A

B

Figure S1 Patients with LN metastases were subjected to poor OS. (A) Tumors with LN metastases were associated with shorter OS. (B) Multivariate cox regression analysis on OS. The multivariate Cox-proportional hazards model included Lymph metastasis, sex, age, and smoke. LN, lymph node; OS, overall survival; HR, hazard ratio; CI, confidence interval.
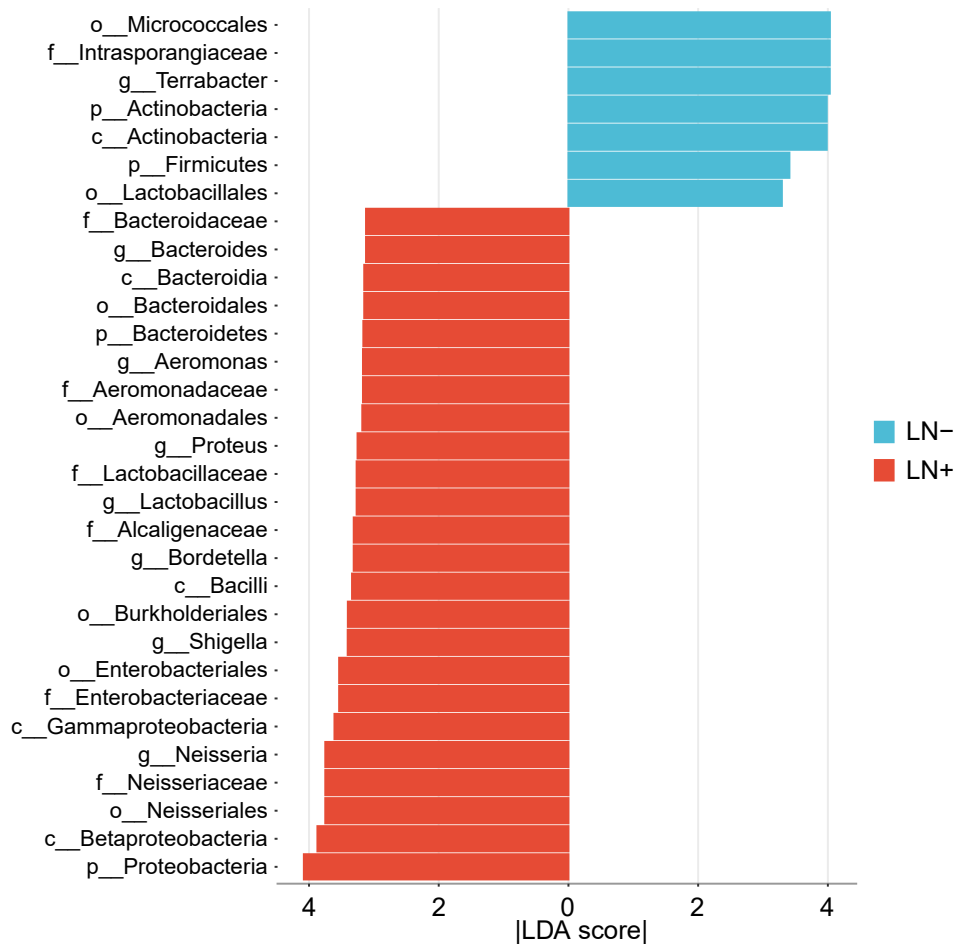

Figure S2 Histogram of the LDA combined with effective size (LEfSe). The LDA score indicates the effect size and ranking of each differentially abundant taxon (|LDA score| >3). The "p" indicate phylum, the "c" indicate class, the "o" indicate order, the "f" indicate family, and the "g" indicate genus. LDA, linear discriminant analysis.
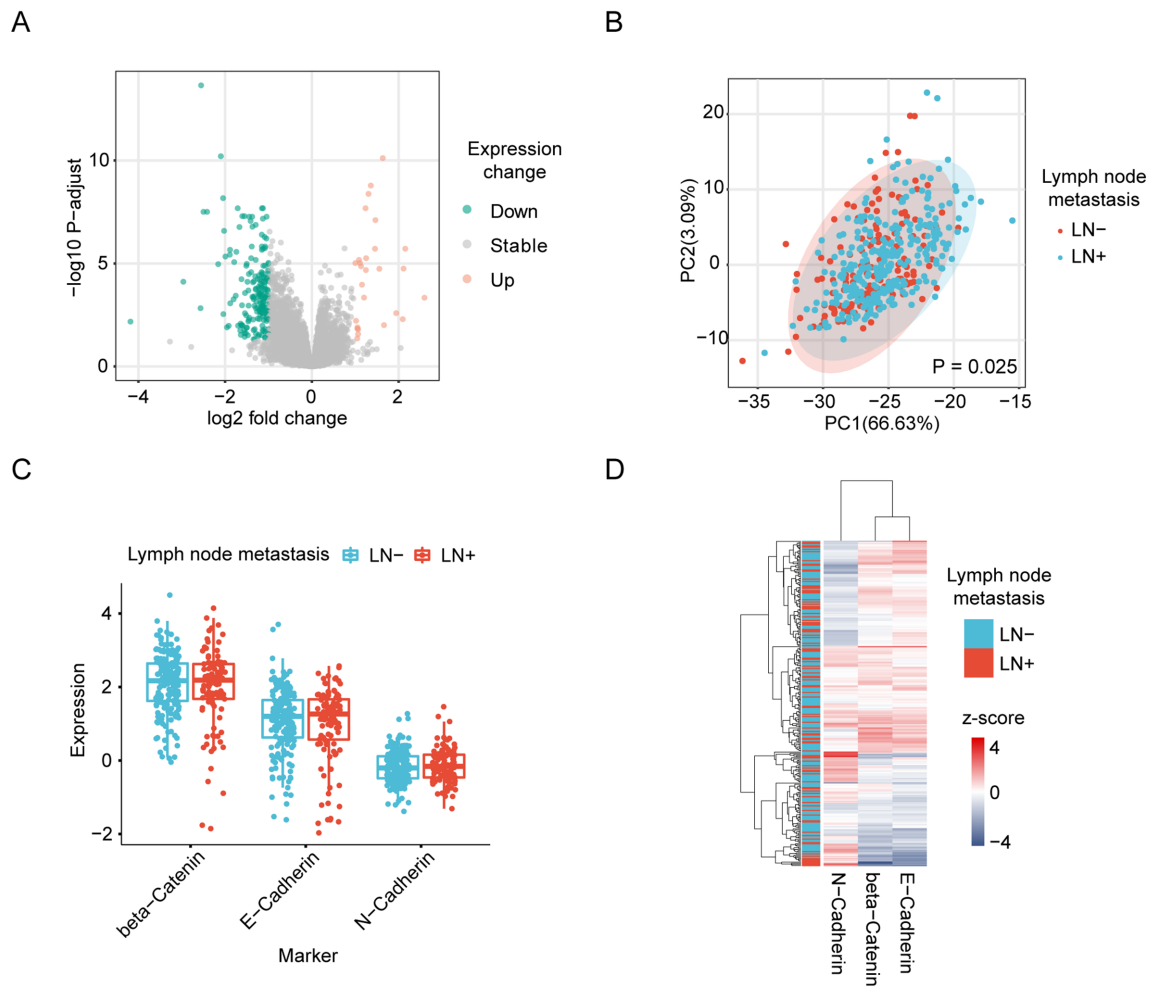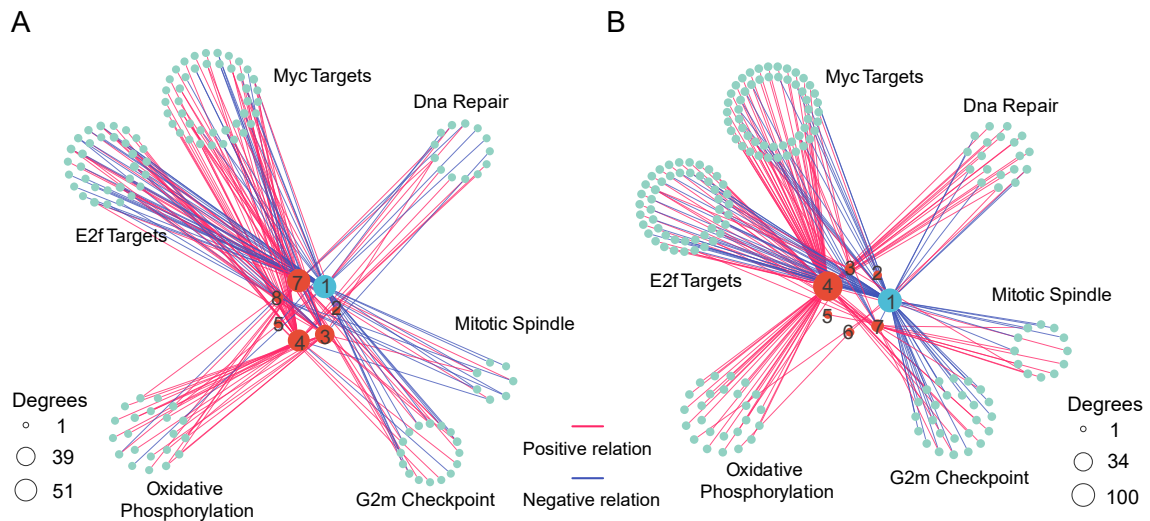
Figure S3 Different profiles in transcripts and epithelial-mesenchymal transition protein markers between LN– and LN+ samples. (A) Differential expression genes between LN– and LN+. (B) PCoA based on hallmark gene sets shows a different cluster between LN+ and LN–. (C) Boxplot displays three EMT protein markers expression. (D) Heatmap clustering of the EMT markers expression. PC, principal component; LN, lymph node; PCoA, principal coordinate analysis; EMT, epithelial mesenchymal transition.
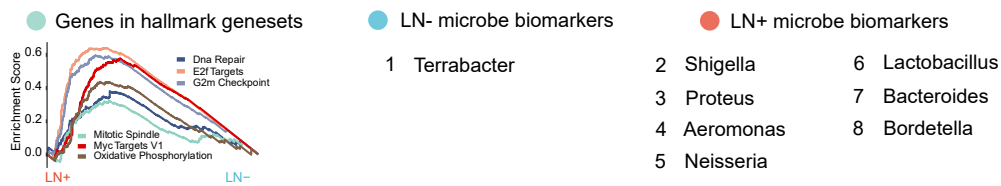
Figure S4 Correlation networks between microbiota and enriched hallmark gene sets. Each line represents a pair of microbe-gene correlation, and the red and blue lines indicate the positive and negative correlations, respectively. The node size is proportional to the correlation degrees, and the green, red and blue nodes indicates the genes in the enriched hallmark gene sets, LN+ microbe biomarkers, and LN– microbe biomarkers, respectively. The enriched hallmark gene sets in GSEA were included in the microbiota-gene sets network. (A) Association between the genes in the enriched hallmark gene sets and microbiota biomarkers in LN+. (B) Association between the genes in the enriched hallmark gene sets and microbiota biomarkers in LN–. LN, lymph node; GSEA, gene set enrichment analysis.