the candidate ranking. MYH9 is therefore a template-dependent benchmark, not an independently discovered candidate. CD2AP, RDX, PDHA1, FDX1, LIAS and PDHB retain relative priority in some models with evidence removed and should be called candidates with higher model-based priority rather than data-supported mechanisms; they belong to different mechanistic layers and cannot be collectively called execution factors. ATP7A-MYH9 is a direction-agnostic transcript association, and MYH9-ANLN is an external interaction hypothesis. The evidence chain is supported at the transcript and pathway levels, is inferential at the TF-target level, and remains unmeasured at the paired AML12 chromatin and cytoskeletal

protein-function levels. This study does not show copper binding, polymerization, acetylation, direct interaction or cell-death execution for any candidate in AML12.

6. Methods

6.1 Data source

GSE300382 contains four AML12 groups with three replicates each: control, CuCl2 300 uM plus elesclomol 40 nM, Cu/ES plus TCAu 1000 uM, and TCAu 1000 uM, all at 24 h. The 12 associated SRA samples were reprocessed into integer counts and used for primary inference; the public FPKM matrix was retained only as a historical comparison.

6.2 Historical FPKM matrix comparison

Genes with FPKM>=1 in at least three samples were retained (11,554 genes). Values were transformed as log2(FPKM+1) and analyzed with limma and robust trended empirical Bayes moderation. Differential expression required BH FDR<0.05 and |log2FC|>=1. The factorial model used copper*tca_u, with interaction (Cu/ES+TCAu - TCAu) - (Cu/ES - Control). All five prespecified comparisons were reported for historical comparability with the public matrix. These results are not used as the primary statistical basis.

6.3 Raw-count primary analysis

All 12 SRA runs were uniformly processed against GRCm39. Reads were aligned with Rsubread::align (`type="rna"`, `phredOffset=33`, `nSubreads=10`, `TH1=3`, `TH2=1`, `maxMismatches=3`, `unique=FALSE`, `nBestLocations=1`, `indels=5`, paired-end orientation `"fr"`) and summarized with Rsubread::featureCounts against the GRCm39 Ensembl 115 GTF (`allowMultiOverlap=FALSE`, `countMultiMappingReads=FALSE`, `fraction=FALSE`, `primaryOnly=TRUE`, `ignoreDup=FALSE`, `strandSpecific=0`, paired-end `countReadPairs=TRUE`, `requireBothEndsMapped=TRUE`). Rsubread version 2.20.0 was recorded in the run logs. The same filtered count matrix (genes with count>=10 in at least three samples; 19,275 genes) was the input for all three count methods. The primary model was DESeq2 with the design `~copper*tca_u`, built with `DESeqDataSetFromMatrix`, default size-factor estimation and the standard `DESeq()` workflow; `results()` was used with default independent filtering and no lfcShrink, and a gene was called differentially expressed at BH FDR<0.05 and |log2FC|>=1. edgeR and limma-voom were run on the same matrix as count-level confirmation. Factor encoding was explicit:

Group	copper	tca_u
Control	0	0
Cu/ES	1	0
TCAu	0	1
Cu/ES+TCAu	1	1

Contrasts were defined as: Cu/ES simple effect = Cu/ES - Control; Cu/ES simple effect under TCAu = Cu/ES+TCAu - TCAu; TCAu interaction = (Cu/ES+TCAu - TCAu) - (Cu/ES - Control). TCAu_at_CuES and CuES_at_TCAu0 were reported as additional pairwise contrasts. No independent library-preparation or sequencing batch variable was available in the public metadata, so no batch covariate was included in the primary design, and this limitation was considered when interpreting the small-sample factorial results. Count results were compared with FPKM-limma by Pearson and Spearman correlation, direction concordance, top-gene overlap and GSEA NES correlation, using FPKM-limma only as a historical matrix comparison. For each comparison, the direction-concordance denominator was the shared set of genes meeting the same prespecified threshold in both methods being compared. Preranked GSEA used the sign(logFC)x-log10(p) statistic with fgsea, gene-set size 10-500 and 10,000 simple permutations. Key genes included MYH9, MYL12A, MYL6, MYO18A, ACTB, TLN1, CD2AP, RDX, ATP7A, FDX1, LIAS, DLAT, PDHA1 and PDHB. TCAu_at_CuES was separated as a method-sensitive technical sensitivity.

6.4 Composition audit

Two composition families were computed and kept separate. In the public FPKM matrix I calculated per-sample FPKM-detected genes (FPKM>=1), ribosomal FPKM fraction, cell-cycle score, oxidative-stress score and housekeeping dispersion. The cell-cycle and oxidative-stress scores were the column means of gene-level z-scores over the GO "cell cycle" and "response to oxidative stress" sets, respectively. Housekeeping dispersion was the per-sample standard deviation of log2(FPKM+1) across a fixed ten-gene housekeeping set, not a within-group coefficient of variation, and the mitochondrial fraction was unavailable because the public matrix contains no mt-transcripts. In the count matrix I calculated library size, DESeq2 size factors, edgeR TMM factors, genes with count>0, genes with count>=10, and the ribosomal read fraction. Because the main analysis is count-based, the count-level ribosomal fraction and count-detected-gene number are the composition metrics referenced in the text where they disagree with the FPKM-derived values. Raw and composition-adjusted interaction models were compared by effect correlation, direction concordance, effect range and FDR change. Leave-one-sample-out analysis assessed interaction stability.

6.5 Candidate scoring and stepwise removal of evidence

The 77-gene candidate universe was defined before final ranking across copper handling, metabolic execution, structural substrate, complex scaffold, antioxidant protection and cell-cycle-associated categories; the complete list, category labels, inclusion source and filter status are in Supplementary Table S4. The universe was informed by the study hypothesis and MYH9-related literature and is therefore not independent of that prior. MYH9 was used only as the fixed external template and does not appear in the candidate ranking excluding MYH9. The full-template model used four evidence domains with weights of 32% (biological prior), 25% (Cu/ES response), 25%

(MYH9/module coupling) and 18% (STRING proximity). These are prespecified expert weights, not parameters learned from the data, not optimized, and not the output of a fitted or cross-validated model. Class priors can favour structural candidates, and the candidate universe is unbalanced across categories; category-level counts and top-10/top-20 representation are therefore reported with the ranks. Reported models include the template-referenced model, leave-MYH9-template, template-reduced class-adjusted, no-literature-prior, no-STRING, no-HPA, no-external-proteomics, random-prior, negative-template and strict response models. Weight and model sensitivity were assessed with equal-domain weights, template weight set to zero, a response-only specification, random weights and leave-one-domain-out. For each candidate I report the minimum, maximum and median rank across models, the rank range and variance, the Monte-Carlo rank interval, and the top-10 and top-20 inclusion frequencies (Supplementary Table S4). The template-reduced class-adjusted model remains a class prior and is therefore not a pure data-only model. Models use different evidence domains; only ranks and rank stability are compared, and no cross-domain sum is reported as a causal score. Model-based priority candidates are defined by models that do not contain MYH9 as a competing candidate, and a higher rank means only higher model-based priority within the prespecified candidate universe; it does not indicate validated molecular function.

6.6 External background evidence

GSE9539, GSE5348, GSE279388, GSE254082, IntAct, PXD076672, PXD076786, GSE236123, GSE337266 and GSE337269 were analyzed as tiered external context. The 76 GSE279388 samples were matched by tissue, age and genotype; only within-age contrasts meeting prespecified sample counts were used, and module directions were reported separately rather than combined into a single cross-system summary score. GSE254082 provided human Wilson-disease liver single-cell context, and PXD076672/PXD076786 provided cross-system protein-layer detection. HPA was used only as human expression background. PXD037278 was excluded because sample mapping could not be verified.

6.7 Interpretation boundary

Model-implied perturbations, model-edge removal and axis projection are design analyses, not causal experiments. Death-mode gene-set scores describe transcriptional resemblance and cannot determine the actual death mode. No new wet-lab validation was performed.

6.8 Evidence-chain audit

To evaluate whether copper stress to molecular/protein changes, pathway remodeling, transcription-factor/chromatin accessibility and cytoskeletal transcript changes forms a complete logic chain, I prespecified six steps: N1 (Cu/ES exposure), N2 (molecular/protein layer), N3 (pathway remodeling), N4 (transcription-factor regulatory inference), N5 (chromatin accessibility) and N6 (cytoskeletal transcripts). Each step and adjacent link was classified with a single internally defined Tier 0-7 framework (Table 1). Tier 0 is literature or database background; Tier 1, same-study transcript association; Tier 2, pathway or target-set association; Tier 3, structural

prediction/input or cross-system multi-omics background; Tier 4, exploratory template-level computational stability; Tier 5, physical interaction in relevant cells; Tier 6, interface-specific perturbation; and Tier 7, causal perturbation with rescue. No other ladder is used in this manuscript, and the framework is presented as a way to organize claims rather than as a field-wide standard.

The statistical basis includes count-level interaction concordance, shared-DEG direction concordance, BH-corrected GO biological-process GSEA, prespecified module scores, composition-bias audits, preranked GTRD transcription-factor target-set enrichment and tiered external chromatin/protein evidence. KEGG pathway enrichment was not performed, and no KEGG-derived statement is made anywhere in this manuscript. N1-to-N2, N2-to-N3 and N3-to-N4 were evaluated as same-study transcript or pathway associations. N4-to-N5 was retained only as external background. N5-to-N6 was classified as still missing because no paired AML12 Cu/ES ATAC-seq, H3K9ac ChIP-seq or CUT&RUN exists. The audit itself did not use co-immunoprecipitation, AP-MS, proximity labelling, colocalization, genetic perturbation or rescue data, and no causal status was assigned to any regulatory link; the separate structural specification and template-level simulation layer is described in 6.9, where AlphaFold3 prediction was not executed.

6.9 Structural specifications and template-level molecular dynamics

To ask whether any MYH9-related candidate interface might be worth experimental testing, I generated 10 AlphaFold3/ColabFold v2 multimer input specifications under `reference/external_refs/af3_job_specs/`, covering MYH9-ACTB, the MYH9 actin-binding peptide against ACTB, MYH9-MYL12A, MYH9-MYL6, MYH9-MYO18A, MYH9-ANLN, MYH9-ATP7A and ATP7A-ANLN, plus DLAT-PDHA1 and FDX1-LIAS as reference pairs representing the canonical lipoylation-associated axis. Each specification carries a planned apo variant and a planned explicit-copper (`CU`) variant, with domain boundaries taken from UniProt and InterPro annotations (Myh9 motor 81-776, actin-binding 654-676, IQ 779-808 and coiled-coil 841-1926; Atp7a HMA1-HMA7; Dlat lipoyl 90-166 and 217-293; full-length Pdha1 and Pdhb). A copper-conditioned input would have to carry an explicit CU entity plus constraints on the curated Cys/His/Met ligands, because AlphaFold3 does not add copper automatically; without those constraints the ion is placed by network preference rather than by coordination chemistry. No AlphaFold3 weights, no licensed server and only a 4 GB GPU were available here, so every specification is labelled `specification_only_not_executed`; neither the apo nor the copper variant was executed and no predicted coordinates exist.

Because prediction could not be executed, three experimentally resolved homolog or ortholog complexes were simulated over 0.30 ns each: PDB 1S5G (scallop myosin S1 heavy chain with regulatory and essential light chains, a template for the MYH9-homolog light-chain interface), PDB 5JLH (a near-atomic-resolution cryo-EM human cytoplasmic actomyosin complex whose motor chain is the MYH14 orthologue, not MYH9 itself) and PDB 3DAW (an actin-ADF-H complex used only as an unrelated actin-binding structural control). After PDBFixer

completed missing heavy atoms and hydrogens were added at pH 7.0, systems were built with `amber14-all.xml` and GBn2 implicit solvent: a 1.6 nm non-periodic cutoff, HBonds constraints, 3.0 amu hydrogen-mass repartitioning, a 2 fs timestep and a 300 K Langevin-middle integrator. No copper ion was placed in any system, because generic AMBER14 ion parameters are not a validated Cu(I)/Cu(II)-protein force field; the trajectories therefore describe apo template behavior only. Twenty uniformly spaced production frames were analysed for native-contact retention Q (the fraction of inter-chain heavy-atom contacts within 0.45 nm of the starting structure that persist in a later frame), interface backbone RMSD, radius of gyration and minimum inter-chain heavy-atom distance.

The statistical basis is that the native-contact set is defined once on the starting structure and held fixed across all frames, so Q is a fixed ratio against a single reference rather than a per-frame refit score; interface backbone RMSD uses Kabsch superposition to remove global translation and rotation; and all metrics are reported as means or extremes over the same prespecified 20 production frames. The design's strengths are a fixed reference set and statistic, a fully recorded force field and simulation protocol, and inclusion of an unrelated actin-binding control, which prevents any actin-binding interface from being read as MYH9-specific evidence. Its limitations are the absence of copper, a single force field, implicit solvent, a single seed and short-timescale sampling, so this layer can support Tier 3 input specification and, at most, Tier 4 exploratory template computation, but cannot replace binding-affinity measurement, co-immunoprecipitation, interface mutation or functional perturbation. The 5JLH radius of gyration was flagged as a template-preparation or Rg-object caveat and was not interpreted as interfacial swelling.

7. Data and code availability

Data availability. No new experimental data were generated in this study. All datasets analysed here were previously deposited in public repositories and are available without restriction. The datasets analysed during this study are available in the NCBI GEO, NCBI SRA, PRIDE and RCSB PDB repositories. Their final accession numbers and public links are listed below and in `DATA_SOURCES.csv`.

The main RNA-seq dataset was NCBI GEO **GSE300382** (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), with associated NCBI SRA runs **SRR34082235-SRR34082246** (https://www.ncbi.nlm.nih.gov/sra?term=SRR34082235). The public FPKM matrix, sample metadata and series metadata used for the historical comparison were downloaded from the same GEO record. The processed integer count matrix, alignment and composition quality-control outputs, differential-expression results and all downstream analysis tables are included with the code archive in this submission.

External public data used in the secondary analyses were NCBI GEO **GSE9539** (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9539), **GSE5348** (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE5348), **GSE279388** (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE279388), **GSE254082**

(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE254082), **GSE236123** (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE236123), **GSE337266** (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE337266) and **GSE337269** (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE337269); PRIDE **PXD076672** (https://www.ebi.ac.uk/pride/archive/projects/PXD076672) and **PXD076786** (https://www.ebi.ac.uk/pride/archive/projects/PXD076786); STRING (https://string-db.org/); IntAct (https://www.ebi.ac.uk/intact/); Human Protein Atlas (https://www.proteinatlas.org/); GTRD (https://gtrd.biouml.org/) and MSigDB (https://www.gsea-msigdb.org/gsea/msigdb); Gene Ontology (https://geneontology.org/) with the Bioconductor annotation package **org.Mm.eg.db 3.20.0** (https://bioconductor.org/packages/org.Mm.eg.db/); UniProtKB (https://www.uniprot.org/); InterPro (https://www.ebi.ac.uk/interpro/); and RCSB PDB structures **1S5G** (https://www.rcsb.org/structure/1S5G), **5JLH** (https://www.rcsb.org/structure/5JLH) and **3DAW** (https://www.rcsb.org/structure/3DAW). OpenMM, PDBFixer and AlphaFold3 are software tools, not data sources; AlphaFold3 was not executed.

PXD037278 was excluded from every analysis and figure because its raw-file-to-condition mapping could not be verified. Large public raw FASTQ files and the mouse reference genome were not duplicated in the submission package; their complete accessions and public links are listed here and in `DATA_SOURCES.csv`. The package also contains the processed count and FPKM matrices, sample metadata, all result tables used for the figures, figure files, figure-generation and analysis code, and a source manifest.

Code availability. At the time of writing, no public repository release, version tag or commit ID, no one-command workflow and no locked environment file had been prepared. A machine-readable molecular-dynamics provenance record (`md_provenance.json`) could not be located during the present audit. The analysis is therefore described as **computationally auditable** rather than computationally reproducible. The software versions and key parameters used for the count-level, enrichment and composition analyses are recorded in the Methods; any claim of full computational reproducibility would require the public repository, version tag, provenance file and locked environment files that are still missing.

8. References

1. Tsvetkov P, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. Science. PMID 35298263.
2. Tang D, Chen X, Kroemer G. Cuproptosis: a copper-triggered modality of mitochondrial cell death. Cell Research. PMID 35322203.
3. Cobine PA, Brady DC. Cuproptosis: cellular and molecular mechanisms underlying copper-induced cell death. Molecular Cell. PMID 35594854.
4. Chen L, et al. Copper homeostasis and cuproptosis in health and disease. Signal Transduction and Targeted Therapy. PMID 36414625.

5. Ge EJ, et al. Targeting cuproplasia and cuproptosis in cancer. Nature Reviews Clinical Oncology. PMID 38486054.

6. Guo Y, et al. The original GSE300382 study on taurocholic acid-enhanced cuproptosis in hepatocytes. Cell Death Discovery. 2026;12(1):12. PMID 41513631.

7. Wu LZ, Mi XY, Liu ZJ, et al. Copper ion-induced MYH9 polymerization executes cuproptosis. Cell. 2026. PMID 42735689.

8. Non-muscle myosin II in disease: mechanisms and therapeutic opportunities. Disease Models and Mechanisms. PMID 26542704.

9. Linking MYH9-related diseases to non-muscle myosin II-A function in cells. Cells. PMID 32545517.

10. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. PMID 25605792.

11. Love MI, et al. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014. PMID 25516281.

12. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-140. PMID 19910308.

13. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Research. 2012;40(10):4288-4297. PMID 22287627.

14. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biology. 2014;15(2):R29. PMID 24485249.

15. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. PMID 19114008.

16. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics. 2006;7:91. PMID 16504092.

17. Marbach D, et al. Wisdom of crowds for robust gene network inference. Nature Methods. 2012;9(8):796-804. PMID 22796662.

18. Wilson disease and ATP7B-dependent hepatic copper overload: clinical genetics and management. PMID 30190489.

19. Copper chelation and tetrathiomolybdate in clinical copper-lowering strategies. PMID 38081365.

20. Disulfiram plus copper in advanced solid tumours: phase 1b clinical evaluation. PMID 41699560.

21. GEO dataset GSE279388. Multi-time-point liver and intestinal epithelial transcriptomes in a mouse copper-transporter model. NCBI Gene Expression Omnibus. Accessed 2026-10-02.

22. Human Wilson-disease liver single-cell transcriptomic study. PMID 39854622. Dataset: GSE254082.

23. Abramson J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630(8016):493-500. PMID 38718835.

24. Eastman P, et al. OpenMM 7: rapid development of high performance algorithms for molecular dynamics. PLoS Computational Biology. 2017;13(7):e1005659. PMID 28278271.

25. Nguyen H, Roe DR, Simmerling C. Improved generalized Born solvent model parameters for protein simulations. Journal of Computational Chemistry. 2013;34(9):770-779. PMID 23339131.

26. Risal D, Gourinath S, Himmel DM, Szent-Gyorgyi AG, Cohen C. Myosin subfragment 1 structures reveal a partially bound nucleotide and a complex salt bridge that helps couple nucleotide and actin binding. PNAS. 2004;101(24):8930-8935. PMID 15184651. Structure: PDB 1S5G.

27. von der Ecken J, Heissler SM, Pathan-Chhatbar S, Manstein DJ, Raunser S. Cryo-EM structure of a human cytoplasmic actomyosin complex at near-atomic resolution. Nature. 2016;534(7609):724-728. PMID 27324845. Structure: PDB 5JLH.

28. Paavilainen VO, Oksanen E, Goldman A, Lappalainen P. Structure of the actin-depolymerizing factor homology domain in complex with actin. Journal of Cell Biology. 2008;182(1):51-59. PMID 18625842. Structure: PDB 3DAW.

9. Figure legends

Figure 1. Study design and global transcriptome. Four-group design, sample correlation, PCA and global expression structure. Treatment is the main source of transcriptomic separation, and composition differences are visible in copper-containing samples.

Data source: NCBI GEO GSE300382 (FPKM matrix and sample metadata; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382). No new data were generated for this figure.

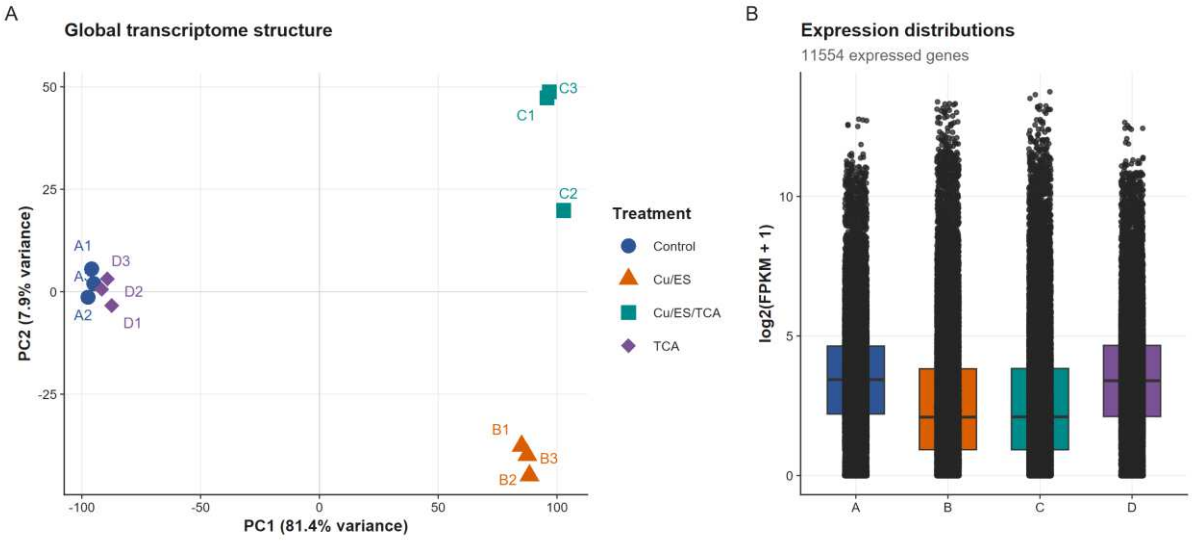


Figure 1. Study design and global transcriptome.

Treatment is the dominant source of global transcriptomic separation, and copper-containing groups also show marked compositional changes. The four groups separate clearly in correlation and PCA structures, indicating that Cu/ES is associated with broad cell-state separation in this dataset rather than isolated gene-level fluctuations. The figure supports a state-separation pattern but cannot determine a death mechanism or establish causality.

Figure 2. Primary raw-count differential expression. A, differential-gene counts for the prespecified contrasts in DESeq2, edgeR and limma-voom; B, DESeq2 volcano plot for Cu/ES versus control with selected structural, copper-handling, mitochondrial and stress genes labelled; C, direction and significance of prespecified key genes in the three count models. TCAu alone has minimal transcriptome impact. FPKM-limma is not used in this main panel and is shown only in ED26 for historical comparison with the public matrix.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382) and NCBI SRA runs SRR34082235-SRR34082246 (https://www.ncbi.nlm.nih.gov/sra?term=SRR34082235); reads were reprocessed into integer counts. No new data were generated for this figure.

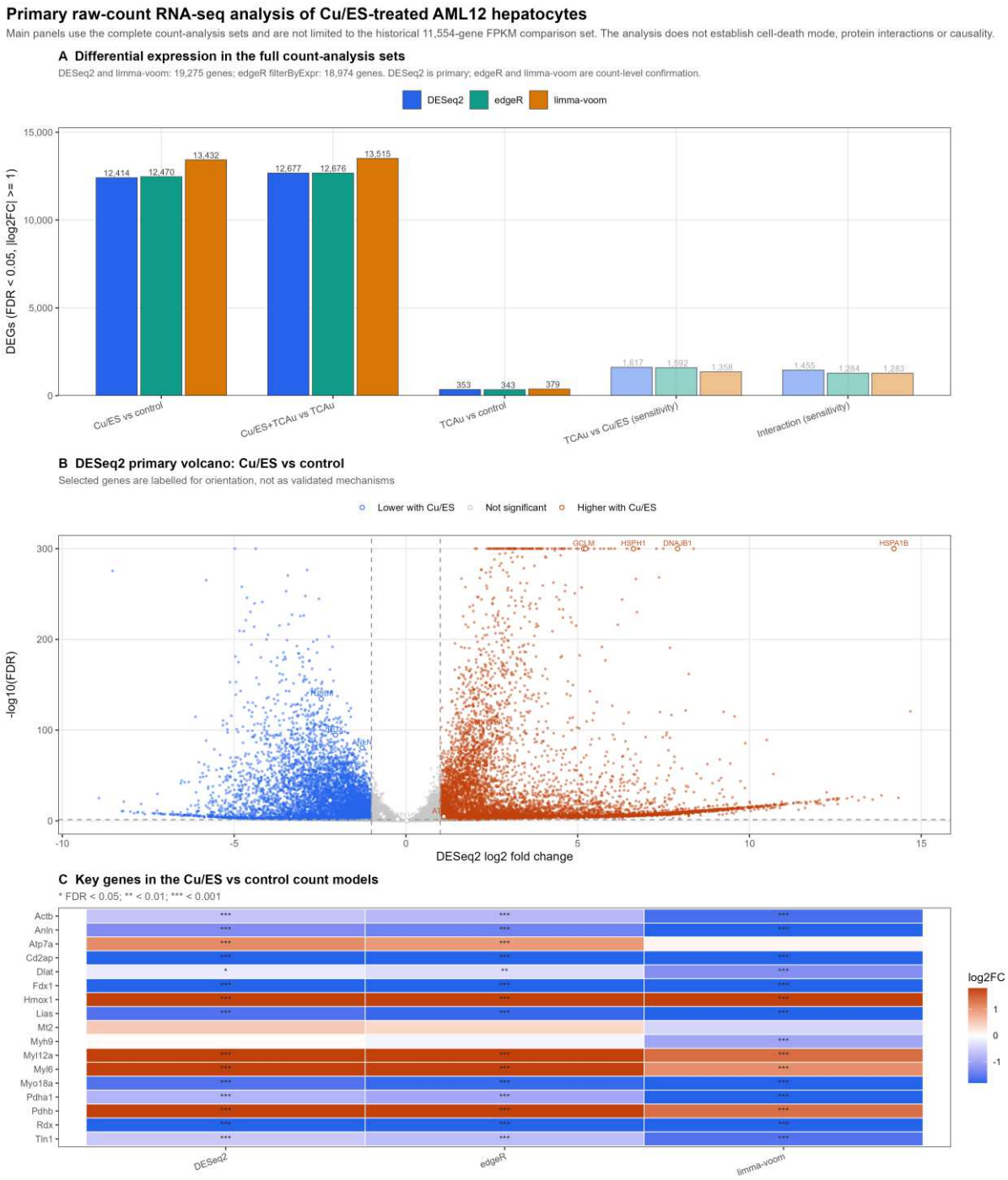


Figure 2. Primary raw-count differential expression.

The count-level analyses consistently identify a broad Cu/ES-associated transcriptional state change rather than an isolated gene response. DESeq2 identified 12,414 and 12,677 genes for Cu/ES versus control and Cu/ES+TCAu versus TCAu, with edgeR and limma-voom giving a near-identical scale (12,470-13,515 genes), whereas TCAu alone versus control changed 353 genes in DESeq2, 343 in edgeR and 379 in limma-voom. The volcano plot supports this scale of remodeling, and the key-gene panel places actin-myosin, adhesion, cell-cycle, mitochondrial and proteostatic genes within the same response: 15 of 17 prespecified key genes agreed in

direction across the three count-based methods, 14 were significant in all three, and Myh9 and Mt2 changed direction between methods. FDX1 was negative and significant in all three methods, while Atp7a was directionally concordant but not significant under limma-voom. This pattern is most consistent with a shift toward proteostatic and redox-adaptive programs accompanied by reduced mitochondrial, cell-cycle and cytoskeletal transcriptional programs; however, count-level differential expression alone cannot distinguish direct suppression from surviving-cell selection or changes in cell composition.

Figure 3. Functional program remodeling by full-count GO biological-process GSEA. A, running-enrichment structure for the Cu/ES versus control contrast; B, representative GO biological-process terms spanning ribosome/translation, chaperone/proteostasis, actin/actomyosin, microtubule, cytokinesis, cell cycle, oxidative phosphorylation/TCA, fatty-acid and copper-homeostasis themes (x, normalized enrichment score, NES; colour, BH FDR below or above 0.05; shape, contrast; size, -log10 FDR). Interpretation must be read with the analytical-concordance audit in Figure 4 and the composition audit in Figure 5.

Data source: NCBI GEO GSE300382 (integer counts; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), Gene Ontology biological-process annotations (https://geneontology.org/) and the Bioconductor annotation package org.Mm.eg.db 3.20.0 (https://bioconductor.org/packages/org.Mm.eg.db/). No new data were generated for this figure.

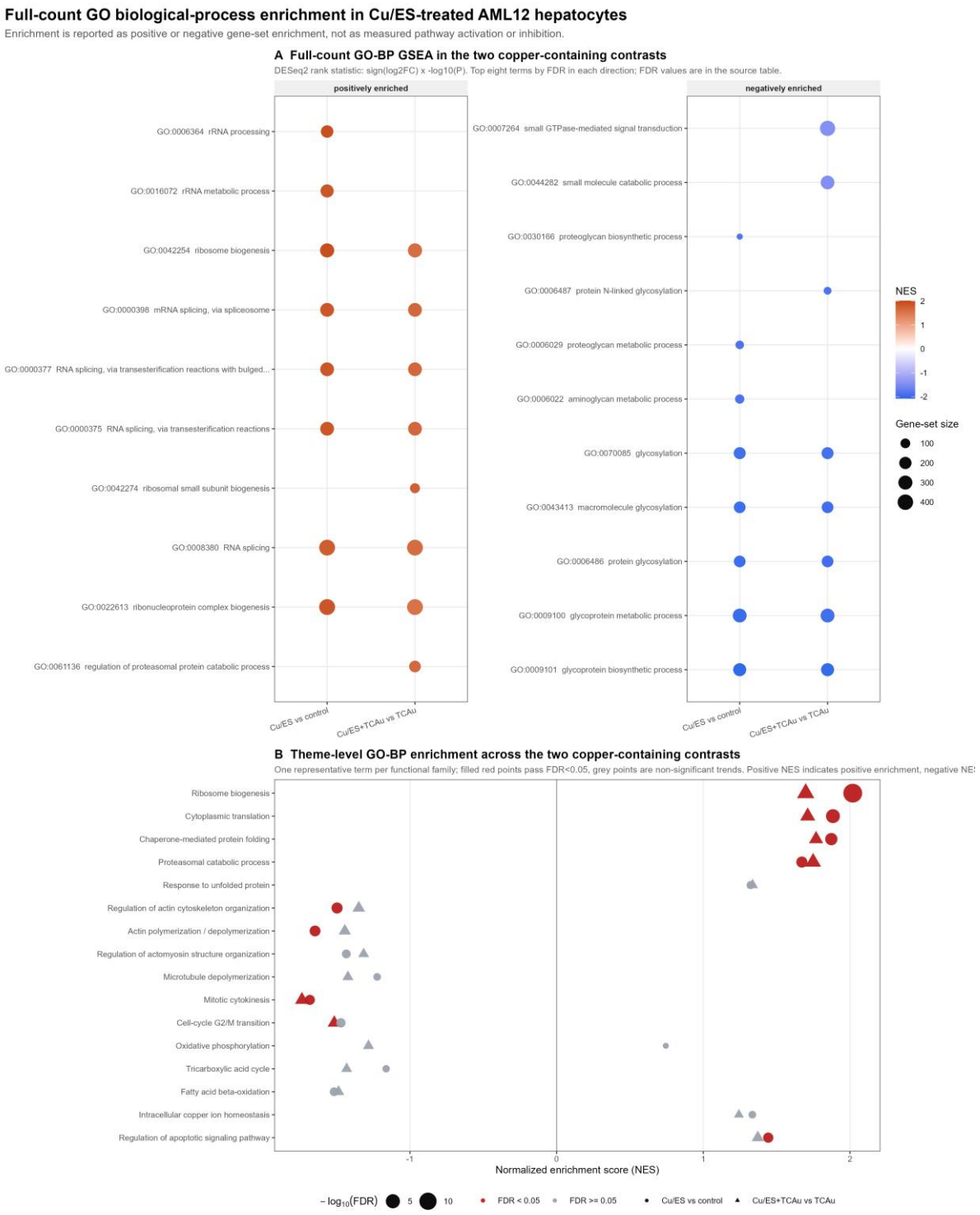


Figure 3. Functional program remodeling by full-count GO biological-process GSEA.

The full-count enrichment analysis places the Cu/ES cell state at positively enriched translation/proteostasis and negatively enriched cytoskeletal and cell-cycle programs, while the mitochondrial and lipid terms are negative but not significant. Positively enriched and FDR-significant terms include ribosome biogenesis (NES=2.02, FDR=1.6e-15), cytoplasmic translation (NES=1.88, FDR=2.0e-06), chaperone-mediated protein folding (NES=1.87, FDR=2.5e-04) and regulation of the proteasomal catabolic

process (NES=1.67, FDR=3.8e-03). Negatively enriched and FDR-significant terms include regulation of actin-cytoskeleton organization (NES=-1.50, FDR=4.2e-03), actin polymerization or depolymerization (NES=-1.65, FDR=6.7e-03) and mitotic cytokinesis (NES=-1.68, FDR=3.1e-02), with cytoskeleton-dependent cytokinesis (NES=-1.63, FDR=4.9e-02) and cell-cycle G2/M transition (NES=-1.52, FDR=4.7e-02) significant in the Cu/ES+TCAu contrast. Oxidative phosphorylation, respiratory electron transport, the TCA cycle and fatty-acid beta-oxidation were negatively oriented but did not reach FDR < 0.05 in this full-count set, and copper-ion homeostasis was not significantly enriched; those terms are therefore reported as negative trends only, and the stronger mitochondrial and lipoylation module signals come from the historical FPKM comparator (Methods Section 6.2 and ED27). Together with Hspa1b, Hsph1, Dnajb1, Gclm and Tpm2 in Figure 2, the significant terms suggest that Cu/ES is accompanied by increased proteostatic and redox-adaptive programs and reduced cytoskeletal and cell-cycle transcriptional programs, with actin-cytoskeleton terms pointing to a state in which actomyosin organization and cytokinesis transcripts are down rather than up. Because the same 12 samples are used and survival selection and composition change are possible, these are pathway-associated transcriptional patterns and not measured pathway activity; the actin enrichment cannot be read as specific inhibition of structural genes.

Figure 4. Analytical concordance among the three count-based RNA-seq workflows. Interaction effects, shared-DEG directions, top-gene overlap and GSEA correlations for DESeq2, edgeR and limma-voom on the same 12 samples. This is an analytical-consistency analysis, not biological replication and not mechanistic validation; the FPKM-limma comparator is reported separately in ED26 and ED27.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382); DESeq2, edgeR and limma-voom were applied to the same reprocessed integer counts. No new data were generated for this figure.

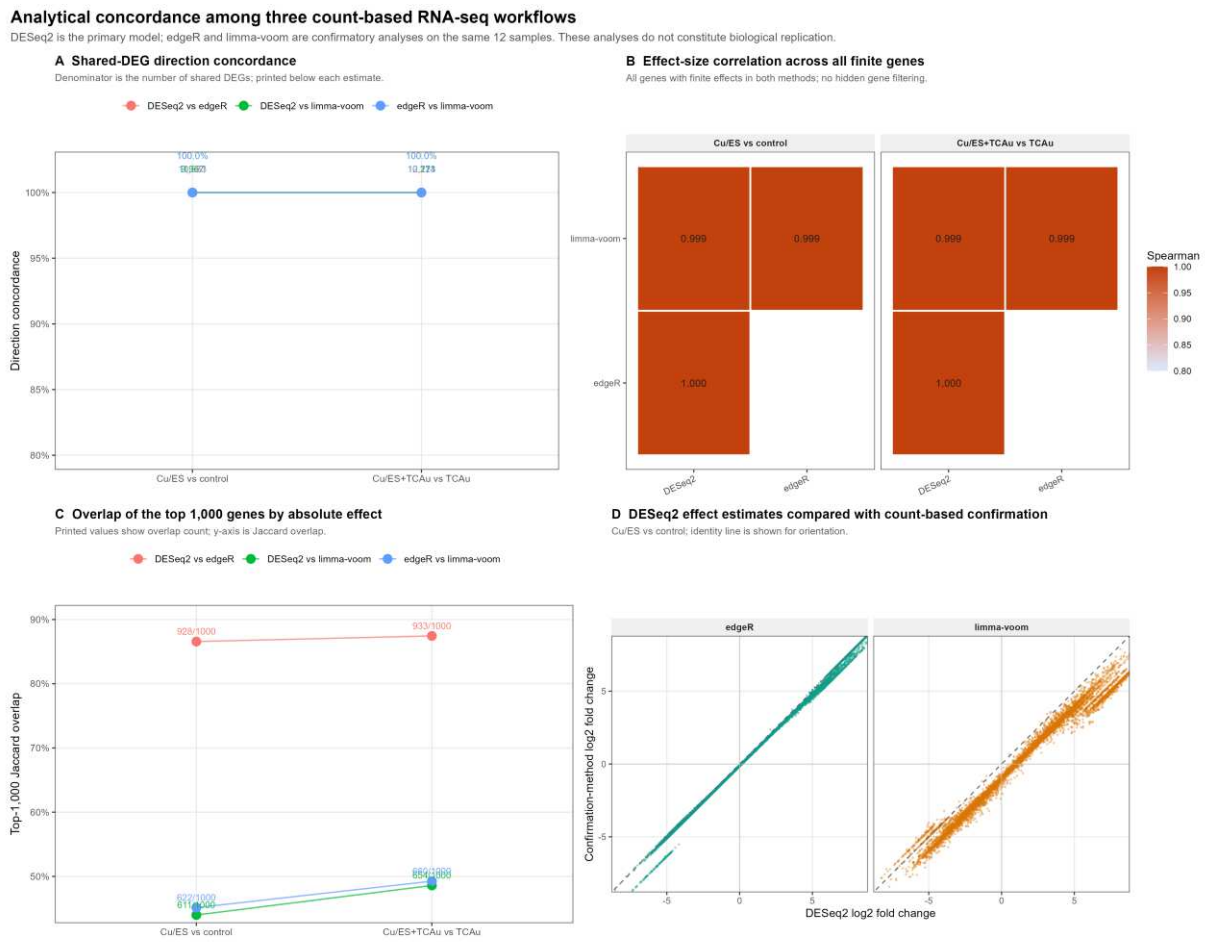


Figure 4. Analytical concordance among the three count-based RNA-seq workflows.

The primary cell-state trends are concordant across the three count-based methods and are not specific to a single normalization or model. DESeq2, edgeR and limma-voom are broadly concordant for interaction effects, shared-DEG direction, top-gene overlap and GSEA correlation. For the 17 prespecified key candidates, 15 agree in direction across all three count methods and 14 reach FDR < 0.05 in all three; Myh9 and Mt2 are directionally discordant, and Atp7a is concordant in direction but not significant under limma-voom. The proteostatic, mitochondrial and cytoskeletal trends are therefore stable across count-based analytical frameworks. This remains a statistical-consistency result: agreement among three methods applied to the same 12 samples is analytical concordance, not independent biological replication, and it does not establish protein-level or functional mechanism.

Figure 5. Composition and interaction sensitivity audit. Detected-gene number, ribosomal fraction, cell-cycle score, housekeeping dispersion and interaction stability. TCAu interaction results are a secondary technical sensitivity and are detailed in Extended Data.

Data source: NCBI GEO GSE300382, using both the reprocessed integer count matrix and the public FPKM matrix (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382). No new data were generated for this figure.

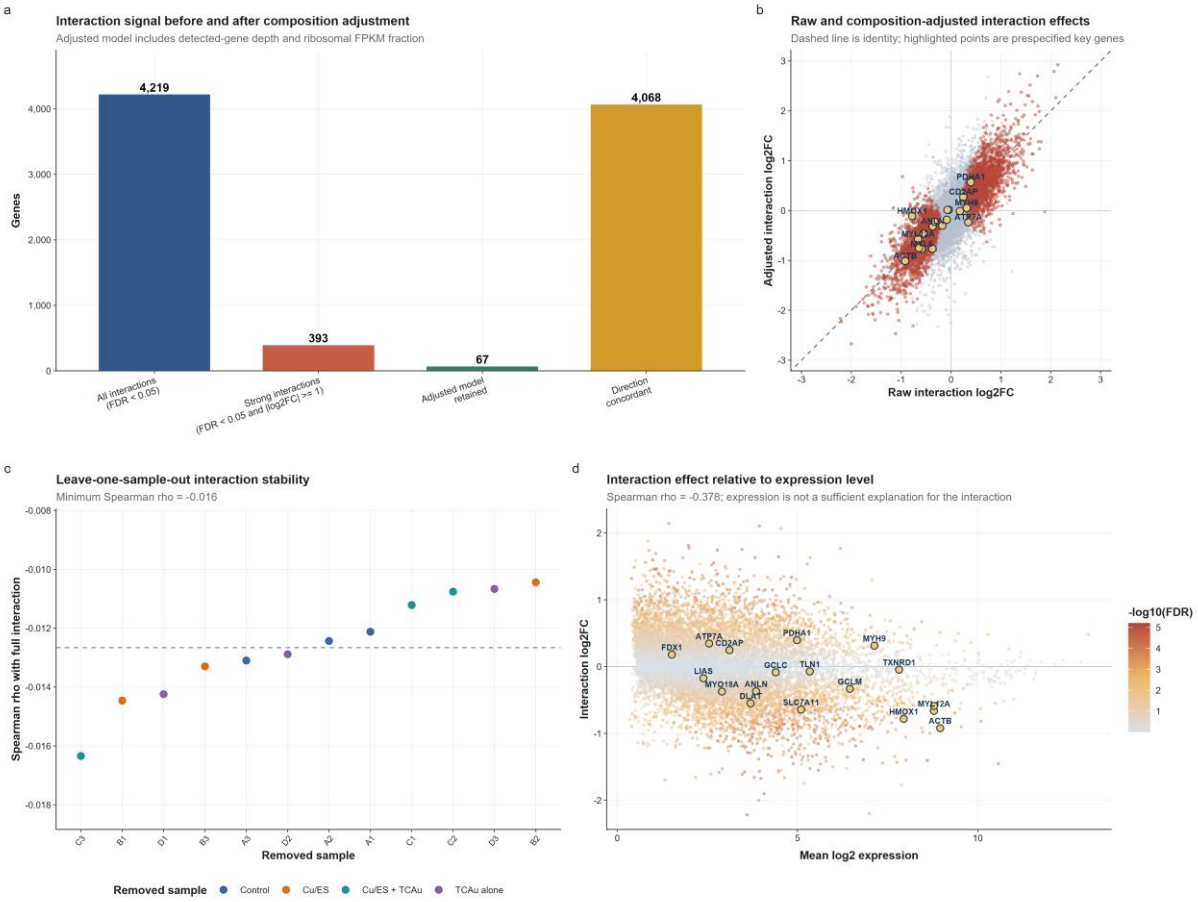


Figure 5. Composition and interaction sensitivity audit.

A higher ribosomal read fraction with a stable count>=10 set indicates altered transcript composition and is compatible with a change in the sampled cell population after copper stress; it does not directly measure survival or cell composition. In the count matrix, the number of genes with count>0 rose from 25,454 in control to 29,325 in Cu/ES, while the count>=10 set was essentially stable (16,066 in control, 16,065 in Cu/ES, 16,691 in Cu/ES+TCAu). The ribosomal read fraction rose from 0.0393 in control to 0.0681 in Cu/ES and 0.0624 in Cu/ES+TCAu. In the public FPKM matrix, FPKM-detected genes fell from about 10,842 to about 8,510 and the ribosomal FPKM fraction rose from about 0.192 to about 0.287; these FPKM values are expression-summary metrics reported only for historical comparability and are not count-based quality-control metrics. Lower cell-cycle scores and greater housekeeping-gene dispersion accompany these changes. These patterns may reflect death of sensitive cells and selection of survivors, a shift in cell state, loss of low-abundance transcripts, and differences in RNA/library quality, sequencing depth or compositional normalization. They support altered

population state but do not prove specific activation of ribosome biogenesis or specific suppression of structural genes; direct viability and cell-composition measurements are still required.

Figure 6. MYH9-template-referenced candidate relationships across structural, copper-handling, metabolic and stress-response layers. Prespecified candidates, expression filtering, evidence domains, rank variation and candidate boundaries. MYH9 is an external template, not a ranked candidate. The main contribution is the auditable separation of template-conditioned similarity from model-based priority within the prespecified candidate universe, not the discovery of a mechanism.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), STRING (https://string-db.org/), Human Protein Atlas (https://www.proteinatlas.org/), UniProtKB (https://www.uniprot.org/), InterPro (https://www.ebi.ac.uk/interpro/) and literature-defined model priors. No new data were generated for this figure.



Figure 6. Candidate-priority framework.

The candidate framework focuses the study on MYH9-template-referenced relationships across structural, copper-handling, metabolic and stress-response layers, but its output is computational priority rather than mechanistic discovery. Candidate ranks change substantially across evidence domains and model specifications, and MYH9 functions as an external mechanism template rather than an independently ranked candidate. The framework is appropriate for generating Cu/ES-cytoskeleton-MYH9 hypotheses, but cannot demonstrate protein function, direct interaction, polymerization or cuproptosis execution.

Figure 7. Stepwise removal of the template and candidates with higher model-based priority. Comparison of the template-referenced, leave-MYH9-template, template-reduced class-adjusted and strict response models. MYH9 is absent from the candidate ranking excluding MYH9; as a template-dependence audit, the MYH9-as-a-

candidate sensitivity term moves to rank 48. This value reflects score construction and prevents template recovery from being read as an independent discovery.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), STRING (https://string-db.org/), Human Protein Atlas (https://www.proteinatlas.org/) and the prespecified model priors listed in Supplementary Table S4. No new data were generated for this figure.

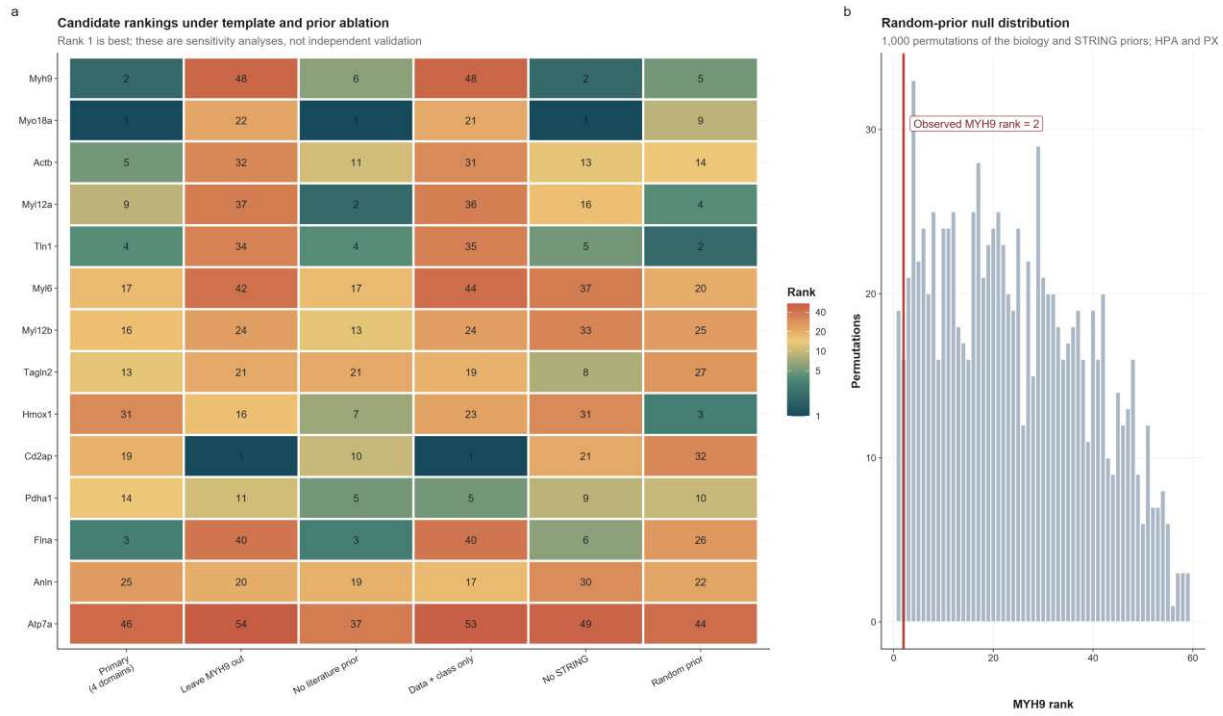


Figure 7. Stepwise removal of the template and candidates with higher model-based priority.

Stepwise removal of the template shows that MYH9-template-referenced relationships must be interpreted separately at the template and model-based priority levels. Removing the MYH9 template moves the MYH9-as-a-candidate sensitivity term from rank 2 to rank 48, while some template-reduced candidates remain relatively stable; this value belongs only to the sensitivity diagnostic, is not a biological candidate rank for MYH9, and prevents the high MYH9 position from being treated as an independent discovery. It supports positioning the paper as Cu/ES-induced AML12 cell-state remodeling with candidate links across structural, copper-handling, metabolic and stress-response layers, not an executor mechanism.

Figure 8. Node-by-node evidence-chain check from copper stress to cytoskeletal transcript remodeling. A, step-by-step evidence status among N1 Cu/ES exposure, N2 molecular/protein layer, N3 pathway remodeling, N4 transcription-factor target-set inference, N5 chromatin accessibility and N6 cytoskeletal transcripts. Solid links indicate same-study transcript or pathway associations, dashed links indicate external background only, and the broken link indicates missing evidence. B, internally defined Tier 0-7 evidence ceiling. The figure shows that the study reaches same-study Tier 1-2 with only partial Tier 3 cross-system background; there is no Tier 5 physical-

interaction and no Tier 6-7 interface or causal-mechanism evidence, the N5-to-N6 link is still missing, and no paired AML12 Cu/ES chromatin measurement exists.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), GTRD (https://gtrd.biouml.org/), MSigDB (https://www.gsea-msigdb.org/gsea/msigdb), Gene Ontology (https://geneontology.org/), and NCBI GEO GSE236123, GSE337266 and GSE337269 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE236123; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE337266; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE337269). No new data were generated for this figure.

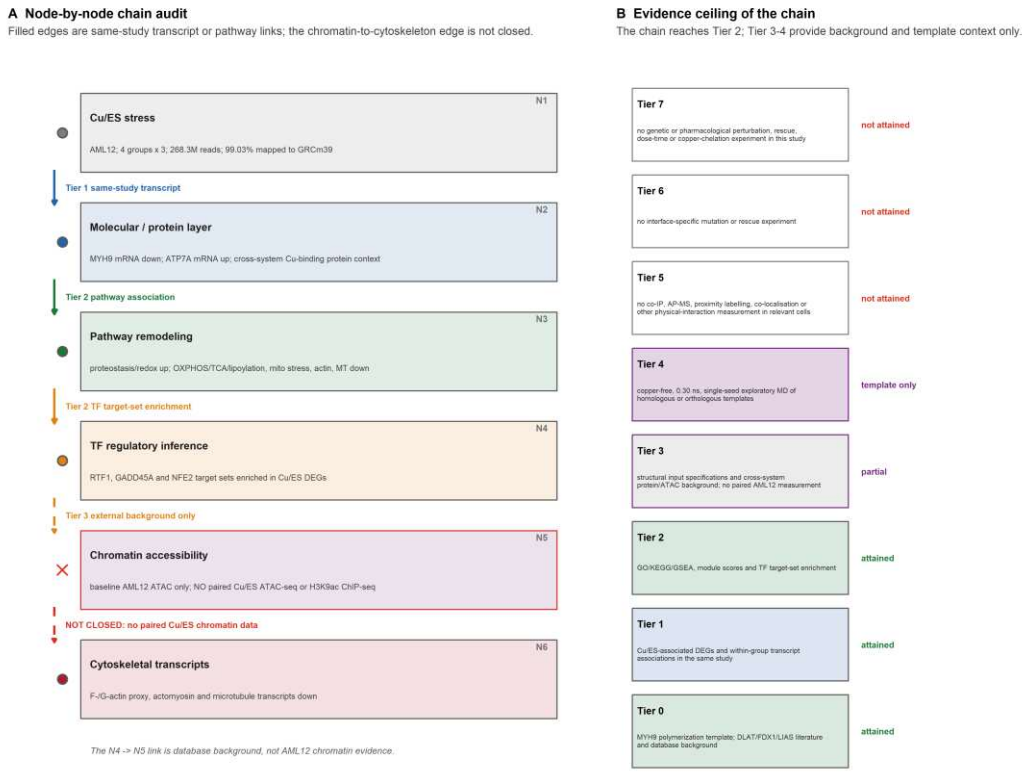


Figure 8. Node-by-node evidence-chain check from copper stress to cytoskeletal transcript remodeling.

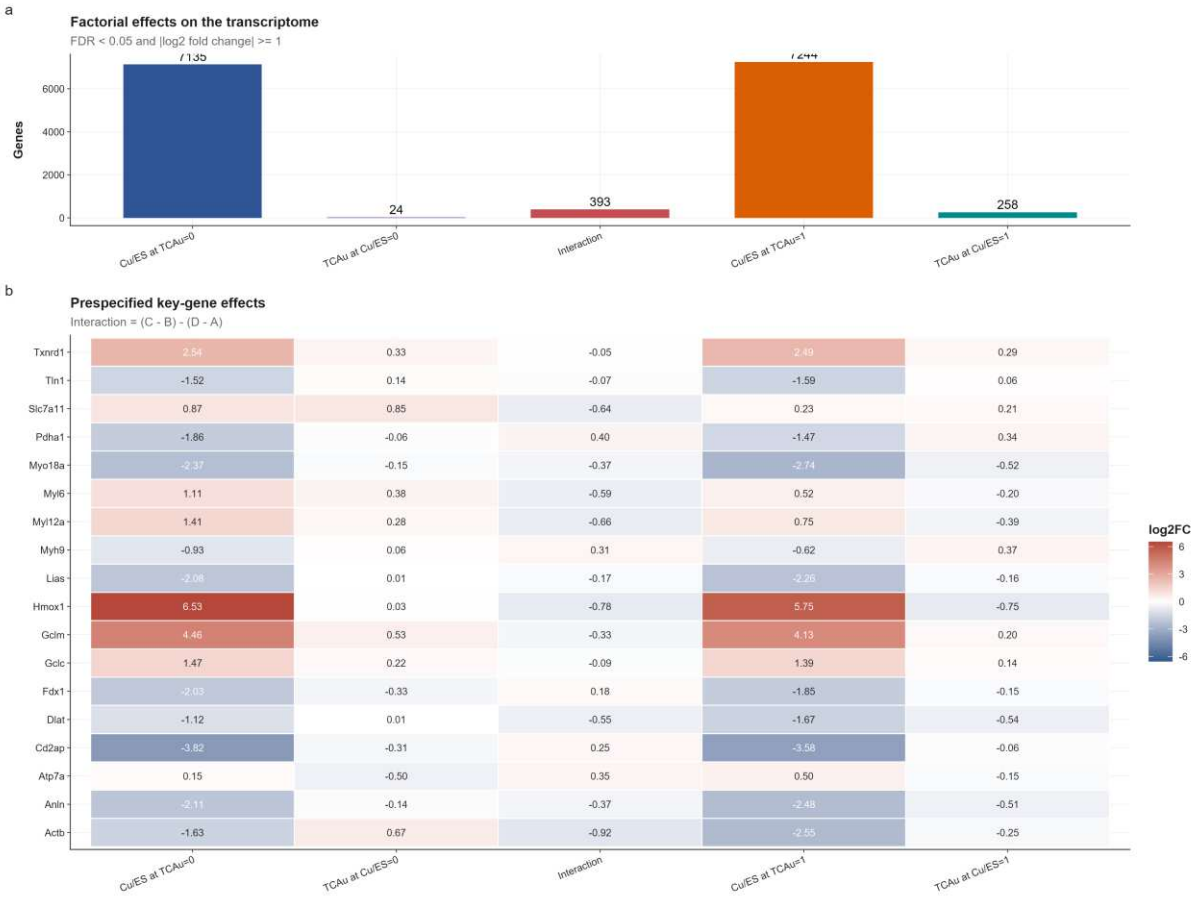
The audit separates analytically concordant association, inferential regulation and missing evidence within the copper-stress-to-cytoskeleton hypothesis. Cu/ES-to-transcript/molecular and pathway remodeling can be supported by count-level differential expression, BH-corrected module scores and enrichment in the same 12 samples. TF target-set enrichment suggests that RTF1-, GADD45A- and NFE2-associated regulatory programs resemble the Cu/ES response, but it does not measure TF activity or chromatin binding. The chromatin-to-cytoskeletal-transcript link has no same-system evidence, so the current data can only propose that copper stress may alter cytoskeletal transcripts through specific regulatory programs. The statistical value of the figure is that it

displays each link with its evidence source, effect estimates, correction strategy and missing measurement, preventing same-matrix correlation or model inference from substituting for the absent experimental layer.

10. Extended Data

ED01. Factorial interaction. Raw and composition-adjusted interaction results. The primary count-level interaction is 1,455 genes in DESeq2 (1,284 in edgeR and 1,283 in limma-voom) and is reported in the main text; the 4,219 values shown here come from the historical FPKM-limma matrix and are an unadjusted audit metric, not the primary interaction estimate.

Data source: NCBI GEO GSE300382, using the reprocessed integer counts and the public FPKM matrix (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382). No new data were generated for this figure.

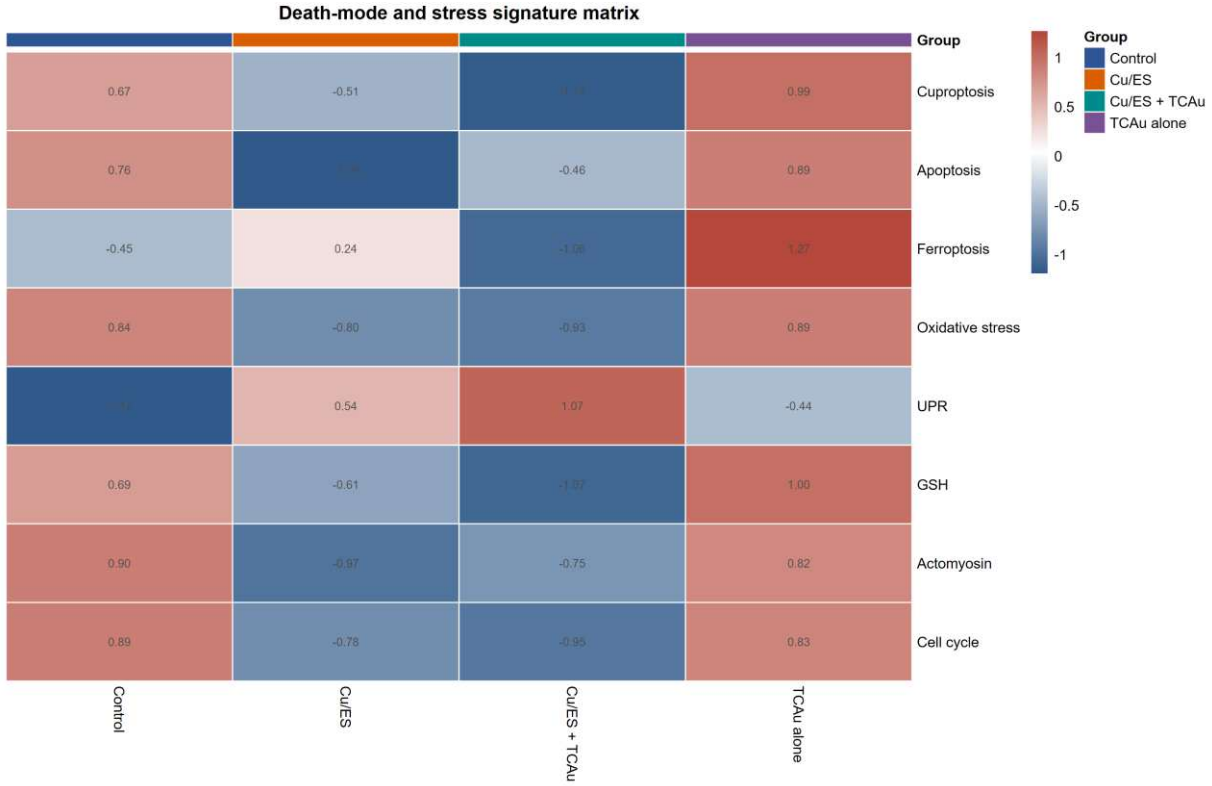


ED01. Factorial interaction.

Stable TCAu interaction with the Cu/ES response is limited, so the primary narrative should focus on Cu/ES-induced cell-state remodeling. In the historical FPKM-limma matrix the unadjusted model reports 4,219 interaction genes, but only 67 remain significant after composition adjustment, and effect size and method sensitivity further limit interpretation. TCAu interaction is a secondary technical audit rather than a substitute for the Cu/ES main effect or evidence of a cuproptosis mechanism.

ED02. Signature and death-mode resemblance. This analysis describes transcriptional resemblance only and cannot determine the actual death mode.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382) and predefined public gene sets from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb). No new data were generated for this figure.

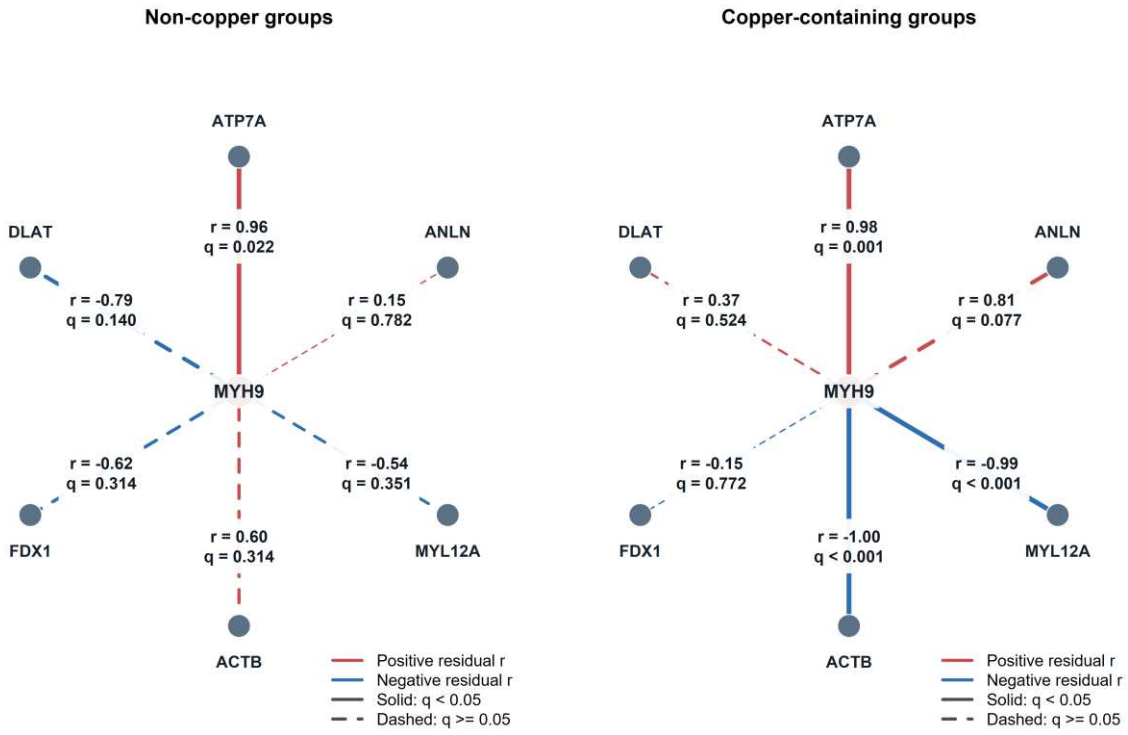


ED02. Signature and death-mode resemblance.

Death-mode gene-set scores provide transcriptional resemblance only and cannot define this study as a cuproptosis mechanism. Preset gene-set scores show varying degrees of similarity, consistent with broad stress-associated state changes after Cu/ES, but they cannot determine the actual death mode or compare the contributions of different death programs. With the composition bias in Figure 5, this analysis is descriptive background rather than cuproptosis evidence.

ED03. Differential residual correlation. Candidate correlations after within-treatment centering. These generate hypotheses and do not establish causal links.

Data source: NCBI GEO GSE300382 (reprocessed integer counts; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382). No new data were generated for this figure.

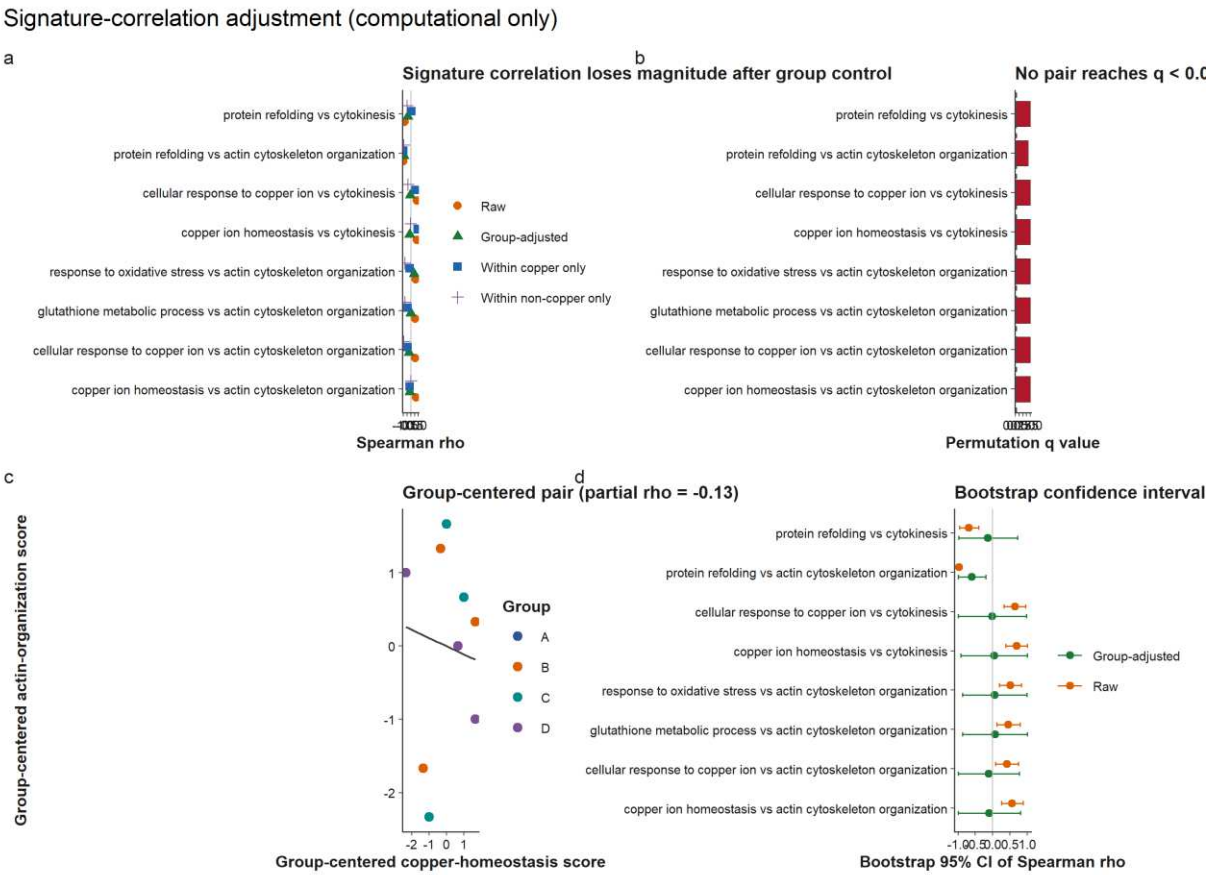


ED03. Differential residual correlation.

Residual-correlation analysis supports ATP7A-MYH9 and MYH9-template-associated candidates as cross-layer relationships for testing. After within-treatment centering, ATP7A-MYH9 retains a residual association, whereas MYH9-ANLN remains mainly a database-supported candidate relationship, consistent with a copper-transport-actin/myosin remodeling research direction. This supports hypothesis generation, not direction, causality or direct protein interaction.

ED04. Gene-set correlation adjustment. High Fig25 correlations largely disappear after treatment-group control and overlap removal.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), with predefined gene sets from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb). No new data were generated for this figure.

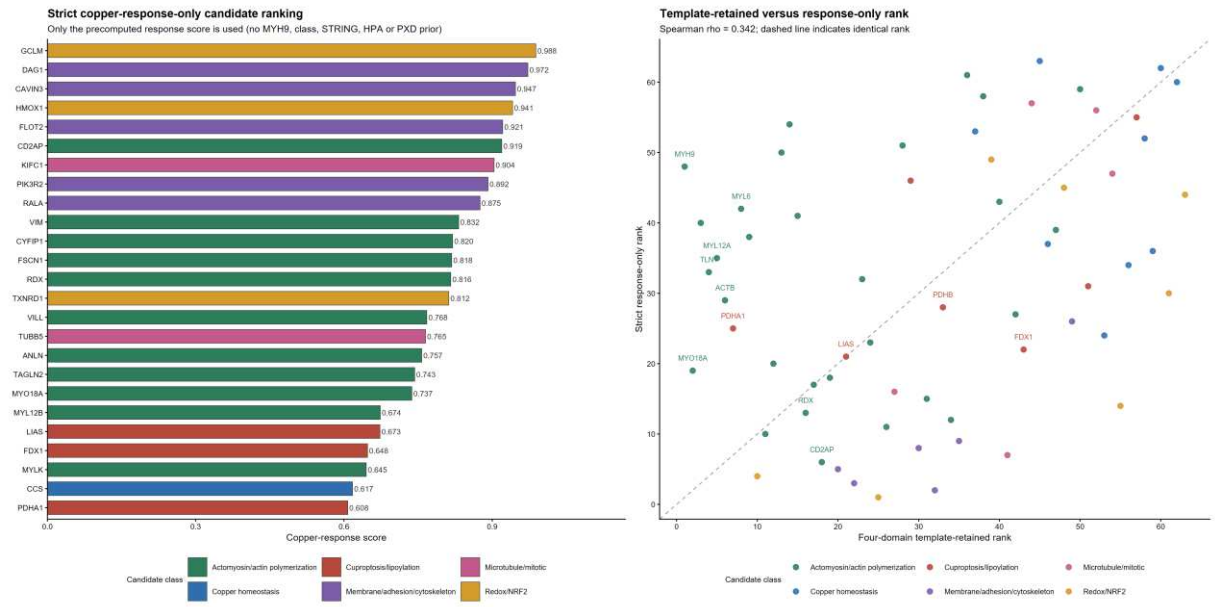


ED04. Gene-set correlation adjustment.

Most strong gene-set score correlations are driven by group separation and gene overlap, so copper-cytoskeleton coupling cannot be treated as an independent discovery. After treatment-group control and overlapping-gene removal, most original correlations disappear, indicating that they largely reflect global between-group state differences rather than stable within-module coupling. Copper-cytoskeleton coupling should therefore remain a weak transcriptomic lead to be tested through candidate-network and experimental work.

ED05. Template ablation and strict response-only ranking. Template-reduced ranking after removal of the MYH9 template, class, STRING, HPA and PXD priors; the MYH9-as-a-candidate sensitivity diagnostic is shown separately and is not part of the candidate ranking.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), STRING (https://string-db.org/), Human Protein Atlas (https://www.proteinatlas.org/) and the prespecified model priors listed in Supplementary Table S4. No new data were generated for this figure.

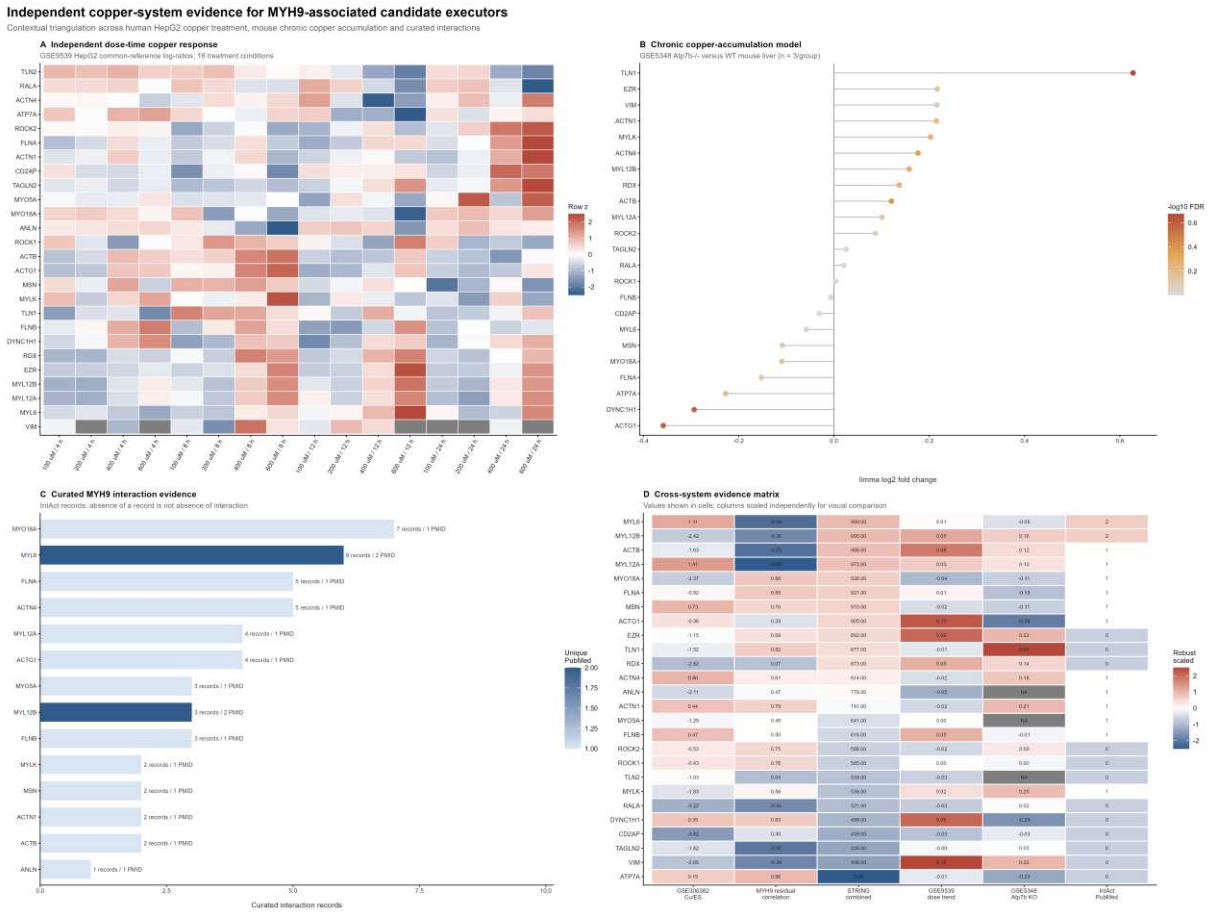


ED05. Template ablation and strict response-only ranking.

Template-reduced ranking retains several candidates that do not depend on the MYH9 template, providing an experimental order for copper-cytoskeleton hypotheses. The strict response model changes candidate order, but CD2AP, RDX, PDHA1, FDX1, LIAS and PDHB can remain relatively high in some models and belong to distinct layers involving cytoskeletal linkage, metabolic execution and protective stress. The figure supports model-based priority status, not structural-execution or cuproptosis function.

ED06. External copper-system context. GSE9539, GSE5348 and IntAct provide cross-system context and do not constitute AML12 validation.

Data source: NCBI GEO GSE9539 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9539), NCBI GEO GSE5348 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE5348) and IntAct (https://www.ebi.ac.uk/intact/). No new data were generated for this figure.

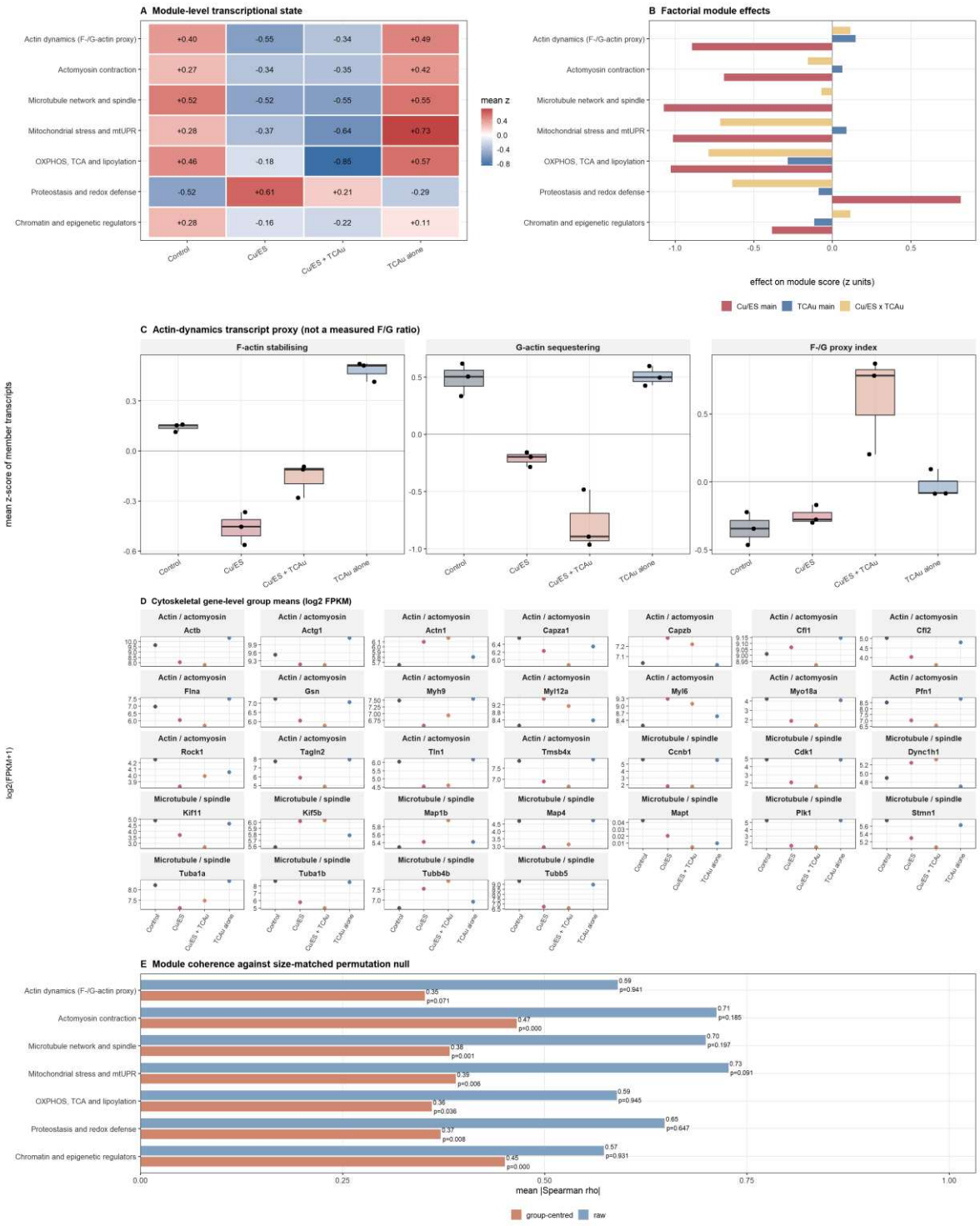


ED06. External copper-system context.

Cross-system evidence supports copper-associated cytoskeletal, mitochondrial and stress programs as a transferable research theme. GSE9539, GSE5348 and IntAct provide heterogeneous context from different cells, chronic copper accumulation and curated interactions, and candidate directions are not uniform, indicating dependence on system, dose and time. These data do not validate AML12 or demonstrate protein-layer changes in this study.

ED07. Actin, microtubule and mitochondrial extension. The F-actin/G-actin score is a transcriptional proxy, not protein fractionation. Module changes remain composition-sensitive.

Data source: NCBI GEO GSE300382 (requantified integer counts and public FPKM matrix; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382). No new data were generated for this figure.

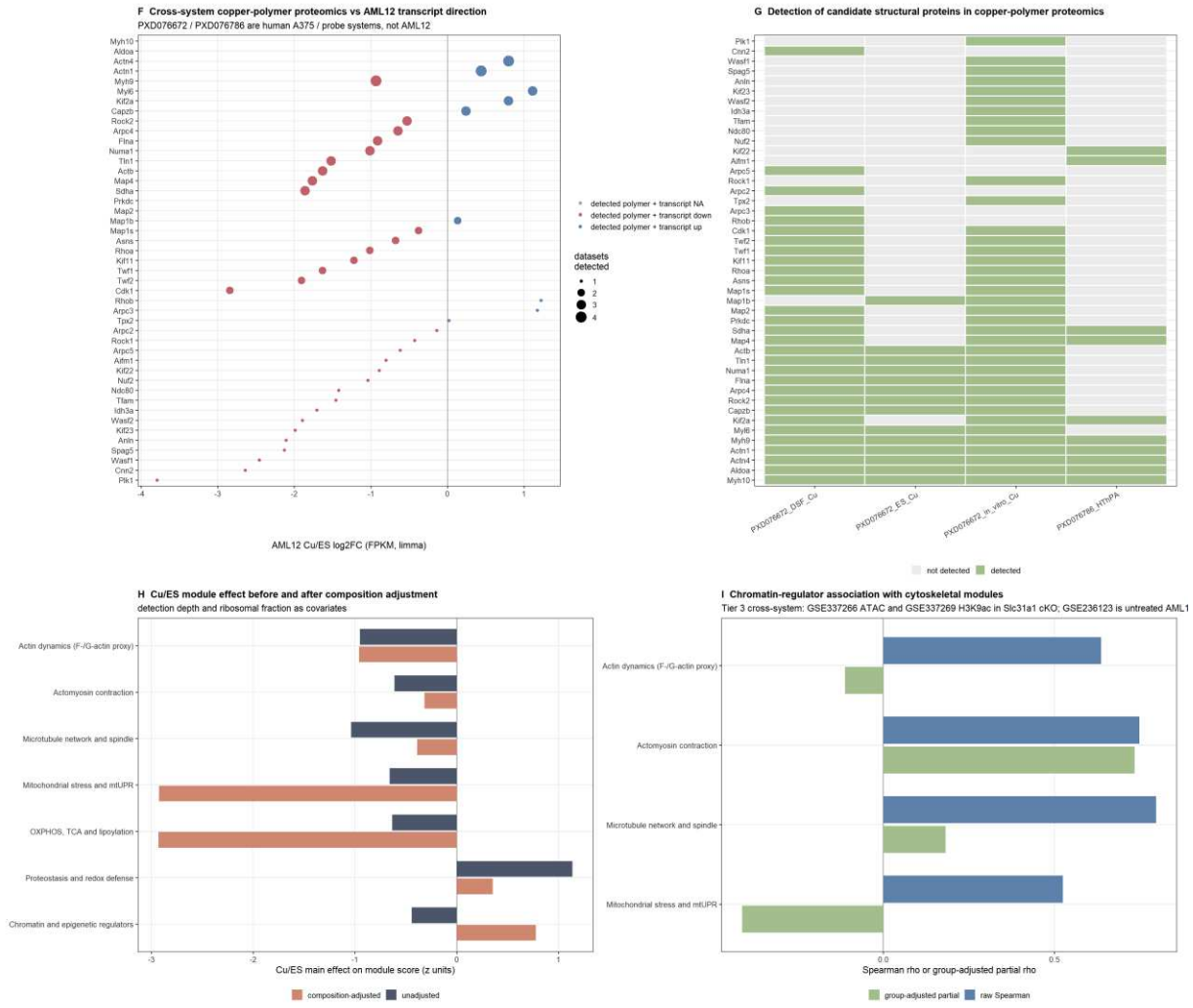


ED07. Actin, microtubule and mitochondrial extension.

Cu/ES is associated with directional transcriptional changes across actin, actomyosin, microtubule and mitochondrial programs. This pattern supports a model in which copper stress involves cytoskeletal and mitochondrial remodeling rather than activation of a single death pathway; reduced microtubule and cytoskeletal programs remain compatible with the composition bias in Figure 5. F-actin/G-actin is only a transcript proxy and is not equivalent to protein fractionation.

ED08. Cross-system proteome and chromatin context. These data provide background only and cannot demonstrate changes in AML12.

Data source: PRIDE PXD076672 (https://www.ebi.ac.uk/pride/archive/projects/PXD076672) and PXD076786 (https://www.ebi.ac.uk/pride/archive/projects/PXD076786); NCBI GEO GSE236123, GSE337266 and GSE337269 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE236123; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE337266; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE337269). No new data were generated for this figure.

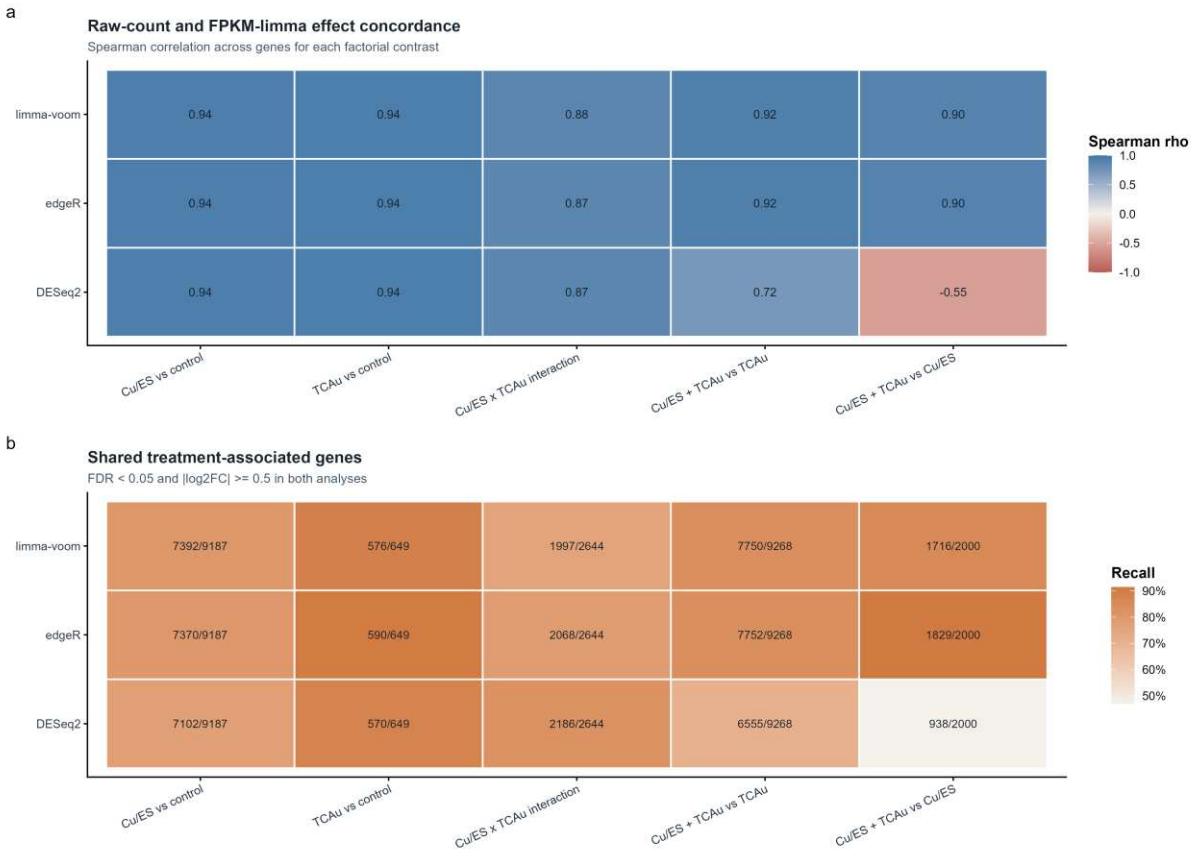


ED08. Cross-system proteome and chromatin context.

External proteomic and chromatin data provide a testable background for MYH9-template-referenced candidates across structural, copper-handling, metabolic and stress-response layers, but not AML12 mechanism evidence. Detection of selected candidates in copper-related proteomic or chromatin datasets supports their suitability for subsequent protein- and regulatory-layer work. Cross-system detection does not demonstrate abundance, localization, polymerization or chromatin regulation in AML12.

ED09. Count-level primary-analysis concordance. Extended companion to main-text Figure 4, reporting Spearman correlation, direction concordance and reference DEG recall across main contrasts. DESeq2 is the primary model, edgeR and limma-voom are count-level confirmation, and FPKM-limma is a historical comparator only.

Data source: NCBI GEO GSE300382, using the reprocessed integer counts and the public FPKM matrix for the historical comparator (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382). No new data were generated for this figure.

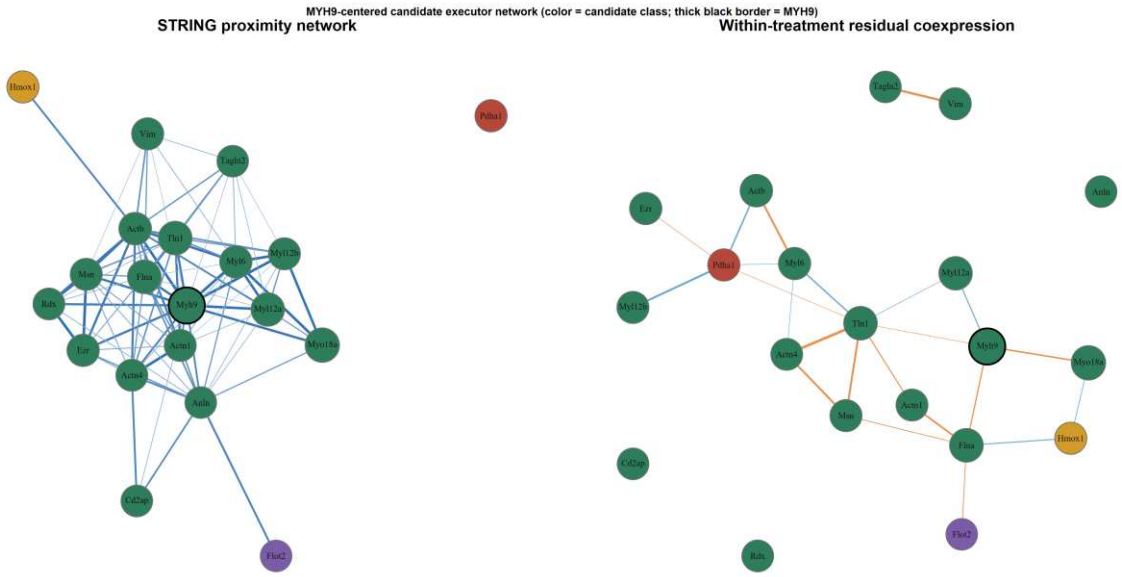


ED09. Raw-count contrast concordance.

Count-level group contrasts are broadly concordant with the historical FPKM matrix, which supports the analytical robustness of the Cu/ES-associated cell-state pattern rather than its biological replication. Correlations, direction concordance and DEG recall remain high across methods, indicating that the central stress, metabolic and cytoskeletal trends are not unique to one normalization strategy. This remains a technical concordance analysis, not a substitute for independent samples or biological validation.

ED10. MYH9-template-referenced candidate relationship network. Within-treatment residual correlations, STRING proximity and candidate boundaries. This network is a hypothesis-generation and experimental-design tool, not a reconstructed regulatory network, and it does not identify causal links or direct protein interactions.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), STRING (https://string-db.org/) and IntAct (https://www.ebi.ac.uk/intact/). No new data were generated for this figure.

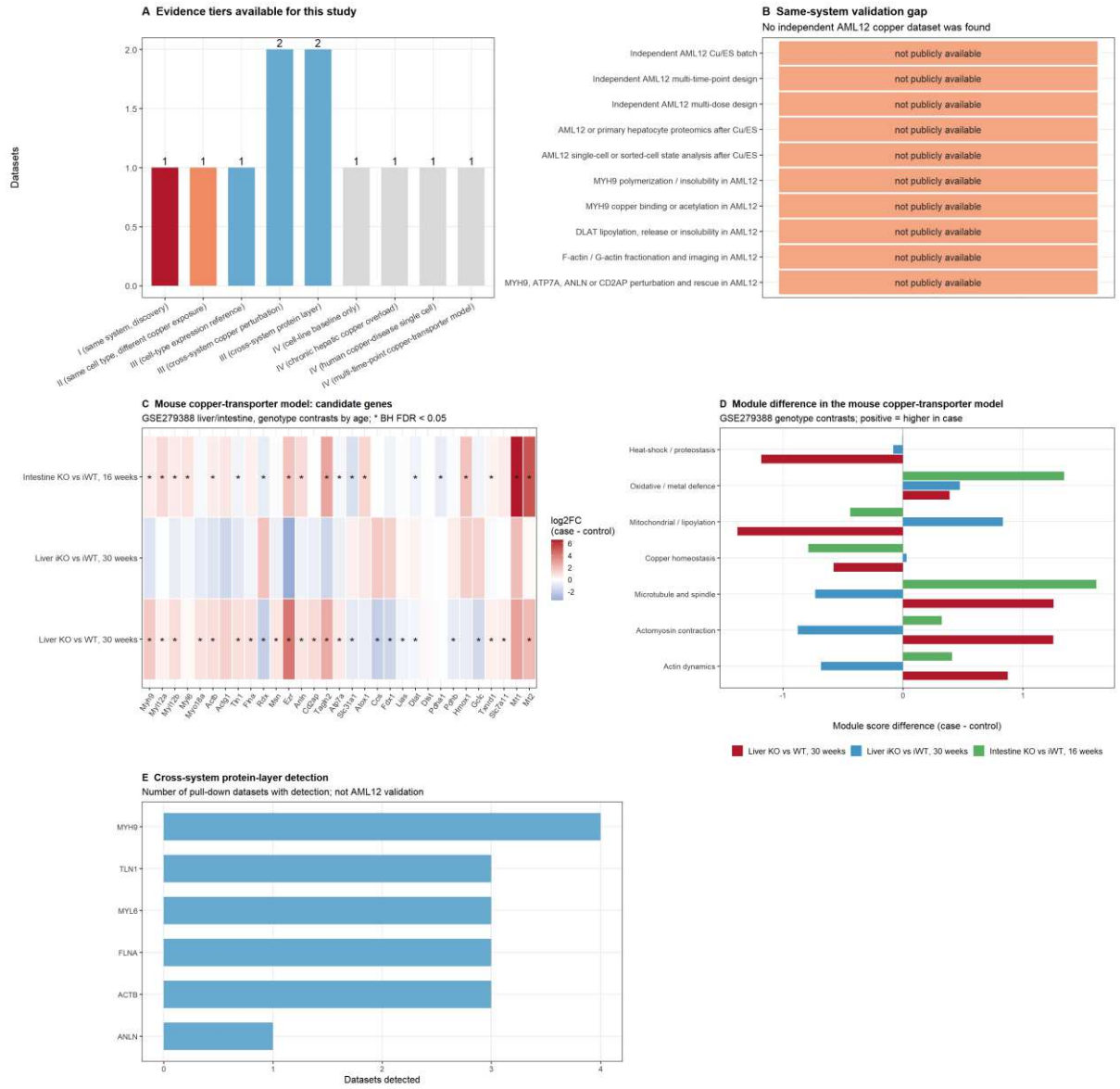


ED10. MYH9-related candidate relationship network.

The network organizes MYH9-template-referenced candidates into relationships for experimental design. Within-treatment residual relationships, STRING proximity and candidate boundaries can guide colocalization, co-immunoprecipitation and perturbation tests involving MYH9, ATP7A, ANLN and actin-myosin-associated proteins. It is not a reconstructed regulatory network, and its links do not represent causal effects or direct protein interactions.

ED21. Same-system replication search and multi-layer background evidence. A, evidence tiers in the available public record; B, current same-system replication gaps; C, age-, tissue- and genotype-matched candidate effects in GSE279388; D, module differences in GSE279388; E, cross-system protein-layer detection. GSE279388 is a chronic genetic copper-transporter model, not an AML12 Cu/ES time course; PXD076672/PXD076786 and GSE254082 are cross-system protein or single-cell background. None constitutes AML12 validation.

Data source: NCBI GEO GSE300382 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), GSE279388 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE279388) and GSE254082 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE254082); PRIDE PXD076672 and PXD076786 (https://www.ebi.ac.uk/pride/archive/projects/PXD076672; https://www.ebi.ac.uk/pride/archive/projects/PXD076786); Human Protein Atlas (https://www.proteinatlas.org/). No new data were generated for this figure.

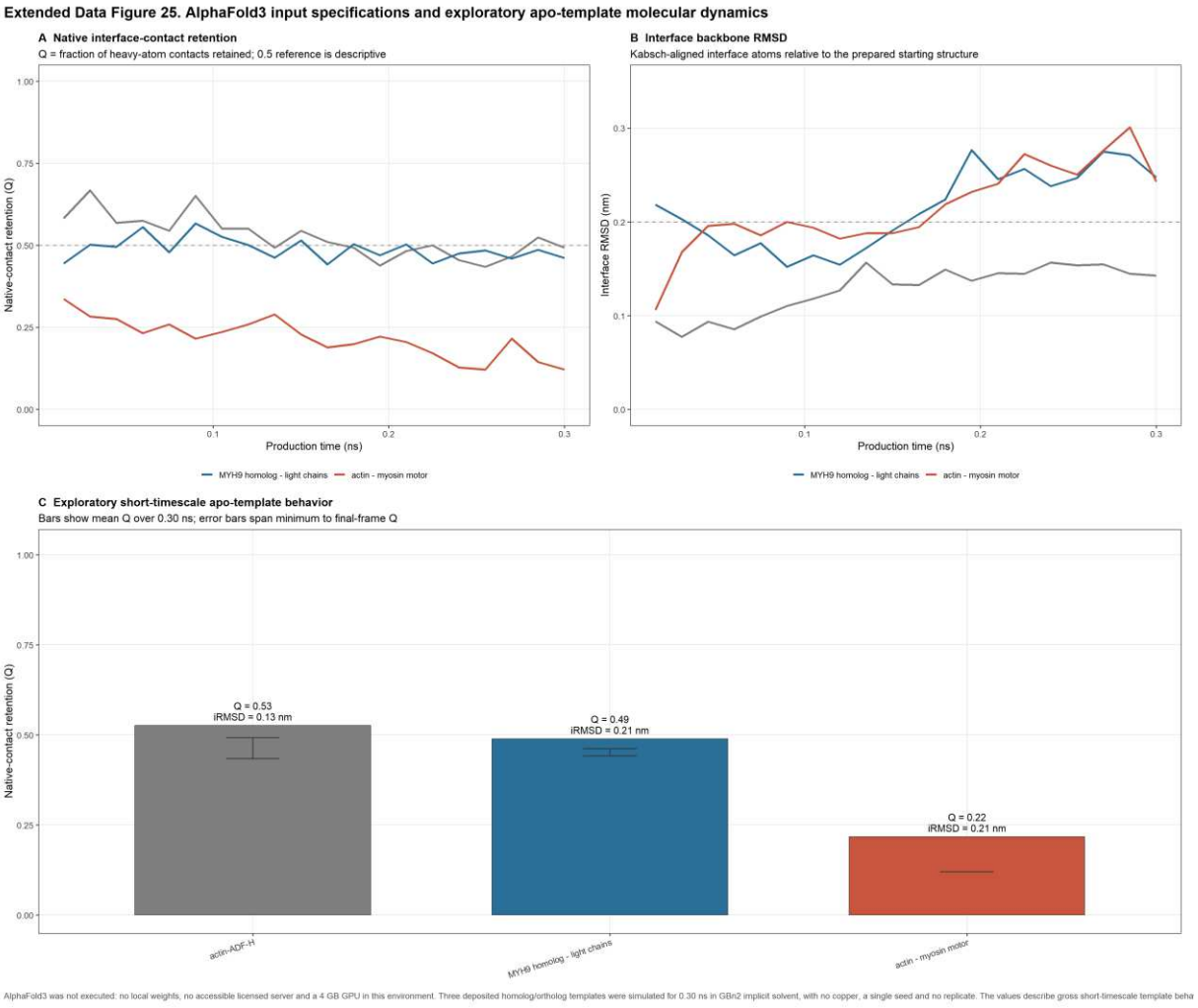


ED21. Same-system replication search and multi-layer background evidence.

No independent AML12 Cu/ES validation dataset was identified, further defining the limits of this study. A and B show the absence of same-cell, same-treatment validation; C and D show that candidate directions in chronic copper-transporter models vary by tissue, age and genotype; E provides cross-system protein-layer detection only. External evidence therefore supports background and experimental priority for copper-cytoskeleton remodeling, but cannot validate the mechanism in AML12.

ED25. AlphaFold3 input specifications and copper-free template molecular dynamics. A, native-contact retention Q versus production time for the three template complexes; the dashed line at 0.5 is descriptive only. B, interface backbone RMSD versus production time; the dashed line at 0.2 nm is descriptive only. C, mean Q per complex (bars) with the minimum-to-final-frame range (error bars). All simulations are copper-free apo systems in GBn2 implicit solvent.

Data source: RCSB PDB structures 1S5G (https://www.rcsb.org/structure/1S5G), 5JLH (https://www.rcsb.org/structure/5JLH) and 3DAW (https://www.rcsb.org/structure/3DAW). OpenMM and PDBFixer are software tools, not data sources. AlphaFold3 was not executed, and no new experimental data were generated for this figure.



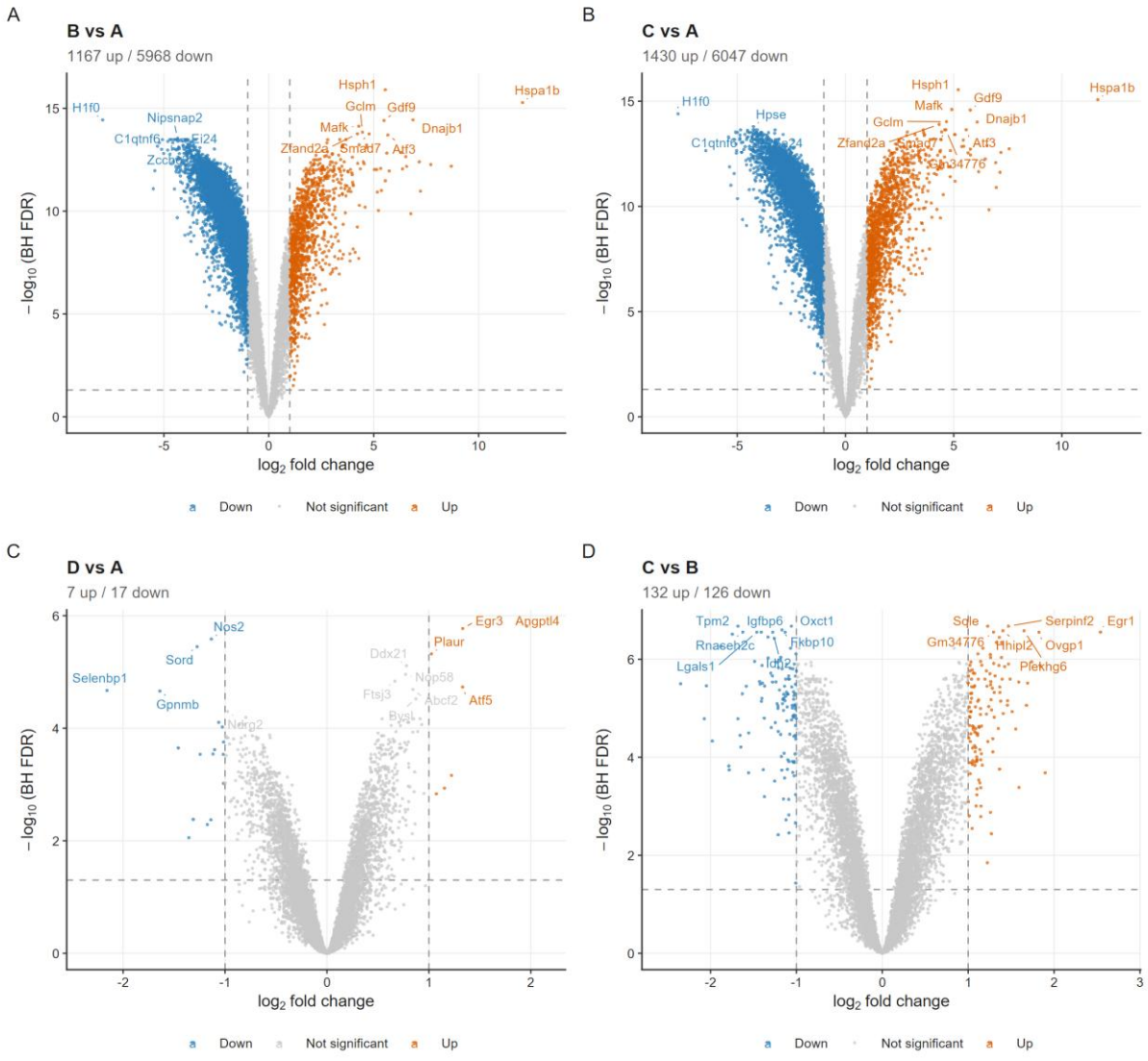
ED25. AlphaFold3 input specifications and copper-free template molecular dynamics.

This figure provides an exploratory short-timescale readout for template interfaces, not an interaction strength and not a copper-dependence result. The MYH9-homolog heavy-chain/light-chain template (PDB 1S5G) retains native inter-chain contacts throughout the production trajectory, with interface backbone RMSD remaining in the sub-nanometre range over this short, copper-free, implicit-solvent sampling window. The unrelated actin-binding structural control (PDB 3DAW) reaches the same order of stability, which indicates that this stability mainly reflects the well-characterized actin-binding fold rather than a copper-responsive property specific to MYH9. The 5JLH actomyosin template produced an anomalous reported radius of gyration and reduced contact retention; because this short trajectory may reflect template-preparation or sampling artifacts, it is not interpreted as evidence of biological interface weakening. Because no copper ion is present, the figure cannot

say whether copper promotes or disrupts these interfaces, and Q must not be read as a binding affinity. Its biological value is that these templates provide starting points for further dynamics and experimental perturbation, not endpoints.

ED26. Historical FPKM-limma differential-expression volcano. This figure shows the differential-expression structure recovered from the public FPKM-limma analysis. It is retained only for historical comparability with the original public expression matrix and is not used for primary inference or candidate prioritization.

Data source: NCBI GEO GSE300382, public FPKM matrix (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382). No new data were generated for this figure.



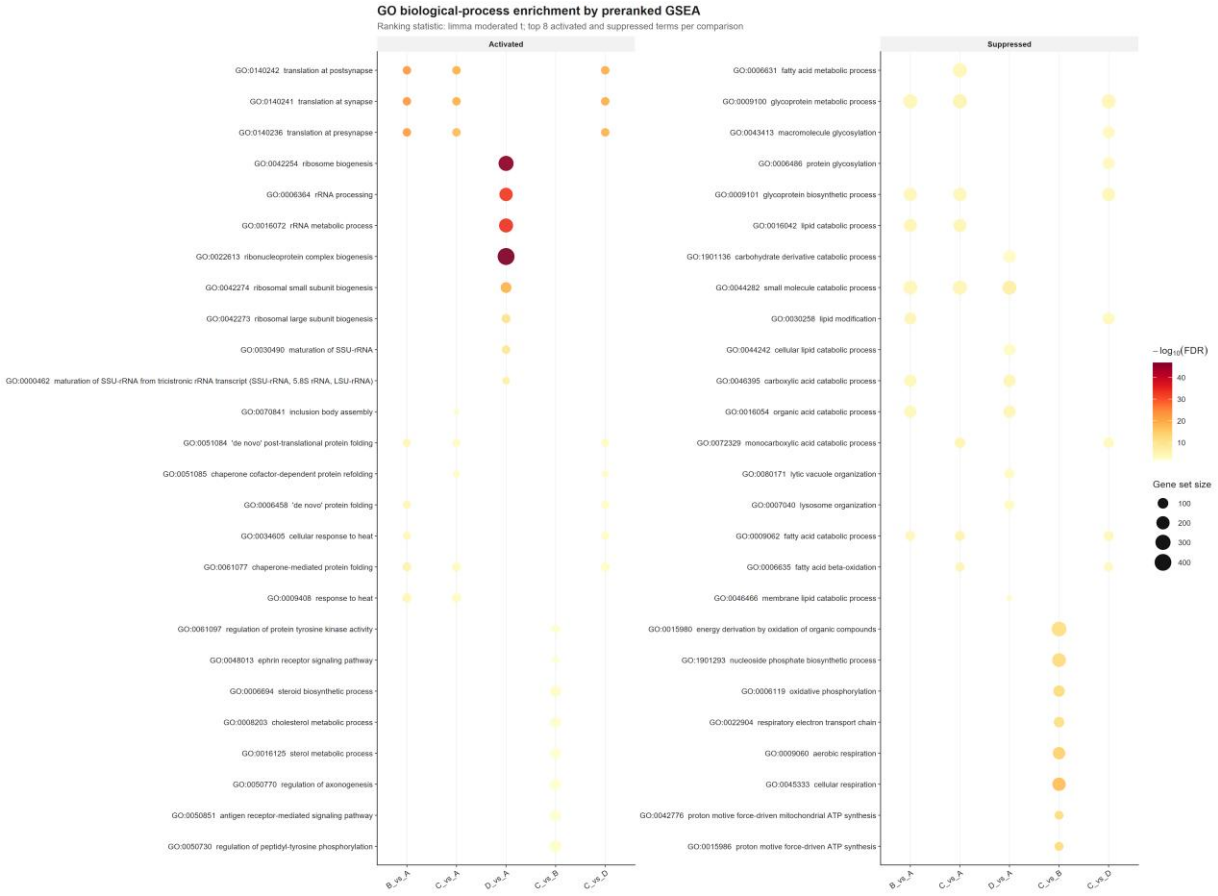
ED26. Historical FPKM-limma differential-expression volcano.

The historical FPKM volcano shows broad copper-stress-associated transcriptional changes, but its statistical scale cannot replace the count-level models. Cu/ES-associated contrasts contain many up- and down-

regulated genes, and the general direction is compatible with the count-level primary analysis. However, FPKM-limma can give an overly optimistic impression of interaction stability. The value of this figure is therefore to show that the revised study remains traceable to the original public matrix, not to provide new mechanism evidence for MYH9, ANLN or ATP7A.

ED27. Historical FPKM gene-set enrichment. This figure shows gene-set enrichment from the public FPKM matrix and is retained as historical cross-scale background only.

Data source: NCBI GEO GSE300382, public FPKM matrix (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE300382), Gene Ontology (https://geneontology.org/) and org.Mm.eg.db 3.20.0 (https://bioconductor.org/packages/org.Mm.eg.db/). No new data were generated for this figure.

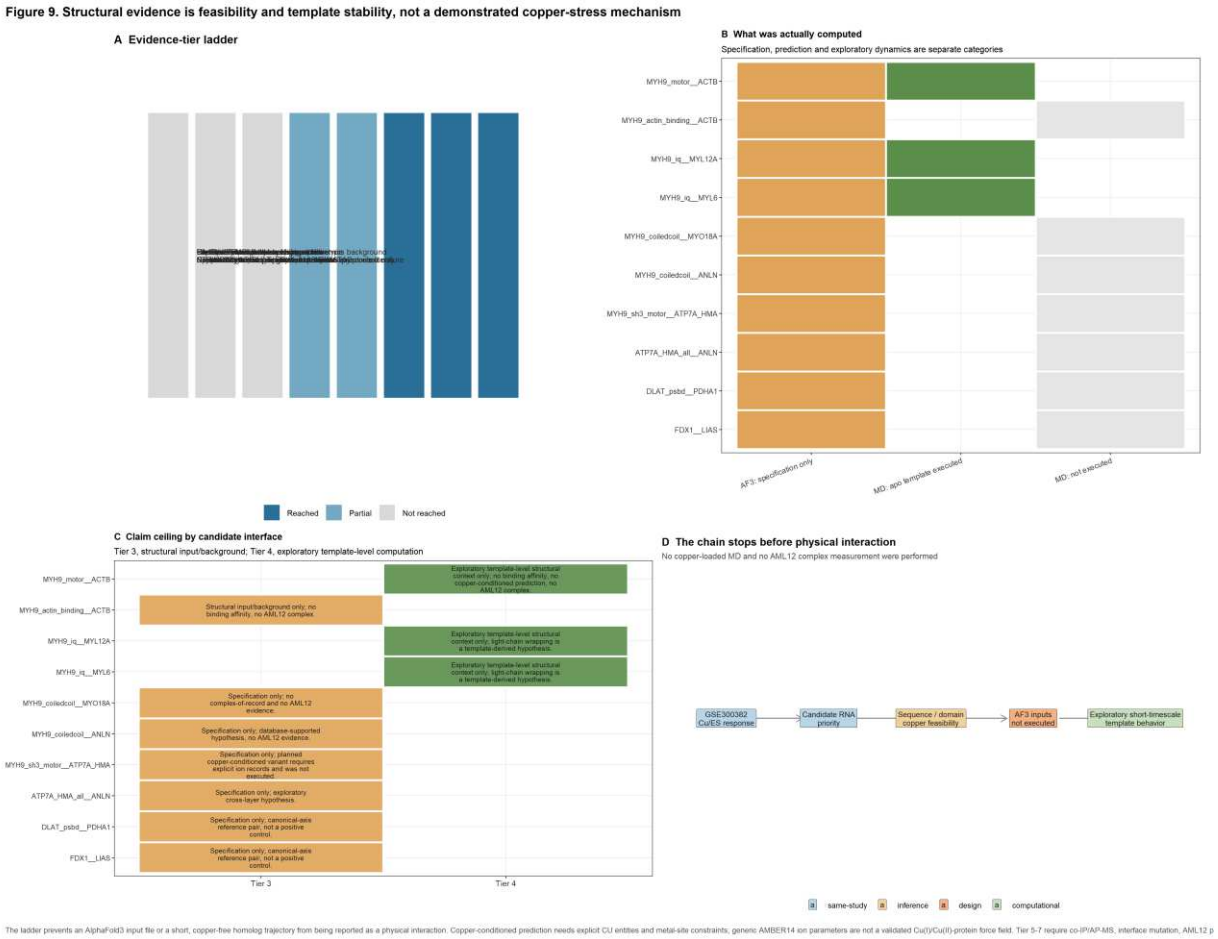


ED27. Historical FPKM gene-set enrichment.

The historical FPKM enrichment results are broadly compatible with the direction of the count-level analysis, but they remain an earlier expression-scale view. The figure supports compatibility between the count-level reanalysis and the published matrix. It does not provide mechanism evidence, does not establish biological replication and does not change the primary conclusions or the definition of candidate priority.

ED28. Structural evidence tier and the template-level computational boundary. A, the internally defined Tier 0-7 evidence framework and the status reached by this study at each tier. B, what was actually computed for the 10 candidate interfaces: AlphaFold3 specifications were generated as planned input files but not executed, and only experimentally resolved homolog/ortholog templates completed copper-free molecular dynamics. C, the limit on what can be claimed for each candidate interface (Tier 3, structural input or cross-system background; Tier 4, exploratory template-level computation). D, the computational chain from Cu/ES response through candidate RNA priority and domain-level copper feasibility to planned AlphaFold3 inputs and copper-free template dynamics, which terminates at the unexecuted copper-conditioned prediction and the unmeasured AML12 complex.

Data source: This is a conceptual summary of the public datasets and computational inputs listed in the Data and code availability section, including NCBI GEO GSE300382, STRING, UniProtKB, InterPro, GTRD/MSigDB and RCSB PDB 1S5G, 5JLH and 3DAW. No new experimental data were generated for this figure.

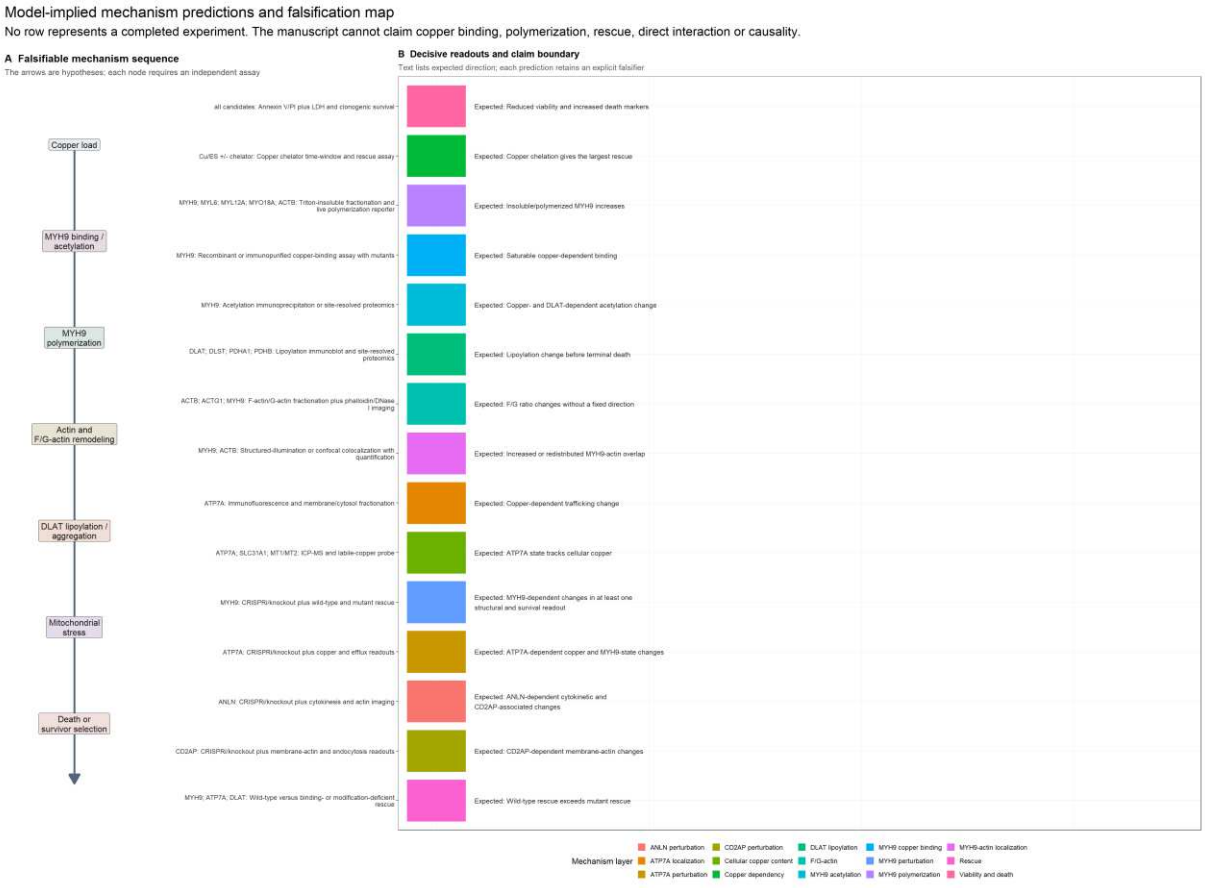


ED28. Structural evidence ladder and the template-level computational boundary.

This claim-boundary figure sets a ceiling on every structural statement in the paper; it is not a mechanistic result. The figure organizes literature/database background (Tier 0), same-study transcript association (Tier 1), pathway or target-set association (Tier 2), structural or cross-system background (Tier 3), exploratory template-

level computation (Tier 4), physical interaction (Tier 5), interface-specific perturbation (Tier 6) and causal perturbation with rescue (Tier 7). This is an internally defined framework for organizing claims, not a field-wide standard. AlphaFold3 was planned but not executed. Only three protein templates completed short, single-seed, copper-free molecular dynamics. This study therefore has same-study transcript and pathway evidence at Tiers 1-2, planned structural input and exploratory template computation at Tiers 3-4, and no Tier 5 physical-interaction evidence and no Tier 6-7 interface or causal evidence. The biological meaning of the figure is that the next experimental priority is Co-IP/AP-MS, interface mutation, perturbation and rescue, rather than further stacking of transcriptional network models.

Supplementary S1. Falsifiable experimental matrix. This item is restricted to the supplement. It describes how future experiments should distinguish structural events, composition changes and the canonical mitochondrial lipoylation axis. It contains no new wet-lab validation, is not a main-text discovery and does not provide data-supported mechanistic evidence. All future experiment matrix content is confined to Supplementary S1.



Supplementary S1. Falsifiable experimental matrix.

Data source: This is a conceptual experimental-design matrix. No dataset was generated or analyzed for this panel.

The experimental matrix converts the copper-cytoskeleton remodeling hypothesis into a falsifiable validation path. Structural events, composition correction, the canonical lipoylation axis and functional perturbation are designed to test MYH9 polymerization, F-actin/G-actin changes, DLAT lipoylation and candidate dependencies. It contains no new experimental result and is therefore not mechanistic evidence; it does not change the hypothesis-generating scope of the paper.

11. Reporting summary

- 鈥?Study type: public-data secondary analysis, hypothesis-generating.
- 鈥?Samples: four groups x three biological replicates.
- 鈥?Primary scale: integer gene-level counts; DESeq2 was the primary model, with edgeR and limma-voom as count-level confirmation.
- 鈥?Historical comparison: log2(FPKM+1)+limma was retained only for comparability with the public matrix.
- 鈥?Raw-count primary analysis: completed; 12/12 SRA samples, 99.03% mapping, and 19,275 filtered genes in DESeq2 and limma-voom (18,974 in edgeR after `filterByExpr`). Count-level analytical concordance: 15/17 key candidates were directionally concordant across DESeq2, edgeR and limma-voom, and 14/17 reached FDR<0.05 in all three count methods. Myh9 and Mt2 were directionally discordant; Atp7a was directionally concordant but not significant under limma-voom. The historical FPKM comparator is reported separately and was not used to define the primary tier. TCAu_at_CuES was method-sensitive.
- 鈥?New wet-lab experiments: none.
- 鈥?Copper binding, MYH9 polymerization, acetylation, DLAT release, F-actin/G-actin and cellular copper: not measured.
- 鈥?Death mode: not determined.
- 鈥?Candidate role: computational priority, not a validated execution factor.
- 鈥?Main contribution: computationally auditable candidate ranking and a framework for stepwise removal of prior evidence domains.
- 鈥?Evidence chain: copper stress to transcript/pathway is same-study Tier 1-2 association and TF target enrichment is Tier 2 inferential support; no paired AML12 Cu/ES chromatin data exists, N5-to-N6 is still missing, and there is no Tier 5 physical-interaction or Tier 6-7 causal-mechanism evidence.
- 鈥?Same-system replication: no independent AML12 copper-treatment dataset was found; GSE279388, PXD076672/PXD076786 and GSE254082 provide multi-time-point, protein-layer and single-cell background evidence only.

鈥?Structural layer: AlphaFold3 was not executed (no weights, no licensed server, insufficient GPU memory); 10 planned apo and explicit-copper input specifications were generated; three experimentally resolved homolog/ortholog templates completed copper-free GBn2 implicit-solvent MD, reaching Tier 3 input specification and at most Tier 4 exploratory template-level computation, and constituting neither physical-interaction nor causal evidence.

Declarations

Funding: No funding was received for this study.

Author contributions: Jinwei Li is the sole author, first author and corresponding author.