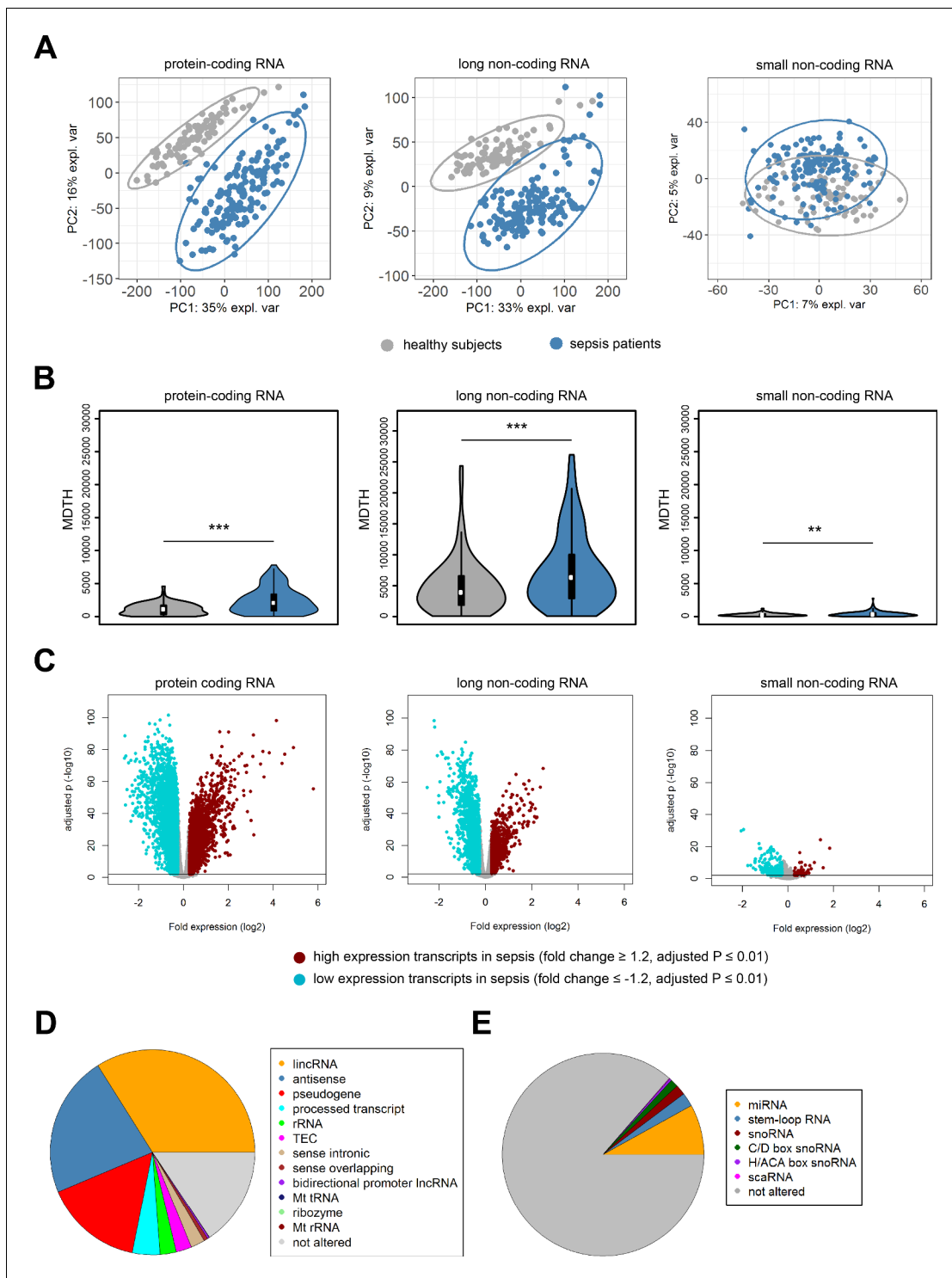


---

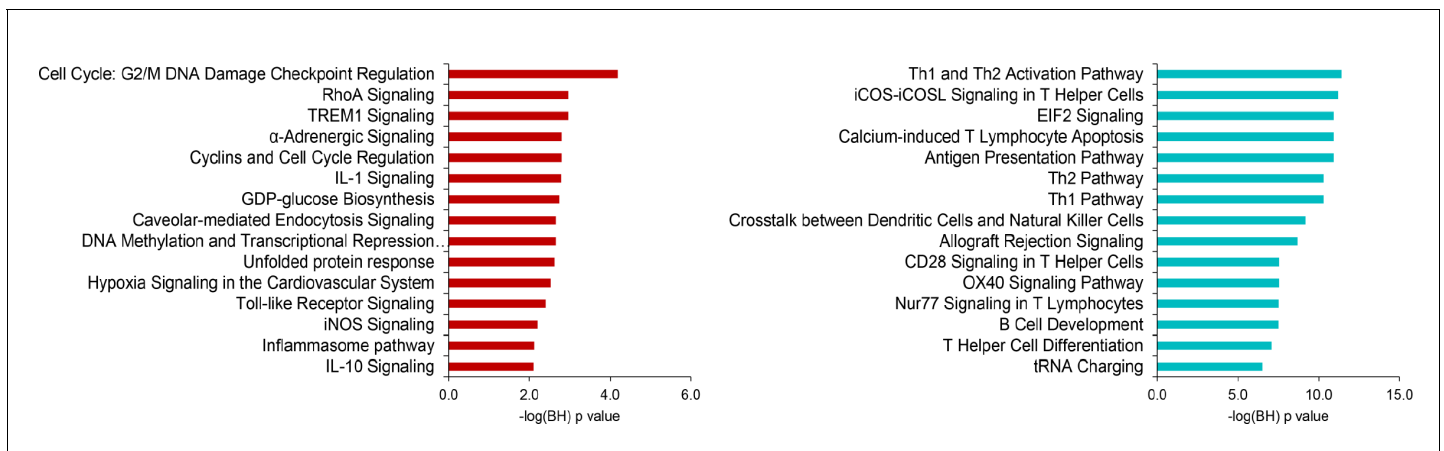
## Figures and figure supplements

The leukocyte non-coding RNA landscape in critically ill patients with sepsis

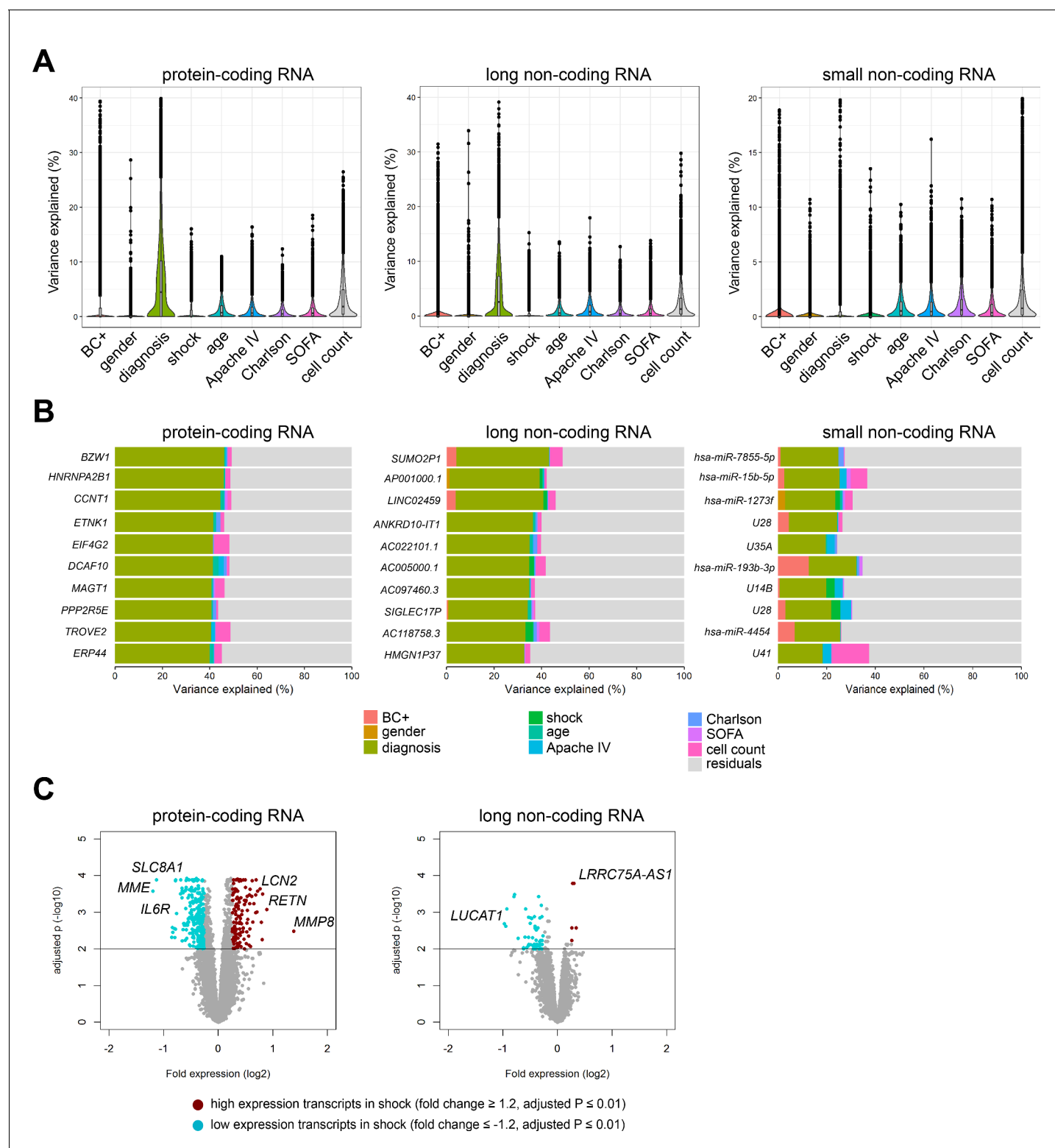
**Brendon P Scicluna *et al***



**Figure 1.** Coding and non-coding RNA expression in leukocytes of sepsis patients and healthy individuals. **(A)** Principal component (PC) plot depicting PC1 and PC2, and **(B)** the molecular distance to health (MDTH) index of protein-coding ( $n = 18,063$ ), long non-coding ( $n = 16,087$ ), and small non-coding RNAs ( $n = 4949$ ) in healthy subjects and sepsis patients.  $**p < 0.01$ ;  $***p < 0.001$ . **(C)** Volcano plot representation of differences in coding and non-coding RNA expression between sepsis patients and healthy subjects. Horizontal (black) line denotes  $-\log_{10}$  transformed adjusted p-value of 0.01. **(D)** Pie chart showing the subclass distribution of significant long non-coding RNA (adjusted  $p < 0.01$ ). LincRNA, long intergenic non-coding RNA; rRNA, ribosomal RNA; TEC, to be experimentally confirmed; Mt tRNA, mitochondrial transfer RNA; Mt rRNA, mitochondrial ribosomal RNA. **(E)** Pie chart showing the subclass distribution of significant small non-coding RNA (adjusted  $p < 0.01$ ). miRNA, microRNA; snoRNA, small nucleolar RNA; C/D box snoRNA, C/D box small nucleolar RNA; H/ACA box snoRNA, H/ACA box small nucleolar RNA; scaRNA, small cajal body-specific RNA.



**Figure 1—figure supplement 1.** Ingenuity pathway analysis of significant protein-coding RNA in sepsis relative to health. Red bars denote pathways harboring protein-coding RNA with elevated expression; turquoise bars denote pathways harboring protein-coding RNA with reduced expression. Significance was demarcated at Benjamini–Hochberg (BH) adjusted  $p < 0.01$ .  $-\log(\text{BH})$  p, negative log-transformed BH p-value.

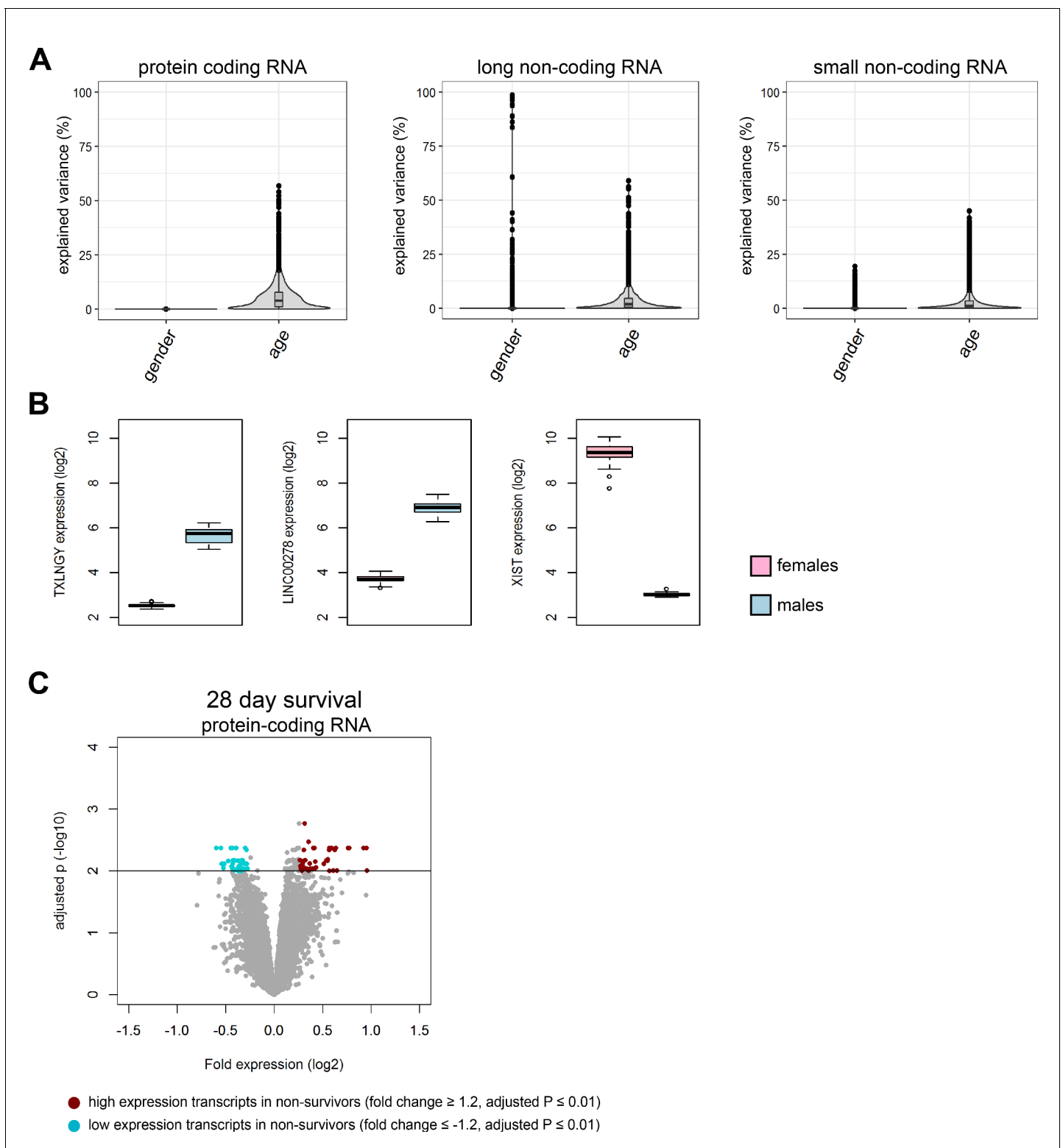


**Figure 2.** Variance in coding and non-coding RNA expression attributed to demographics and clinical characteristics of sepsis patients. **(A)** Violin plots of percent variation in protein-coding, long non-coding, and small non-coding RNA expression explained by sepsis patient demographics and clinical variables. Black dots depict outlier RNA transcripts. **(B)** Percent variance of select protein-coding and long non-coding RNA partitioned into the segment attributable to each demographic and clinical variable ranked by percent variation ( $>20\%$ ) for primary diagnosis (site of infection and place of acquisition). **(C)** Volcano plots depicting the changes in protein-coding and long non-coding RNA in patients discordant for septic shock on ICU

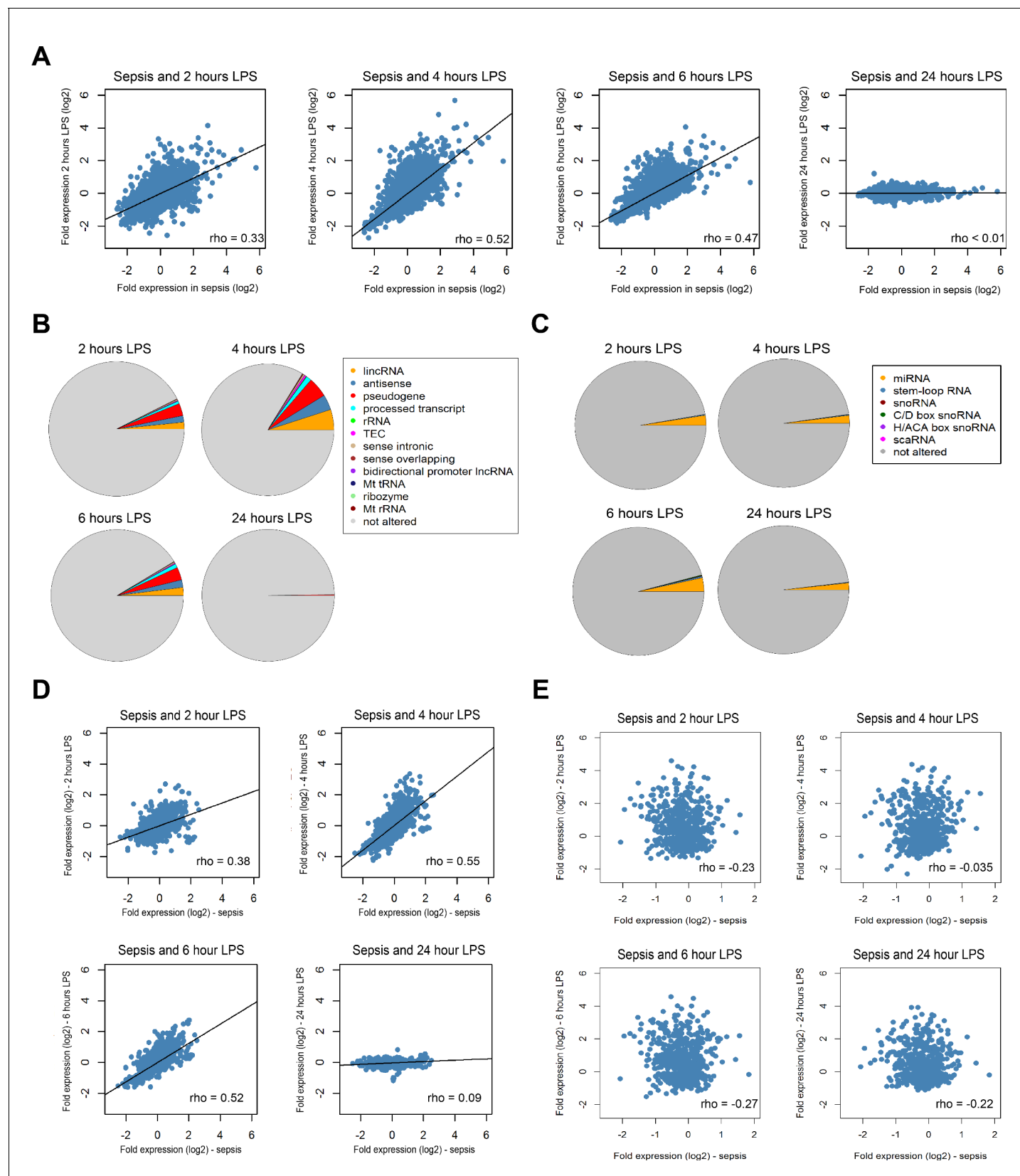
Figure 2 continued on next page

*Figure 2 continued*

admission. Horizontal (black) line denotes the adjusted p-value threshold for significance (adjusted  $p \leq 0.01$ ). Abbreviations: BC+, blood culture positive microbiology; diagnosis, infection site (lung or abdomen) and source (community or hospital); Charlson, Charlson comorbidity index; Apache IV, Acute Physiology and Chronic Health Evaluation; ICU, Intensive care unit; SOFA, Sequential Organ Failure Assessment.



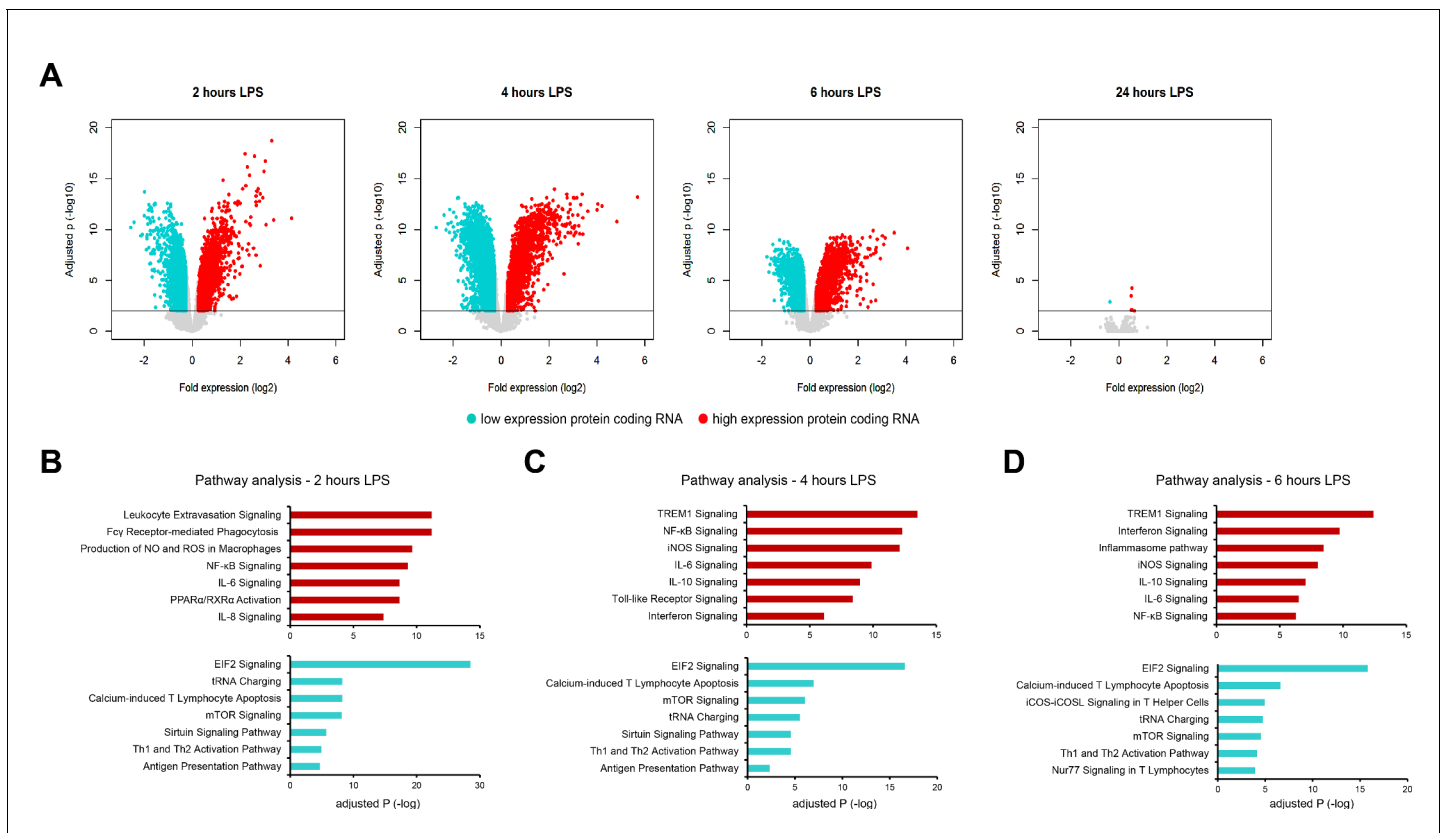
**Figure 2—figure supplement 1.** Variance partition of protein-coding and non-coding RNA expression in health and differential expression in sepsis non-survivors relative to survivors. **(A)** Violin plots of percent variation in protein-coding, long non-coding, and small non-coding RNA expression explained by age in gender in healthy subjects ( $n = 82$ ). Black dots depict outlier RNA. **(B)** Expression of long non-coding RNA *TXLNGY*, *LINC00278*, and *XIST* in healthy males and females. **(C)** Volcano plot of significantly altered protein-coding RNA in non-survivors relative to survivors after 28 days since ICU admission. Horizontal (black) line denotes  $-\log_{10}$  transformed adjusted p-value thresholds.



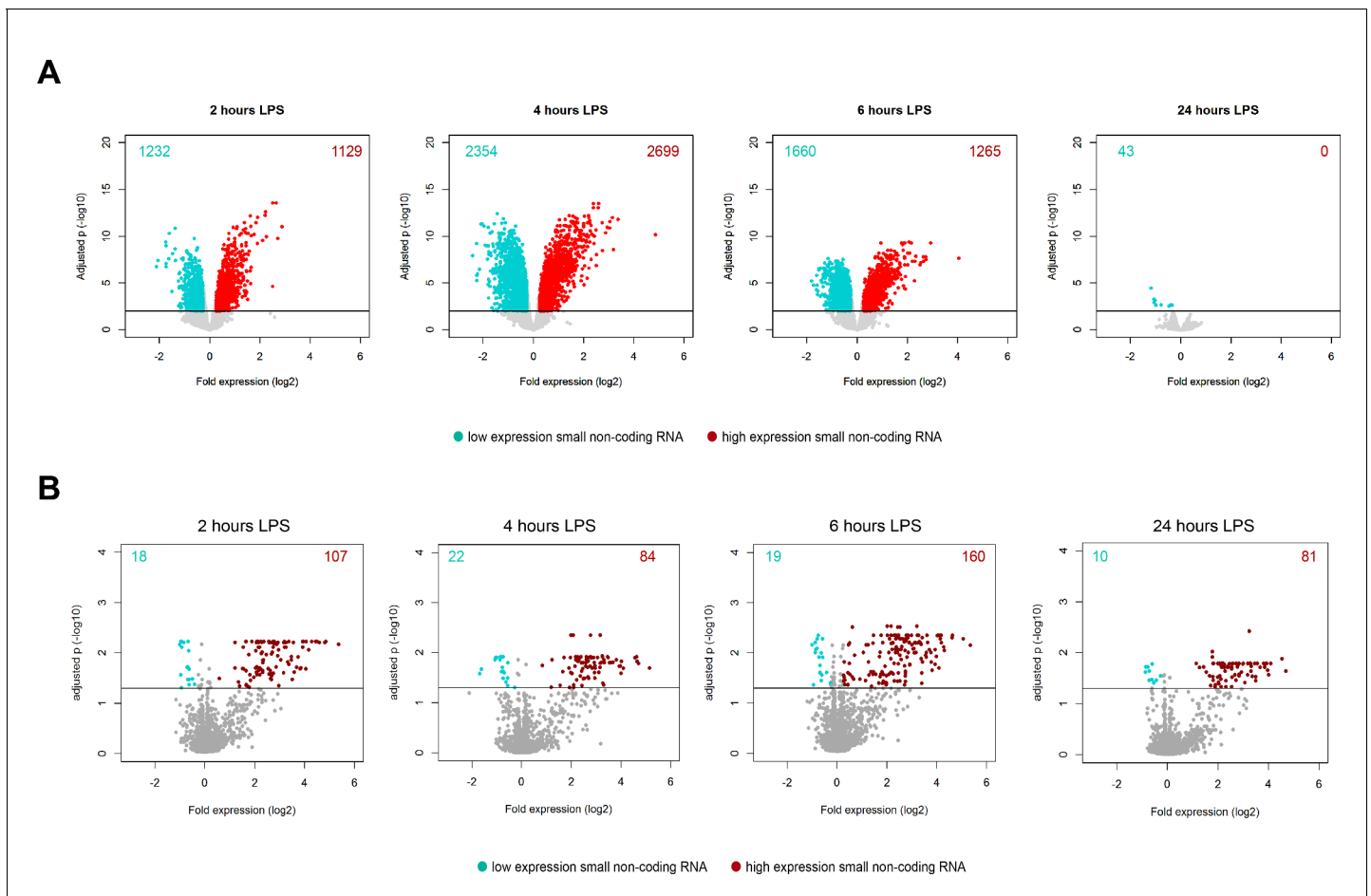
**Figure 3.** Comparison of the coding and non-coding transcriptome in sepsis to human endotoxemia. **(A)** Dot plots depicting the correlation between protein-coding RNA fold expression indices in sepsis (compared to health) and fold expression after 2, 4, 6, and 24 hr lipopolysaccharide (LPS) infusion. *Figure 3 continued on next page*

*Figure 3 continued*

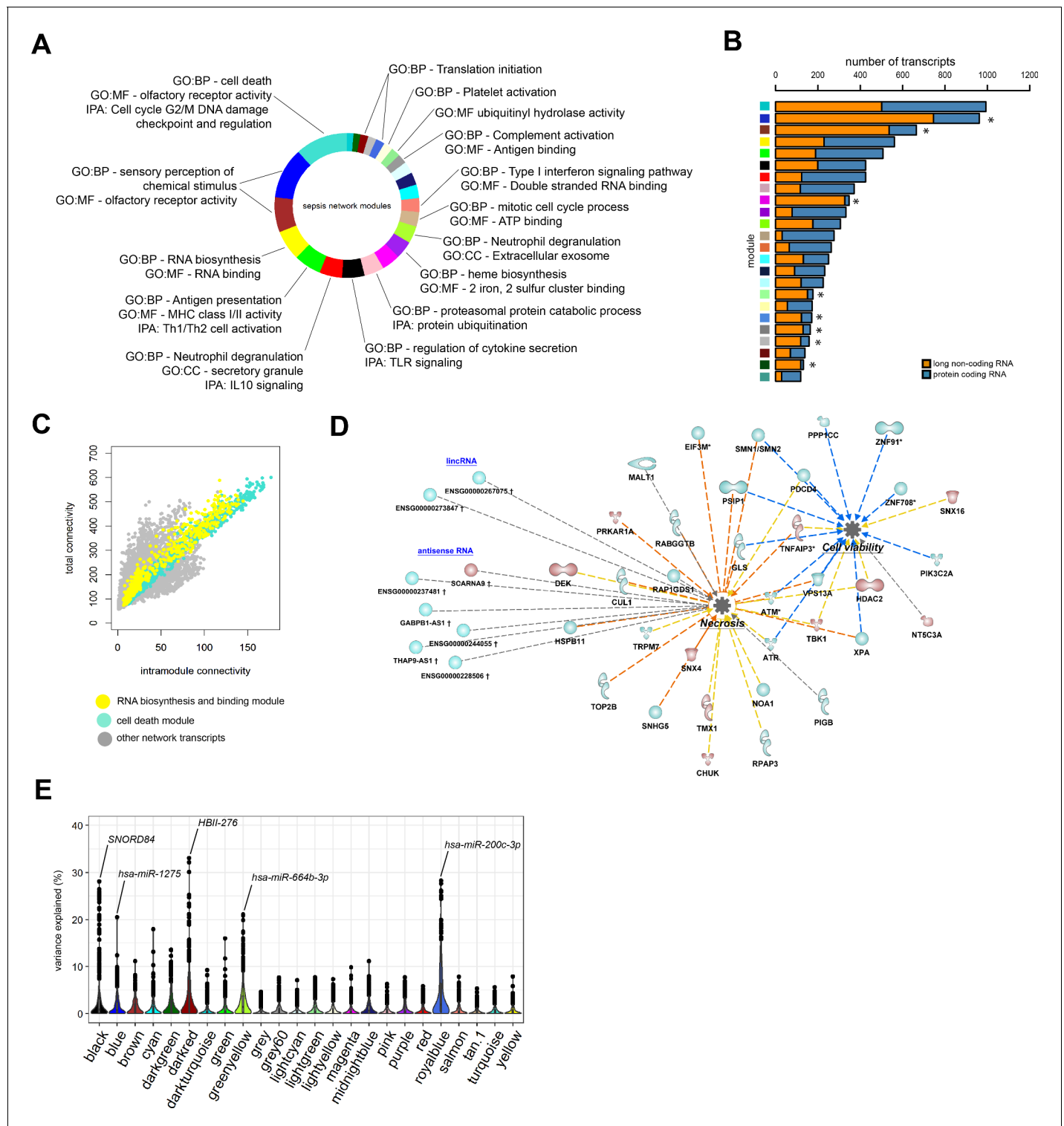
relative to pre-LPS. **(B)** Pie chart illustrating the biotypes of significantly altered long non-coding RNA (adjusted  $p < 0.01$ ) across endotoxemia time points (2, 4, 6, and 24 hr after 2 ng/kg LPS). LincRNA, long intergenic non-coding RNA; rRNA, ribosomal RNA; TEC, to be experimentally confirmed; Mt tRNA, mitochondrial transfer RNA; Mt rRNA, mitochondrial ribosomal RNA. **(C)** Pie chart showing the biotypes of significantly altered small non-coding RNA (adjusted  $p < 0.05$ ) in human endotoxemia. miRNA, microRNA; snoRNA, small nucleolar RNA; C/D box snoRNA, C/D box small nucleolar RNA; H/ACA box snoRNA, H/ACA box small nucleolar RNA; scaRNA, small cajal body-specific RNA. **(D)** Dot plots illustrating the correlation between long non-coding RNA fold expression indices in sepsis (compared to health) and fold expression of 2, 4, 6, and 24 hr after LPS relative to pre-LPS. rho, Spearman's coefficient. **(E)** Dot plots depicting the correlation between small non-coding RNA fold expression indices in sepsis (compared to health) and 2, 4, 6, and 24 hr after LPS relative to pre-LPS. rho, Spearman's coefficient.



**Figure 3—figure supplement 1.** Comparing fold expression in sepsis (relative to health) to human endotoxemia. (A) Volcano plot representation of significantly altered protein-coding RNA after 2, 4, 6, and 24 hr lipopolysaccharide (LPS) infusion relative to pre-LPS. Horizontal (black) line denotes  $-\log_{10}$  transformed adjusted p-value threshold of 0.01. (B–D) Ingenuity pathway analysis of significant protein-coding RNA after 2, 4, and 6 hr human endotoxemia. Red bars denote pathways harboring protein-coding RNA with elevated expression; turquoise bars denote pathways harboring protein-coding RNA with reduced expression. Significance was demarcated at Benjamini–Hochberg (BH) adjusted  $p < 0.01$ . Adjusted P, negative log-transformed BH p-value.



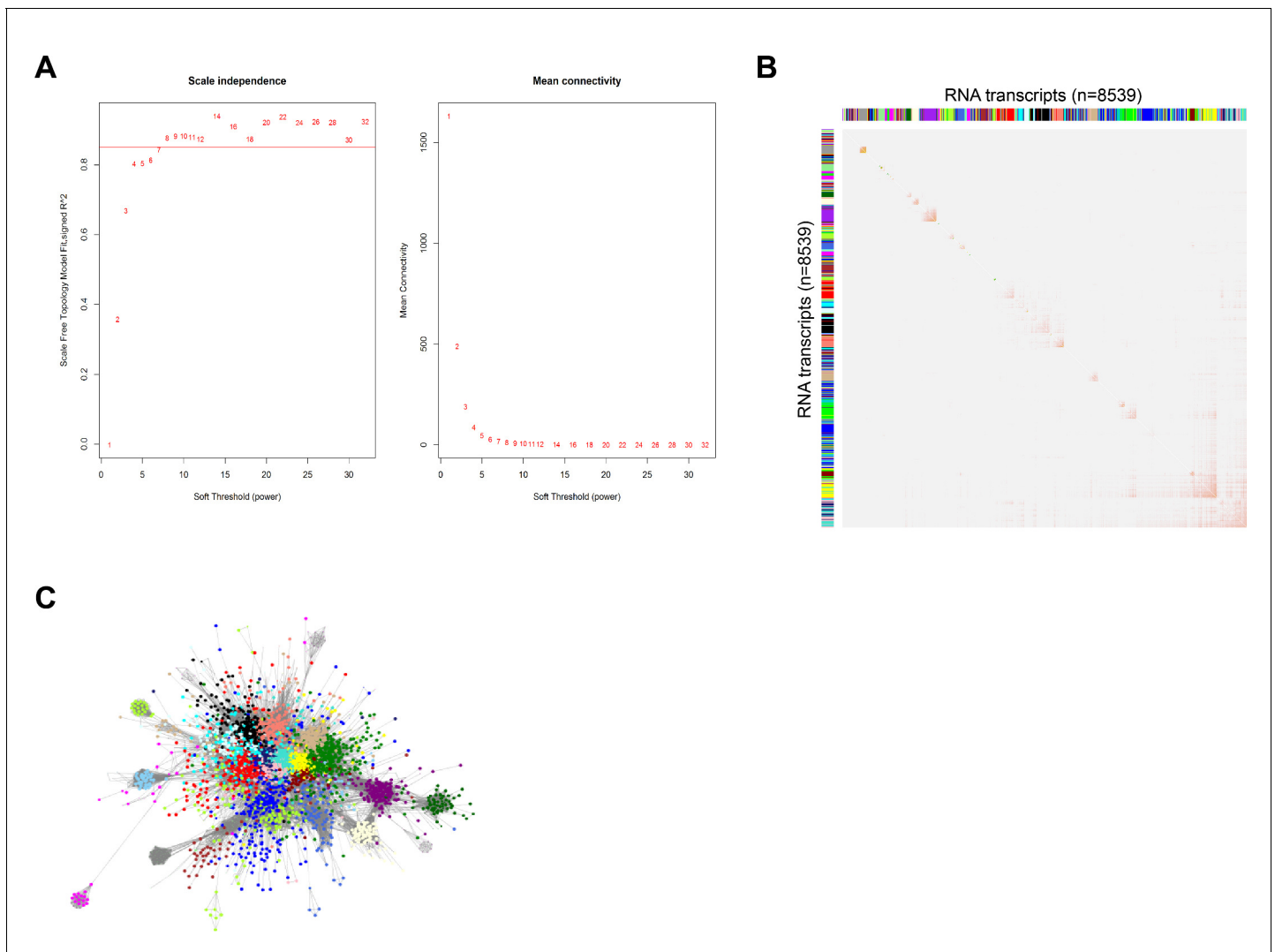
**Figure 3—figure supplement 2.** Volcano plot representations of significantly altered. (A) Long non-coding RNA and (B) small non-coding RNA after 2, 4, 6, and 24 hr lipopolysaccharide (LPS) relative to pre-LPS. Horizontal (black) line denotes  $-\log_{10}$  transformed adjusted p-value thresholds.



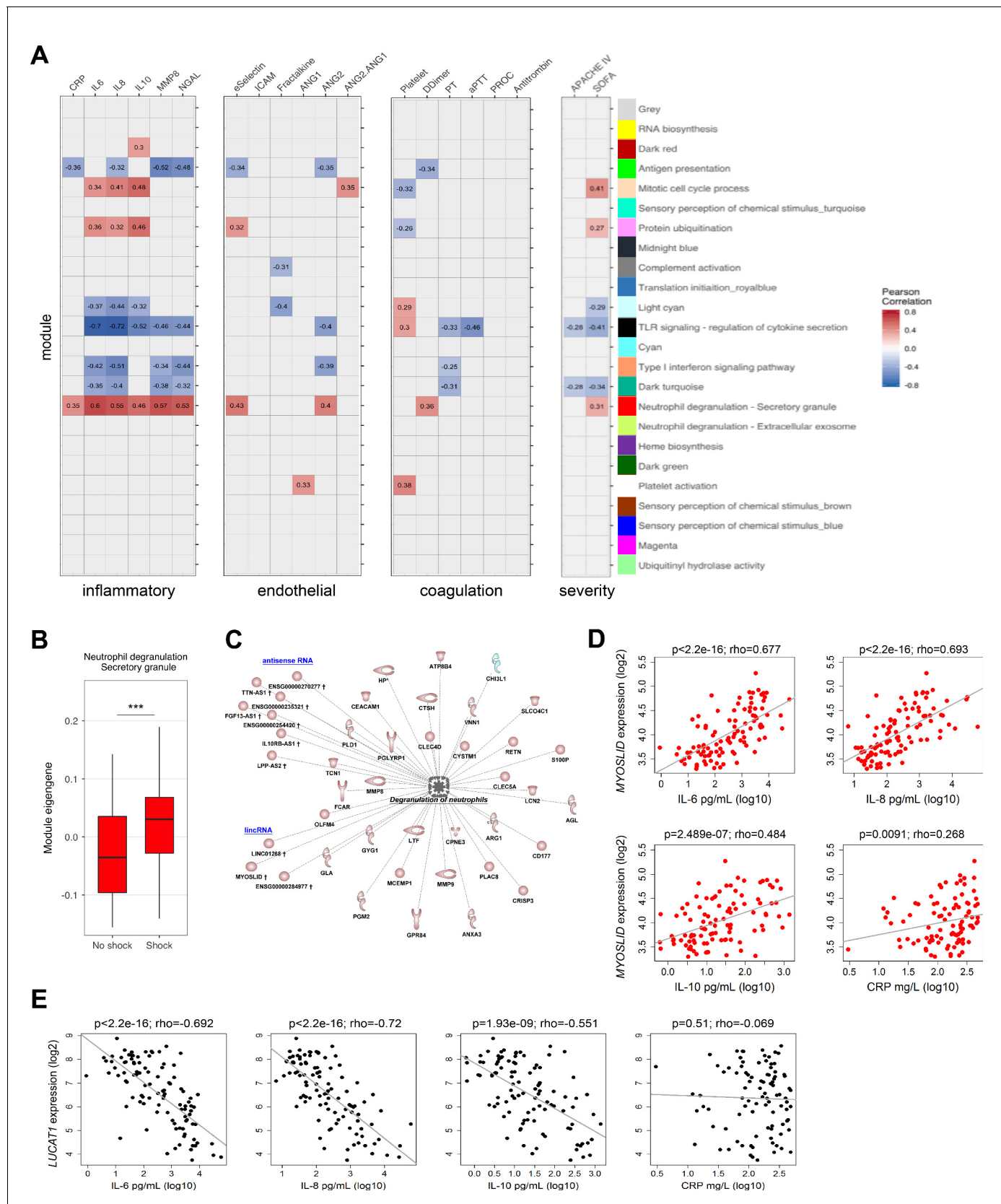
**Figure 4.** Network analysis of coding and non-coding RNA expression. (A) Circular plot of protein-coding and long non-coding co-expression network modules characterized by significantly associated (Fisher's adjusted  $p < 0.01$ ) gene ontologies and Ingenuity canonical signaling pathways. Seventeen modules were associated with specific ontologies or canonical signaling pathways. (B) Bar plot depicting the distribution of protein coding and long non-coding RNA in each network module. \*Fisher's Benjamini-Hochberg adjusted  $p < 0.01$ . (C) Dot plot illustrating the correlation between intramodular and total connectivities of each RNA transcript in their respective network module. Yellow dots illustrate protein-coding and long non-coding RNA in the RNA biosynthesis/RNA binding module; Turquoise dots depict the cell death and olfactory receptor activity module. (D) Diagrammatic representation of Ingenuity's biofunctions (z-score  $< 2$  or  $> 2$  and adjusted  $p < 0.05$ ) together with predicted long intergenic non-coding RNA (lincRNA) Figure 4 continued on next page

Figure 4 continued

and antisense RNA in the cell death/olfactory receptor activity/cell-cycle G2/M DNA damage checkpoint and regulation module (turquoise). Blue, reduced expression; red, elevated expression in sepsis relative to health (fold change  $\geq 1.2$  or  $\leq -1.2$ ; adjusted p-value  $< 0.01$ ). (E) Violin plots of network module eigengene (first principal component) percent variance attributable to small non-coding RNA.



**Figure 4—figure supplement 1.** Co-expression network analysis. **(A)** Evaluation of scale free topology model fit and mean connectivities (y-axes) across various soft threshold powers (x-axis) with scale independence denoted at  $R^2 > 0.85$  (red horizontal line) for protein-coding and long non-coding RNA expression in sepsis patients ( $n = 8539$ ). **(B)** Topological overlap plot of adjacencies calculated for 8539 protein-coding and long non-coding RNA expression and module colors. **(C)** Cytoscape plot (organic layout) of protein-coding and long non-coding RNA (nodes) and connectivities (edges; weight  $> 0.2$ ). Turquoise and yellow modules were visibly central to the co-expression network.



**Figure 5.** Relationship of protein-coding, non-coding RNA network modules to soluble mediators and clinical severity. (A) Heatmap representation of Pearson correlation coefficients (adjusted  $p < 0.05$ ) calculated for each network module eigengene (first principal component) against soluble mediators. Figure 5 continued on next page

*Figure 5 continued*

of inflammation, endothelial function, coagulation, as well as clinical parameters of disease severity. APACHE IV, Acute Physiology and Chronic Health Evaluation; SOFA, Sequential Organ Failure Assessment. Red denotes direct correlations and blue denotes indirect correlations **(B)** Boxplot showing differences in neutrophil degranulation (red) module eigengene values in sepsis patients discordant for septic shock on intensive care unit admission. High module eigengene values mean overall elevated RNA expression; low module eigengene values mean reduced expression. **(C)** Diagrammatic representation of the neutrophil degranulation (secretory; red) module (Ingenuity's biofunction z-score <2 or >2; adjusted  $p < 0.05$ ) together with predicted long intergenic non-coding RNA (lincRNA) and antisense RNA. Red or blue nodes denote high or low expression RNA transcripts in sepsis relative to health, respectively. \*\*\*Mann–Whitney  $p < 0.001$ . **(D and E)** Dot plots of **(D)** *MYOSLID* expression and **(E)** *LUCAT1* expression against soluble mediators of inflammation IL-6, IL-8, and IL-10, as well as the acute phase response protein CRP. Rho, Spearman's coefficient.