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.

Two-module architecture of ribosome-associated quality control in human brain: divergent co-expression networks of PELO, LTN1, and NEMF

Drake H. Harbert

drake@innerarchitecturellc.com

Inner Architecture LLC

Research Article

Keywords: Ribosome-associated quality control, Co-expression network, PELO, LTN1, NEMF, Human brain, GTEx, Translational surveillance, mRNA decay, Proteostasis, Epigenetic regulation, Hippocampus

Posted Date: February 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8887367/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Ribosome-associated quality control (RQC) resolves stalled translational complexes through sequential ribosome rescue (PELO), nascent chain ubiquitination (LTN1), and C-terminal alanine–threonine tail modification (NEMF). While the biochemical sequence of these events is well characterized, the broader transcriptional coordination of individual RQC components with other cellular systems in the human brain remains unexplored.

Methods

We performed genome-wide Pearson co-expression analysis for PELO, LTN1, and NEMF across 16,225 expressed genes in the GTEx v8 dataset, using Brain–Frontal Cortex (BA9; $n = 209$) as the primary region. The top 5% of co-expressed genes for each target were compared using Jaccard similarity, Fisher's exact test, and Spearman rank correlation, with permutation-based null models (10,000 random gene pairs) and bootstrap confidence intervals (1,000 resamples) providing formal statistical assessment. Gene Ontology enrichment was performed via gProfiler. Multi-region replication was assessed across four additional brain regions, with cell-type deconvolution, covariate adjustment, and agonal stress sensitivity analysis as robustness checks.

Results

LTN1 and NEMF shared near-identical co-expression networks (Jaccard $= 0.461$, 95% CI [0.357, 0.546]; Spearman $\rho = 0.986$; permutation $p = 0.006$), while PELO diverged substantially (PELO–LTN1 $J = 0.165$, 95% CI [0.095, 0.242]; PELO–NEMF $J = 0.156$, 95% CI [0.082, 0.227]; permutation $p > 0.10$ for both). Bootstrap confidence intervals confirmed non-overlapping separation between the LTN1–NEMF module and PELO pairs. The two-module partition was robust across network thresholds from 1% to 20% (partition ratio 1.64×–8.78×). Genes unique to PELO (528 genes) were enriched for protein metabolic process ($p = 9.35 \times 10^{-22}$) and RNA metabolism (Reactome $p = 1.58 \times 10^{-20}$). Genes unique to NEMF (246 genes) were enriched for histone modifying activity ($p = 2.16 \times 10^{-5}$). Multi-region replication confirmed the two-module architecture (median cross-region $\rho = 0.89$–0.90), with PELO exhibiting uniformly high cross-region consistency (10/10 pairs > 0.80) and the LTN1–NEMF module showing hippocampus-specific divergence (all six sub-0.80 pairs involved hippocampus). All patterns were robust to agonal stress adjustment (rank preservation $\rho \geq 0.998$).

Conclusions

RQC in human brain operates as a two-module system: a surveillance–mRNA decay module (PELO) is transcriptionally distinct from a nascent chain processing module (LTN1–NEMF). The hippocampus-specific divergence of the processing module suggests region-dependent transcriptional tuning of nascent chain quality control.

Background

Ribosome-associated quality control (RQC) is a conserved surveillance mechanism that detects and resolves ribosomes stalled during translation elongation [1, 2]. Translational stalling can arise from mRNA damage, problematic codon sequences, defective tRNAs, or aberrant mRNA structures, and if unresolved, generates truncated polypeptides that are potentially cytotoxic through aggregation or gain-of-function mechanisms [3, 4].

The canonical RQC pathway proceeds through a defined biochemical sequence. First, the stalled ribosome is recognized and split into its 40S and 60S subunits. In eukaryotes, the protein Pelota (PELO, the mammalian ortholog of yeast Dom34) cooperates with the GTPase HBS1L to catalyze this splitting step, enabling recycling of the 40S subunit and exposing the nascent chain–tRNA complex on the 60S subunit [5, 6]. Second, the E3 ubiquitin ligase Listerin (LTN1, the mammalian ortholog of yeast Ltn1) is recruited to the 60S–nascent chain complex, where it ubiquitinates the exposed nascent polypeptide, targeting it for proteasomal degradation via the AAA-ATPase VCP/p97 and its cofactors [7, 8]. Third, the nuclear export mediator factor (NEMF, the mammalian ortholog of yeast Rqc2) catalyzes the addition of C-terminal alanine–threonine (CAT) tails to the stalled nascent chain, a modification that both facilitates ubiquitination by LTN1 and serves as a signal for protein aggregation quality control [4, 9].

The functional importance of RQC in neuronal tissues is underscored by genetic evidence. Loss-of-function mutations in LTN1 cause the *lister* neurodegeneration phenotype in mice, characterized by progressive motor neuron loss [10]. NEMF mutations have been linked to juvenile neuromuscular disease in both mice and humans [11, 12]. These observations suggest that neurons, with their extreme polarization, high translational demand, and post-mitotic nature, are particularly vulnerable to failures in co-translational quality control.

Despite the well-characterized biochemistry of RQC, virtually nothing is known about how individual RQC components are transcriptionally coordinated with other cellular systems in the human brain. The three core RQC proteins (PELO, LTN1, NEMF) act sequentially on the same stalled ribosomal complex, which might suggest they would be co-regulated as a single transcriptional unit. Alternatively, each component might be independently

coordinated with distinct downstream effector systems, reflecting functional specialization beyond the immediate RQC reaction.

Genome-wide co-expression analysis offers a systematic approach to distinguish between these possibilities. Genes that are consistently co-expressed across individuals within a tissue tend to participate in shared biological processes, as coordinated transcriptional regulation reflects functional relationships that extend beyond direct physical interactions [13, 14]. The Genotype-Tissue Expression (GTEx) project provides a unique resource for such analyses, offering RNA-seq expression data from hundreds of post-mortem human brain samples across multiple anatomical regions [15].

Here, we performed genome-wide co-expression analysis for PELO, LTN1, and NEMF across more than 16,000 genes in five GTEx brain regions. We report that RQC in human brain operates as a two-module system: PELO maintains a transcriptional program distinct from LTN1–NEMF, with the surveillance module (PELO) coordinated with proteostasis, mRNA surveillance, and catabolic machinery, while the nascent chain processing module (LTN1–NEMF) is coordinated with intracellular transport, ubiquitin-mediated proteolysis, and, in the case of NEMF, histone modification. These findings represent the first systematic characterization of genome-wide co-expression architecture for individual RQC components in human tissue.

Methods

Data source and sample selection

Gene expression data were obtained from the Genotype-Tissue Expression (GTEx) project version 8 (dbGaP accession phs000424.v8.p2). Transcripts per million (TPM) values were extracted from the GTEx_Analysis_2017-06-05_v8_RNASeQCv1.1.9_gene_tpm.gct.gz file. Sample attributes were used to identify brain tissue samples. Brain–Frontal Cortex (Brodmann Area 9) was selected as the primary analysis region due to its largest sample size among cortical regions ($n = 209$ samples after quality filtering). Four additional brain regions served as replication cohorts: Putamen (basal ganglia; $n = 205$), Hippocampus ($n = 197$), Nucleus accumbens (basal ganglia; $n = 246$), and Anterior cingulate cortex (BA24; $n = 176$).

Gene filtering and expression thresholds

Genes with median TPM < 1.0 across samples within each brain region were excluded from analysis. Where multiple entries mapped to the same gene symbol, the entry with the highest median expression was retained. In the primary region (BA9), this threshold

retained 16,225 genes for analysis. All three target genes were expressed above threshold in all brain regions examined.

Genome-wide co-expression analysis

For each target gene (PELO, LTN1, NEMF), Pearson correlation coefficients were computed between the target and all other expressed genes across samples within the primary brain region. Correlations were computed on $\log_2(\mathrm{TPM}+1)$-transformed expression values. Genes were ranked by correlation strength (descending), and the top 5% of ranked genes were designated as the co-expression network for each target. In the primary region, this corresponded to 811 genes per target at the following thresholds: PELO ($r \geq 0.899$), LTN1 ($r \geq 0.891$), NEMF ($r \geq 0.874$). The 5% threshold was selected as a balance between capturing functionally coherent co-expression partners and excluding the broad housekeeping signal that dominates at more permissive thresholds, consistent with standard practice in co-expression network analysis [14]; threshold sensitivity across 1–20% is reported below.

Network comparison statistics

Pairwise comparisons between the three co-expression networks were quantified using three complementary metrics. Jaccard similarity index was computed as the ratio of the intersection to the union of the top 5% gene sets for each pair. Fisher's exact test assessed whether the overlap between pairs of top 5% gene sets was significantly greater than expected by chance, given the total number of expressed genes (16,225). Spearman rank correlation was computed on the genome-wide correlation rankings (all 16,222 non-target genes) for each pair of target genes, providing a measure of global network similarity that is independent of the 5% threshold. Three-way overlap analysis identified genes present in all three top 5% sets, genes shared by exactly two targets, and genes unique to each target.

Permutation null model

To contextualize the observed Jaccard indices against the natural co-expression structure of the genome, a permutation null distribution was constructed. For each of 10,000 iterations, two genes were randomly selected (without replacement) from the 16,222 non-target genes. Genome-wide Pearson correlations were computed for both genes against all other genes using vectorized dot-product Pearson on pre-standardized expression matrices, and the top 5% co-expression networks were defined and compared using the Jaccard index. This procedure generates a null distribution that captures the full range of

pairwise co-expression overlap expected from random gene pairs, including the inflated overlap arising from housekeeping co-regulation. Empirical p-values were computed as the proportion of null Jaccard values equaling or exceeding each observed value. Z-scores were computed as (observed − null mean) / null standard deviation.

Bootstrap confidence intervals

Stability of the Jaccard indices and Spearman rank correlations was assessed via nonparametric bootstrap resampling. For each of 1,000 iterations, the 209 BA9 samples were resampled with replacement, and the complete co-expression analysis was repeated: genome-wide Pearson correlations were recomputed for all three target genes using vectorized operations, top 5% gene sets were redefined, and pairwise Jaccard indices and genome-wide Spearman rank correlations were calculated. Bias-corrected percentile 95% confidence intervals were derived from the bootstrap distributions. Non-overlap of the LTN1–NEMF lower confidence bound with the maximum upper confidence bound of either PELO pair was used as a formal test of two-module separation robustness.

Threshold sensitivity analysis

To assess whether the two-module architecture depended on the specific choice of a 5% network threshold, pairwise Jaccard indices were recomputed at eight threshold values: 1%, 2%, 3%, 5%, 7%, 10%, 15%, and 20% of the genome-wide correlation ranking. For each threshold, a partition ratio was computed as the LTN1–NEMF Jaccard divided by the maximum of the two PELO-pair Jaccards, with values above 2.0 indicating strong two-module separation and values above 1.5 indicating moderate separation. This analysis complements the threshold-independent Spearman rank correlations by demonstrating robustness across a range of discrete network definitions.

Gene Ontology enrichment

Functional enrichment analysis was performed using the gProfiler g:GOSt tool [16] via the public REST API. Enrichment was tested against Gene Ontology Biological Process (GO:BP), Molecular Function (GO:MF), Cellular Component (GO:CC), KEGG pathway, and Reactome pathway databases. Background was set to the full list of expressed genes in the primary brain region (16,225 genes). Statistical significance was determined using the g:SCS multiple testing correction method native to gProfiler, with a significance threshold of adjusted $p < 0.05$. Enrichment was performed separately for (i) the full top 5% gene set of each target, (ii) genes unique to each target, (iii) genes shared by all three targets, and (iv) genes shared pairwise.

Multi-region replication

The complete genome-wide co-expression analysis was independently repeated for each target gene in four additional brain regions. Cross-region consistency was quantified by computing Spearman rank correlations of the genome-wide correlation rankings between all pairs of brain regions for each target gene, restricted to the common gene universe expressed across all five regions. Replication was characterized by the median, minimum, and maximum pairwise ρ across all ten region pairs for each target, and the proportion of pairs exceeding $\rho = 0.80$. Complete pairwise decomposition was performed to identify specific region pairs exhibiting reduced network similarity.

Cell-type deconvolution

To assess whether observed co-expression patterns were driven by cell-type composition variation across samples, a reference-based deconvolution approach was employed. Expression levels of 53 established cell-type marker genes spanning six major brain cell types (neurons, astrocytes, oligodendrocytes, microglia, endothelial cells, and oligodendrocyte precursor cells) were used to estimate cell-type proportions for each sample. Partial correlation analysis was then performed, computing the Pearson correlation between each target gene and all other genes while controlling for the six estimated cell-type proportion variables. Genes were re-ranked based on partial correlation coefficients, and the top 5% were compared with the unadjusted results.

Covariate and agonal stress adjustment

To control for potential demographic confounds, partial correlation analysis was repeated with donor age and sex as covariates. To address the concern that co-expression patterns in post-mortem tissue may reflect agonal stress rather than constitutive biology, additional partial correlation models were computed controlling for Hardy Scale (DTHHRDY; death circumstance classification) and ischemic time (SMTSISCH; interval between death and tissue preservation). Four sensitivity models were evaluated: (i) Hardy Scale alone, (ii) ischemic time alone, (iii) Hardy Scale + ischemic time, and (iv) the full model (age + sex + Hardy Scale + ischemic time). Genome-wide rank preservation between unadjusted and adjusted co-expression rankings was quantified by Spearman ρ for each target gene.

Software and reproducibility

All analyses were performed in Python 3.10 using pandas (v1.5.3), scipy (v1.11.4), and numpy (v1.24.3) in the Google Colaboratory environment. Visualization was performed with matplotlib (v3.7.1). Complete analysis code is publicly available at https://github.com/nwharbert8-ui/rqc-two-module and archived at https://doi.org/10.5281/zenodo.18616697.

Results

PELO, LTN1, and NEMF are widely expressed in human brain

All three RQC components were expressed above the median TPM > 1.0 threshold across all five brain regions examined. In the primary analysis region (Brain–Frontal Cortex, BA9; n = 209 samples), 16,225 genes met the expression threshold and were included in the genome-wide co-expression screen.

LTN1 and NEMF share a co-expression network distinct from PELO

Pairwise comparison of the top 5% co-expression networks revealed a striking asymmetry among the three RQC components (Table 1). LTN1 and NEMF shared 512 of their top 811 genes (Jaccard = 0.461; Fisher's exact OR = 86.55, $p < 10^{-324}$), with a genome-wide rank correlation of ρ = 0.986. In contrast, PELO shared only 230 genes with LTN1 (Jaccard = 0.165; OR = 10.10, $p = 7.14 \times 10^{-116}$) and 219 genes with NEMF (Jaccard = 0.156; OR = 9.26, $p = 5.04 \times 10^{-105}$). While all pairwise overlaps were statistically significant by Fisher's exact test, the magnitude of the LTN1–NEMF overlap was approximately three-fold greater than the PELO overlaps with either gene, indicating that RQC partitions into two transcriptional modules rather than operating as a single co-regulated unit (Fig. 1).

Table 1 Pairwise comparison of PELO, LTN1, and NEMF co-expression networks in Brain–Frontal Cortex (BA9).

Gene A	Gene B	Jaccard	Shared genes	Fisher OR	Fisher p	Rank ρ
PELO	LTN1	0.165	230	10.10	7.14 × 10□□□□	0.913
PELO	NEMF	0.156	219	9.26	5.04 × 10□□□□	0.900
LTN1	NEMF	0.461	512	86.55	< 10□□□□	0.986

Permutation null model confirms exceptional LTN1–NEMF convergence

The permutation null distribution (10,000 random gene pairs) had a mean Jaccard of 0.055 (SD = 0.092; median = 0.009), reflecting the right-skewed distribution expected when most gene pairs share minimal co-expression overlap but a minority of functionally related pairs exhibit substantial convergence (Additional file 5). The LTN1–NEMF Jaccard of

0.461 was significantly above the null (z = 4.4; empirical p = 0.006), exceeding the 99th percentile of the null distribution (0.419) and confirming that the convergence between LTN1 and NEMF is exceptional even against the natural co-expression landscape of the genome. In contrast, the PELO–LTN1 (J = 0.165; z = 1.2; p = 0.11) and PELO–NEMF (J = 0.156; z = 1.1; p = 0.12) Jaccards did not reach significance against the permutation null (Fig. S1).

These results sharpen the two-module interpretation. The Fisher's exact test evaluates overlap against the assumption of independent random gene selection, yielding highly significant p-values for all three pairs. The permutation null evaluates overlap against the empirical co-expression structure of the genome, where correlated gene pairs naturally produce elevated Jaccard values. That LTN1–NEMF is significant under both frameworks, while the PELO pairs are significant only under the independence assumption, indicates that PELO's overlap with LTN1 and NEMF falls within the range of typical genomic co-expression, whereas the LTN1–NEMF convergence represents a biologically exceptional level of transcriptional coordination.

Bootstrap confidence intervals confirm non-overlapping module separation

Bootstrap resampling (1,000 iterations) produced 95% confidence intervals that formally confirmed the robustness of the two-module partition (Fig. S2). The LTN1–NEMF Jaccard 95% CI [0.357, 0.546] did not overlap with either PELO pair CI (PELO–LTN1: [0.095, 0.242]; PELO–NEMF: [0.082, 0.227]). The lower bound of the LTN1–NEMF CI (0.357) exceeded the upper bound of both PELO pair CIs (0.242) by a margin of 0.115, confirming that the two-module separation is robust to sampling variability. Spearman rank correlation CIs showed similar separation (LTN1–NEMF: [0.971, 0.991]; PELO–LTN1: [0.851, 0.942]; PELO–NEMF: [0.830, 0.936]; Additional file 7).

Two-module architecture is robust across network thresholds

The two-module partition was maintained across all eight network thresholds tested, from 1% to 20% of the genome-wide ranking (Fig. S3; Additional file 8). The LTN1–NEMF Jaccard exceeded both PELO pair Jaccards at every threshold, with partition ratios ranging from 8.78× at the most stringent threshold (1%) to 1.64× at the most permissive (20%). Strong separation (partition ratio > 2.0×) was observed at thresholds from 1% through 7%, with moderate separation (1.64×–1.97×) at 10–20%. The monotonic attenuation of the partition ratio at more permissive thresholds is expected, as increasingly broad networks

accumulate housekeeping and ubiquitous co-expression partners that converge across all gene pairs. The threshold-independent Spearman rank correlations (ρ = 0.986 for LTN1–NEMF versus ρ = 0.900–0.913 for PELO pairs) corroborate this pattern without reliance on any threshold.

Three-way overlap analysis

Three-way overlap analysis confirmed this architecture (Table 2). Of the 1,638 unique genes represented across all three top 5% networks, only 166 genes (10.1%) were shared by all three components. An additional 346 genes were shared exclusively between LTN1 and NEMF, while only 64 and 53 genes were shared exclusively between PELO–LTN1 and PELO–NEMF, respectively. PELO contributed 528 unique genes not found in either LTN1 or NEMF networks, whereas LTN1 and NEMF contributed only 235 and 246 unique genes, respectively (Fig. 1b).

Table 2 Three-way overlap analysis of top 5% co-expression networks.

Category	Gene count	% of total
All three (PELO ∩ LTN1 ∩ NEMF)	166	10.1
LTN1 ∩ NEMF only	346	21.1
PELO ∩ NEMF only	53	3.2
PELO ∩ LTN1 only	64	3.9
PELO unique	528	32.2
LTN1 unique	235	14.3
NEMF unique	246	15.0

The shared RQC core is enriched for ubiquitin–proteasome function

Gene Ontology analysis of the 166 genes shared by all three RQC components revealed enrichment for macromolecule catabolic process (GO:BP; $p = 5.31 \times 10^{-10}$), protein modification by small protein conjugation (GO:BP; $p = 1.44 \times 10^{-8}$), ubiquitin-protein transferase activity (GO:MF; $p = 8.61 \times 10^{-4}$), and cytosol localization (GO:CC; $p = 1.89 \times 10^{-14}$). This shared core represents the expected ubiquitin–proteasome degradation machinery that all three RQC components must interface with (Fig. 2).

PELO’s unique network reflects surveillance–mRNA decay function

The 528 genes unique to PELO’s co-expression network were enriched for protein metabolic process (GO:BP; $p = 9.35 \times 10^{-22}$), macromolecule catabolic process (GO:BP; $p = 4.08 \times 10^{-20}$), and catabolic process broadly (GO:BP; $p = 1.95 \times 10^{-15}$). Molecular function analysis identified protein binding ($p = 3.38 \times 10^{-25}$) and catalytic activity acting on RNA ($p = 9.85 \times 10^{-7}$). Reactome analysis identified metabolism of RNA ($p = 1.58 \times 10^{-20}$) and antigen processing via ubiquitination and proteasome degradation ($p = 1.20 \times 10^{-10}$) as the most significantly enriched pathways.

The RNA metabolism enrichment (Reactome $p = 1.58 \times 10^{-20}$) combined with catalytic activity acting on RNA ($p = 9.85 \times 10^{-7}$) positions PELO not merely as a proteostasis factor but as a node at the interface of translational surveillance and mRNA quality control. This is consistent with PELO’s known dual function: ribosome splitting by Pelota-HBS1L simultaneously frees the nascent chain for ubiquitin-dependent degradation and releases the mRNA for nonstop decay (NSD) and no-go decay (NGD) [5, 6, 18].

LTN1–NEMF module coordinates with intracellular transport and transcriptional regulation

Analysis of genes unique to LTN1 (235 genes) revealed enrichment for macromolecule modification (GO:BP; $p = 3.70 \times 10^{-4}$), organelle organization (GO:BP; $p = 3.85 \times 10^{-4}$), small molecule binding (GO:MF; $p = 2.91 \times 10^{-8}$), and metal ion binding (GO:MF; $p = 6.32 \times 10^{-7}$). The full LTN1 network showed strong enrichment for membrane trafficking (Reactome; $p = 2.13 \times 10^{-11}$), vesicle-mediated transport (Reactome; $p = 1.93 \times 10^{-8}$), and ubiquitin-mediated proteolysis (KEGG; $p = 2.88 \times 10^{-6}$).

Genes unique to NEMF (246 genes) were enriched for regulation of nucleobase-containing compound metabolic process (GO:BP; $p = 1.00 \times 10^{-8}$), regulation of RNA metabolic process (GO:BP; $p = 3.58 \times 10^{-6}$), and notably, histone modifying activity (GO:MF; $p = 2.16 \times 10^{-5}$). NEMF’s unique association with histone modification was independently supported through custom gene set enrichment analysis (epigenetic regulators: 5.4-fold enrichment, $p = 1.6 \times 10^{-7}$ in the primary region).

Multi-region replication reveals hippocampus-specific divergence of the processing module

The two-module architecture was broadly consistent across all five brain regions, though the degree of cross-region consistency differed between modules in a manner that reinforces their transcriptional independence (Table 3). PELO's surveillance module displayed uniformly high cross-region replication, with all 10 pairwise correlations exceeding ρ = 0.80 (range: ρ = 0.84–0.97; median ρ = 0.89). The highest PELO correlation was between BA9 and BA24 (ρ = 0.968), consistent with the anatomical proximity of these cortical regions (Fig. 3).

The LTN1–NEMF processing module exhibited high median cross-region consistency (LTN1: median ρ = 0.89; NEMF: median ρ = 0.90) but greater variability across region pairs. LTN1 produced 8/10 pairwise correlations above ρ = 0.80 (80%), while NEMF produced 6/10 (60%). Pairwise decomposition (Additional file 6) revealed that all six sub-0.80 pairs involved the hippocampus: LTN1 showed reduced cross-region consistency for BA9–Hippocampus (ρ = 0.680) and Hippocampus–BA24 (ρ = 0.691); NEMF showed reduced consistency for BA9–Hippocampus (ρ = 0.668), Putamen–Hippocampus (ρ = 0.788), Hippocampus–NAC (ρ = 0.781), and Hippocampus–BA24 (ρ = 0.703). No non-hippocampal pair fell below ρ = 0.80 for any RQC gene. The maximum cross-region correlations for LTN1 (ρ = 0.976) and NEMF (ρ = 0.974), both occurring between BA9–BA24, exceeded that of PELO (ρ = 0.968), confirming that the processing module replicates as strongly as the surveillance module between anatomically proximate cortical regions.

This hippocampus-specific divergence is consistent with the known transcriptional and cellular distinctiveness of the hippocampal formation. The hippocampus has a unique laminar architecture with specialized neuronal populations (granule cells, pyramidal cells in CA1–CA3) and distinct patterns of synaptic plasticity, activity-dependent translation, and proteostatic demand. That the surveillance module (PELO) replicates uniformly across all regions including hippocampus, while the processing module (LTN1–NEMF) diverges specifically at hippocampus, suggests that upstream ribosome surveillance is governed by a conserved transcriptional program whereas nascent chain processing is transcriptionally tuned to region-specific demands.

NEMF's enrichment for epigenetic regulators was significant in the primary region and directionally consistent in all four replication regions, with the Nucleus accumbens independently reaching significance (7/52 epigenetic regulators in top 5%; 2.7-fold enrichment; p = 0.014).

Table 3 Cross-region Spearman rank correlation ranges for each RQC gene.

Gene	Min ρ	Max ρ	Median ρ	All > 0.80
PELO	0.84	0.97	0.89	Yes
LTN1	0.68	0.98	0.89	No
NEMF	0.67	0.97	0.90	No

Agonal stress sensitivity and covariate adjustment

Cell-type deconvolution analysis confirmed that the primary co-expression patterns were not driven by cell-type composition variation. Covariate adjustment for donor age and sex produced minimal changes in the genome-wide rankings.

Agonal stress sensitivity analysis demonstrated that the two-module architecture reflects constitutive biology rather than post-mortem artifacts. The most conservative full model (age + sex + Hardy Scale + ischemic time) yielded rank preservation of $\rho \geq 0.998$ (PELO: $\rho = 0.998$; LTN1: $\rho = 0.998$; NEMF: $\rho = 0.998$). The two-module partition ratio (LTN1–NEMF Jaccard / max PELO Jaccard) ranged from 2.70× to 2.82× across all five sensitivity models, confirming that the two-module partition is not attributable to differential agonal stress responses.

Discussion

This study presents a systematic genome-wide co-expression analysis of the three core ribosome-associated quality control components in human brain tissue. Our principal finding is that the three core RQC proteins—PELO, LTN1, and NEMF—do not form a single co-regulated transcriptional unit despite acting sequentially on the same stalled ribosomal complex. Instead, RQC operates as a two-module system in human brain: a surveillance–mRNA decay module centered on PELO and a nascent chain processing module defined by the tightly co-regulated LTN1–NEMF pair.

Two modules within a single biochemical pathway

The most striking result is the magnitude of the LTN1–NEMF co-expression overlap (Jaccard = 0.461; $\rho = 0.986$) compared with the relatively modest overlap of either gene with PELO (Jaccard = 0.156–0.165; $\rho = 0.900$–0.913). The permutation null model provides formal statistical context: the LTN1–NEMF Jaccard exceeds the 99th percentile of the genomic null distribution (0.419) with an empirical p-value of 0.006, whereas the PELO pair Jaccards fall within the normal range of genomic co-expression ($p > 0.10$). This distinction is critical: Fisher’s exact test indicates that all three pairwise overlaps exceed chance expectation under an independence assumption, but the permutation test reveals

that only LTN1–NEMF exceeds what is expected from the natural co-expression architecture of the genome. Bootstrap 95% confidence intervals confirm that the separation is robust to sampling variability, with non-overlapping intervals between the LTN1–NEMF module (CI [0.357, 0.546]) and both PELO pairs (maximum upper CI = 0.242). The two-module partition is maintained across all network thresholds tested (1–20%), with partition ratios ranging from 8.78× at 1% to 1.64× at 20%, demonstrating that the architecture is not an artifact of the specific 5% threshold.

Biochemically, PELO acts first (ribosome splitting), followed by LTN1 (ubiquitination) and NEMF (CAT-tailing), with the latter two operating on the same 60S–nascent chain complex. The co-expression data suggest that this biochemical relationship is reflected in transcriptional coordination: LTN1 and NEMF, which physically co-occupy the 60S subunit, are transcriptionally co-regulated with nearly identical genome-wide profiles, while PELO, which acts at an earlier and mechanistically distinct step, is regulated by a partially independent transcriptional program. This architecture is consistent with the modular organization principle observed in other multi-step pathways, where sensors, mediators, and effectors occupy the same pathway but are embedded in distinct co-expression modules reflecting their specific downstream functions [17].

PELO as bridge between translational surveillance and mRNA decay

PELO’s unique co-expression partners are overwhelmingly enriched for both protein catabolism ($p = 9.35 \times 10^{-22}$) and RNA metabolism (Reactome $p = 1.58 \times 10^{-20}$; catalytic activity acting on RNA $p = 9.85 \times 10^{-7}$). This dual enrichment is consistent with PELO’s biochemical mechanism: when Pelota-HBS1L splits a stalled ribosome, the reaction generates two products that must be separately processed: the nascent chain on the 60S subunit (protein fate) and the aberrant mRNA released from the 40S subunit (RNA fate). The mRNA is subsequently targeted for degradation by the NSD and NGD pathways [18, 19]. Our co-expression data suggest that PELO is transcriptionally embedded within this broader mRNA surveillance network, not merely the protein quality control arm.

KEGG enrichment for amyotrophic lateral sclerosis ($p = 1.89 \times 10^{-8}$) and spinocerebellar ataxia ($p = 1.08 \times 10^{-5}$) is notable because PELO itself has not been directly implicated in these neurodegenerative conditions. The co-expression data suggest that PELO’s transcriptional program is embedded within the broader proteostasis–RNA surveillance network that is disrupted in motor neuron and cerebellar diseases. The EIF2S1 co-expression network exhibits a convergence with PELO ($J = 0.526$) that exceeds the LTN1–NEMF convergence reported here ($J = 0.461$), suggesting a closed-loop transcriptional axis linking translation initiation surveillance (eIF2α) to ribosome rescue (Pelota) that we characterize in a companion study (Harbert, D.H., manuscript in preparation).

NEMF’s connection to histone modification

The enrichment of NEMF’s unique co-expression partners for histone modifying activity ($p = 2.16 \times 10^{-5}$), independently confirmed through custom gene set analysis (5.4-fold enrichment; $p = 1.6 \times 10^{-7}$), is among the most unexpected findings. No direct biochemical interaction between NEMF and chromatin-modifying enzymes has been reported. While no direct mechanistic link is established by co-expression data, a plausible biological rationale exists: CAT-tailing by NEMF tags aberrant nascent polypeptides for degradation, and a substantial fraction of ribosome-translated proteins are transcription factors and chromatin modifiers. The observed transcriptional coordination between NEMF and epigenetic regulators may reflect shared upstream regulatory control or a quality control requirement to ensure that defective chromatin-modifying proteins are effectively degraded before they can alter the epigenomic landscape. Distinguishing between these possibilities will require experimental approaches beyond co-expression analysis.

Constitutive architecture confirmed by agonal stress analysis

Adjustment for Hardy Scale, ischemic time, and the full covariate model all produced rank preservation exceeding $\rho = 0.998$, and the two-module partition ratio remained between 2.70× and 2.82× across all models. The robustness to agonal covariates indicates that the transcriptional separation of PELO from LTN1–NEMF reflects constitutive brain architecture rather than differential stress responses during the peri-mortem period.

Hippocampus-specific divergence of the processing module

The multi-region analysis revealed that all six sub-0.80 cross-region correlations for LTN1 and NEMF involved hippocampus, while PELO replicated uniformly across all regions including hippocampus (10/10 pairs > 0.80). This pattern is striking because hippocampus has sufficient sample size ($n = 197$) to produce stable correlation estimates, and PELO achieves strong replication at hippocampus (lowest hippocampal pair for PELO: PUT–HIP, $\rho = 0.839$), ruling out a general quality issue with the hippocampal dataset. The divergence is specific to the processing module’s co-expression neighborhood, not to hippocampal data quality.

Several features of hippocampal biology may explain this selective divergence. The hippocampus has uniquely high rates of activity-dependent translation associated with long-term potentiation and memory consolidation, generating region-specific demands on co-translational quality control. The hippocampal formation also contains highly

specialized neuronal populations (dentate gyrus granule cells, CA1–CA3 pyramidal neurons) with distinct proteostatic and transcriptional profiles. Additionally, adult neurogenesis in the dentate gyrus creates a population of immature neurons with different quality control requirements than the post-mitotic neurons that predominate in other brain regions. These features could collectively drive hippocampus-specific transcriptional tuning of the nascent chain processing machinery (LTN1, NEMF) while leaving the upstream surveillance decision (PELO) governed by a conserved program. This interpretation generates testable predictions: single-cell RNA-seq analysis of hippocampal neuronal subtypes should reveal cell-type-specific variation in LTN1 and NEMF co-expression neighborhoods that is absent for PELO. We note that bulk tissue deconvolution cannot fully resolve whether the hippocampal divergence reflects genuine region-specific regulatory biology, cell-type compositional differences not captured by our marker-gene panel, or an interaction between these factors; resolving this will require spatially resolved or single-nucleus transcriptomic analysis.

Implications for RQC biology

The two-module architecture is consistent with a model in which cells could independently modulate the ribosome surveillance arm (PELO) and the nascent chain processing arm (LTN1–NEMF), potentially enabling differential responses to translational stress versus protein misfolding stress. The tight co-regulation of LTN1 and NEMF ($\rho = 0.986$; bootstrap CI [0.971, 0.991]) implies strong selective pressure to maintain stoichiometric balance between ubiquitination and CAT-tailing. PELO’s integration with both the proteostasis and mRNA surveillance networks suggests that ribosome rescue is a bifurcation point that determines the fate of both the protein and the mRNA, consistent with a more interconnected quality control architecture than is typically emphasized in current pathway models.

Interpretive framework

We emphasize that co-expression analysis identifies transcriptional coordination, not direct regulatory or physical interactions. The modular architecture described here reflects shared higher-order transcriptional control programs rather than demonstrated biochemical mechanisms. Genes that are consistently co-expressed across individuals within a tissue are inferred to participate in shared or functionally related biological processes, but the nature of that relationship—whether mediated by common transcription factors, chromatin state, signaling cascades, or indirect coupling—cannot be resolved from expression data alone. The term "module" as used throughout this study refers to a statistically defined co-expression partition, not to a functionally validated regulatory circuit. All mechanistic interpretations offered in the Discussion represent hypotheses generated by the co-expression data that require independent experimental validation.

Limitations

Several limitations should be considered. First, co-expression analysis identifies transcriptional coordination but cannot establish causality; the observed architecture requires experimental validation. Second, GTEx data are derived from bulk tissue RNA-seq; while cell-type deconvolution suggests the major patterns are not driven by composition variation, single-cell resolution would provide cell-type-specific network maps. Third, while the two-module partition is robust across thresholds from 1% to 20% and confirmed by threshold-independent Spearman correlations, bootstrap CIs, and permutation testing, different analytical approaches (e.g., weighted gene co-expression network analysis) could yield complementary perspectives. Future work incorporating weighted gene co-expression network analysis (WGCNA) or independent developmental datasets such as BrainSpan could provide additional validation of the modular partition through orthogonal analytical frameworks. Fourth, the PELO pair Jaccards, while highly significant by Fisher's exact test, did not reach significance against the permutation null ($p > 0.10$); this indicates that PELO's co-expression overlap with LTN1 and NEMF, though greater than expected by chance under independence, falls within the range of normal genomic co-expression. Fifth, the hippocampus-specific divergence of the processing module warrants further investigation to determine whether this reflects genuine region-specific biology, differences in cellular composition not captured by bulk deconvolution, or hippocampus-specific technical factors in the GTEx dataset.

Conclusions

Genome-wide co-expression analysis of RQC components in human brain reveals a two-module transcriptional architecture that is robust to agonal stress, cell-type composition, demographic covariates, sampling variability (bootstrap 95% CIs non-overlapping), network threshold selection (partition ratio 1.64×–8.78× across 1–20%), and permutation-based null modeling (LTN1–NEMF $p = 0.006$). PELO operates within a surveillance–mRNA decay module enriched for protein catabolism and RNA metabolism, functioning as a transcriptional bridge between translational surveillance and mRNA quality control. LTN1 and NEMF share a near-identical nascent chain processing module coordinated with intracellular transport, ubiquitin-mediated proteolysis, and histone modification. The hippocampus-specific divergence of the processing module provides evidence for region-dependent transcriptional tuning of nascent chain quality control in human brain.

Abbreviations

BA9
Brodmann Area 9
BA24

Brodmann Area 24
CAT
C-terminal alanine–threonine
CI
confidence interval
GO
Gene Ontology
GTEx
Genotype-Tissue Expression
KEGG
Kyoto Encyclopedia of Genes and Genomes
NGD
no-go decay
NMD
nonsense-mediated decay
NSD
nonstop decay
OR
odds ratio
RQC
ribosome-associated quality control
TPM
transcripts per million
ρ
Spearman rank correlation coefficient

Declarations

Ethics approval and consent to participate

Not applicable. This study exclusively analyzed publicly available, de-identified gene expression data from the Genotype-Tissue Expression (GTEx) project (dbGaP accession phs000424.v8.p2). The GTEx project was reviewed and approved by each biospecimen source site's Institutional Review Board, and all donors or their next-of-kin provided written informed consent. As this study involved only secondary analysis of existing de-identified data with no direct interaction with human subjects, it is exempt from further ethics review under 45 CFR § 46.104(d)(4) of the U.S. Department of Health and Human Services regulations governing the protection of human subjects. No additional Institutional Review Board approval was required.

Consent for publication

Not applicable.

Competing interests

The author declares that they have no competing interests.

Funding

This research was conducted independently by Inner Architecture LLC. No external funding was received.

Author Contribution

DH conceived and designed the study, developed the analytical pipeline, performed all computational analyses, interpreted the results, generated all figures, and wrote the manuscript. The author read and approved the final manuscript.

Acknowledgements

The author acknowledges the Genotype-Tissue Expression (GTEx) Project, supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.

Data Availability

The GTEx v8 dataset is publicly available through the GTEx Portal (https://gtexportal.org/) under dbGaP accession phs000424.v8.p2. All analysis code is publicly available at https://github.com/nwharbert8-ui/rqc-two-module and archived at https://doi.org/10.5281/zenodo.18616697. Processed co-expression rankings, network comparison statistics, Gene Ontology enrichment results, the complete list of 166 shared RQC core genes (Additional file 4), and all intermediate data files are included in the repository.

References

1. Brandman O, Hegde RS. Ribosome-associated protein quality control. Nat Struct Mol Biol. 2016;23:7–15.

2. Joazeiro CAP. Mechanisms and functions of ribosome-associated protein quality control. Nat Rev Mol Cell Biol. 2019;20:368–83.
3. Choe YJ, Park SH, Hassemer T, Körner R, Hartl FU, Raber L, et al. Failure of RQC machinery causes protein aggregation and proteotoxic stress. Nature. 2016;531:191–5.
4. Yonashiro R, Tahara EB, Bengtson MH, et al. The Rqc2/Tae2 subunit of the RQC complex marks ribosome-stalled nascent polypeptide chains for aggregation. eLife. 2016;5:e11794.
5. Shoemaker CJ, Eyler DE, Green R. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. Science. 2010;330:369–72.
6. Pisareva VP, Skabkin MA, Hellen CU, Pestova TV, Pisarev AV. Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. EMBO J. 2011;30:1804–17.
7. Bengtson MH, Joazeiro CAP. Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. Nature. 2010;467:470–3.
8. Defenouillere Q, Yao Y, Mouaikel J, et al. Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. Proc Natl Acad Sci U S A. 2013;110:5046–51.
9. Shen PS, Park J, Qin Y, et al. Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. Science. 2015;347:75–8.
10. Chu J, Hong NA, Masuda CA, et al. A mouse forward genetics screen identifies LISTERIN as an E3 ubiquitin ligase involved in neurodegeneration. Proc Natl Acad Sci U S A. 2009;106:2097–103.
11. Martin PB, Kigoshi-Tansho Y, Sher RB, et al. NEMF mutations that impair ribosome-associated quality control are associated with neuromuscular disease. Nat Commun. 2020;11:4625.
12. Ahmed A, Wang M, Tran T, Joazeiro CAP. NEMF mutations cause LISTERIN-mediated proteotoxicity and atypical progressive neurodegeneration. Neuron. 2021;109:2721–35.
13. Stuart JM, Segal E, Koller D, Kim SK. A gene-coexpression network for global discovery of conserved genetic modules. Science. 2003;302:249–55.
14. van Dam S, Vösa U, Franke L, de Magalhães JP. Gene co-expression analysis for functional classification and gene–disease predictions. Brief Bioinf. 2018;19:575–92.
15. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020;369:1318–30.
16. Raudvere U, Kolberg L, Kuzmin I, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 2019;47:W191–8.

17. Pearl LH, Schierz AC, Ward SE, Al-Lazikani B, Pearl FMG. Therapeutic opportunities within the DNA damage response. Nat Rev Cancer. 2015;15:166–80.
18. Shoemaker CJ, Green R. Translation drives mRNA quality control. Nat Struct Mol Biol. 2012;19:594–601.
19. Doma MK, Parker R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature. 2006;440:561–4.

Figures

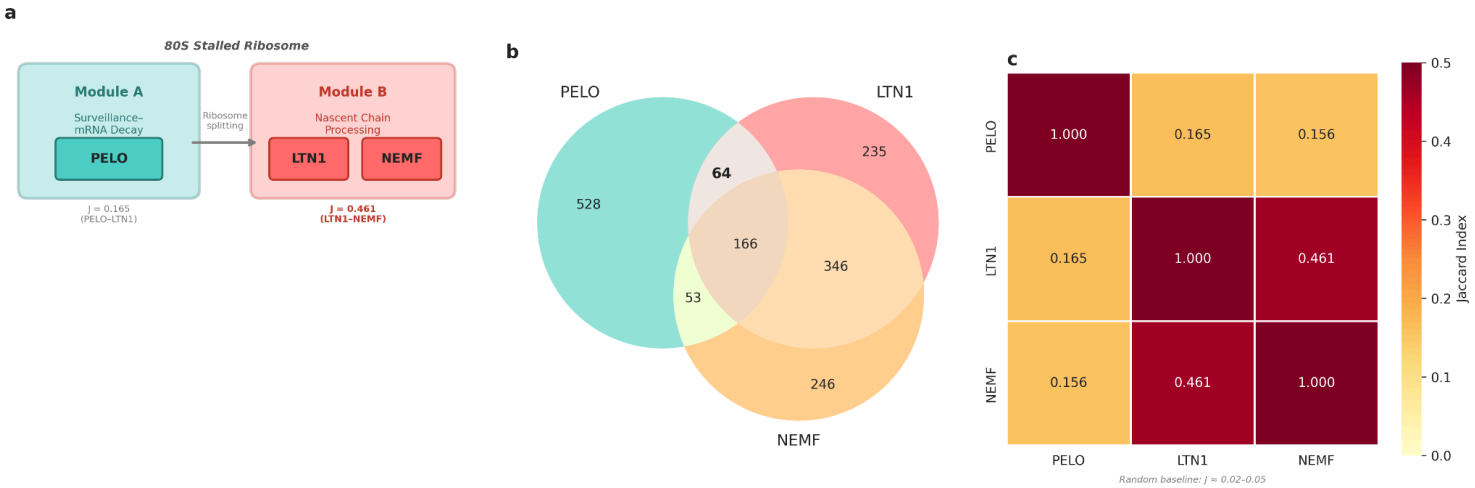


Figure 1

Two-module architecture of ribosome-associated quality control in human brain. **a** Schematic of the canonical RQC pathway showing the sequential action of PELO (ribosome splitting), LTN1 (ubiquitination), and NEMF (CAT-tailing). The dashed box indicates the high-overlap LTN1–NEMF module (Module B; Jaccard = 0.461), while PELO operates as a distinct surveillance–mRNA decay module (Module A). **b** Three-way Venn diagram of top 5% co-expression network overlap. **c** Heatmap of pairwise Jaccard similarity indices.

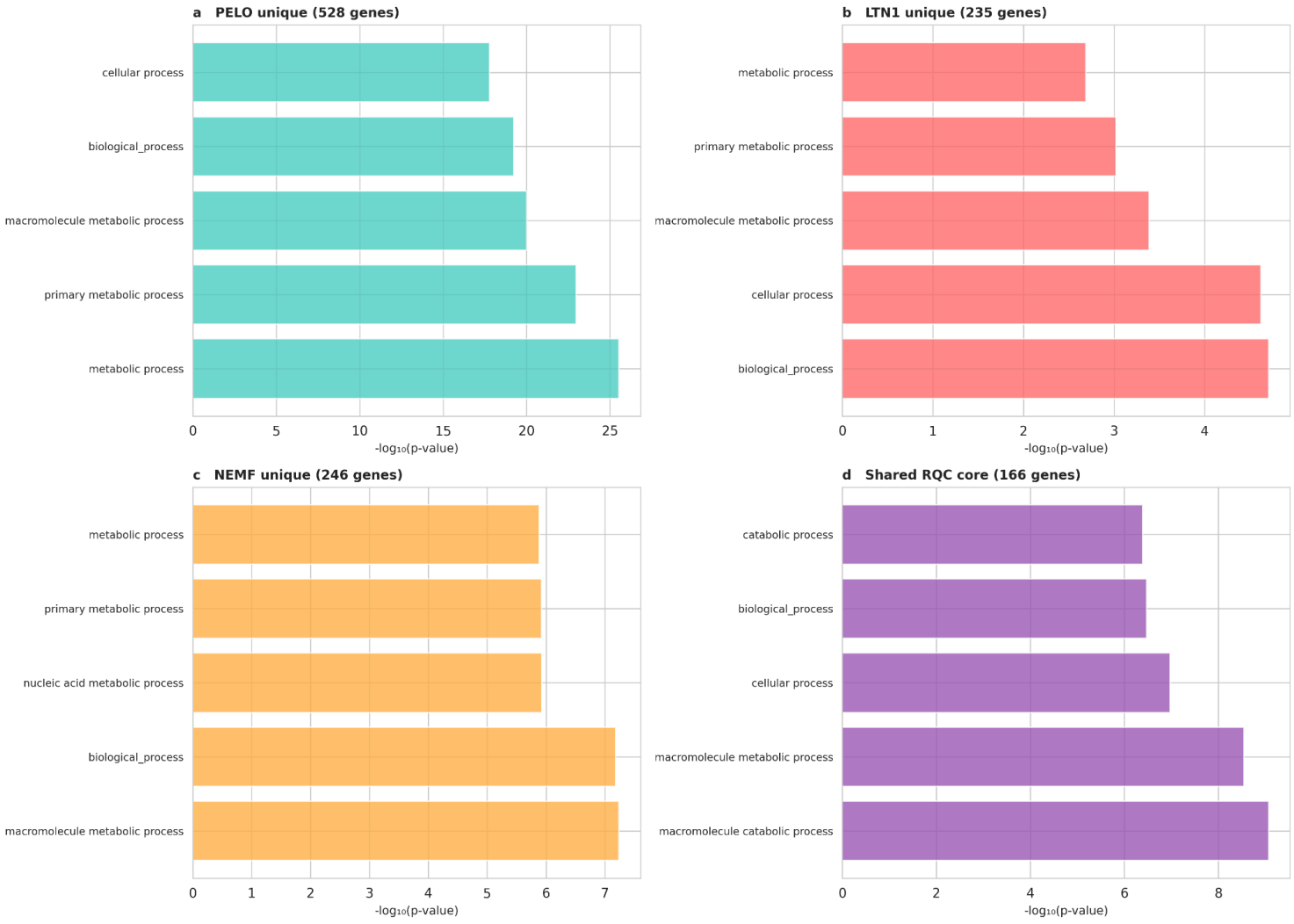


Figure 2

Gene Ontology Biological Process enrichment of unique gene sets. Top five GO:BP terms for genes unique to **a** PELO (528 genes), **b** LTN1 (235 genes), **c** NEMF (246 genes), and **d** the shared RQC core (166 genes).

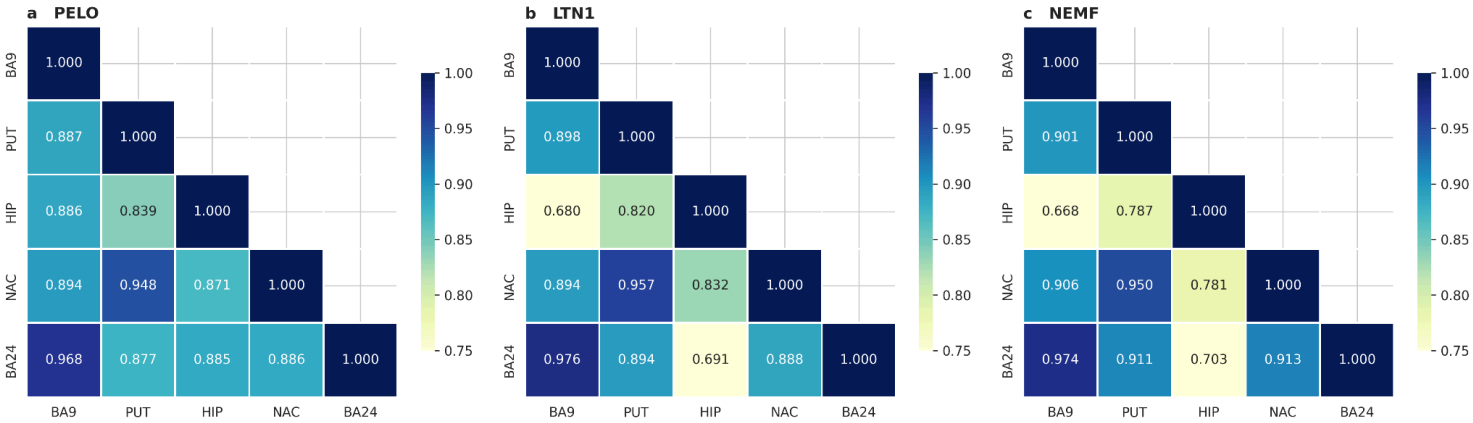


Figure 3

Multi-region replication of co-expression patterns. Cross-region Spearman rank correlation heatmaps for **a** PELO, **b** LTN1, and **c** NEMF across five brain regions. PELO exhibits uniformly high cross-region consistency (all pairwise ρ > 0.80). LTN1 and NEMF show hippocampus-specific divergence, with all sub-0.80 pairs involving hippocampus (LTN1 minimum ρ = 0.68; NEMF minimum ρ = 0.67). Color scale: ρ = 0.60–1.00.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS1permutationnull.tiff
- FigureS2bootstrapjaccard.tiff
- FigureS3thresholdsensitivity.tiff
- TableS1genomewiderankings.csv
- TableS2enrichmentall.csv
- TableS3agonaljaccard.csv
- TableS3customenrichment.csv
- TableS3sensitivityrankpreservation.csv
- TableS4shared166genes.csv
- TableS5nulldistribution.csv
- TableS6crossregionallpairs.csv
- TableS7bootstrapCIs.csv
- TableS8thresholdsensitivity.csv
- Table1pairwisecomparison.csv
- Table2threewayoverlap.csv
- Table3crossregionranges.csv
- Supplementaryinformation.docx