

Figure Legends

Supplementary Figure S1 OTU rarefaction curve of 85 NSCLC patients.

Supplementary Figure S2 Taxonomic composition of sputum microbiota community of the 85 NSCLC subjects. **(A)** Dominant phyla of the 85 NSCLC subjects; **(B)** Dominant genera of of the 85 NSCLC subjects.

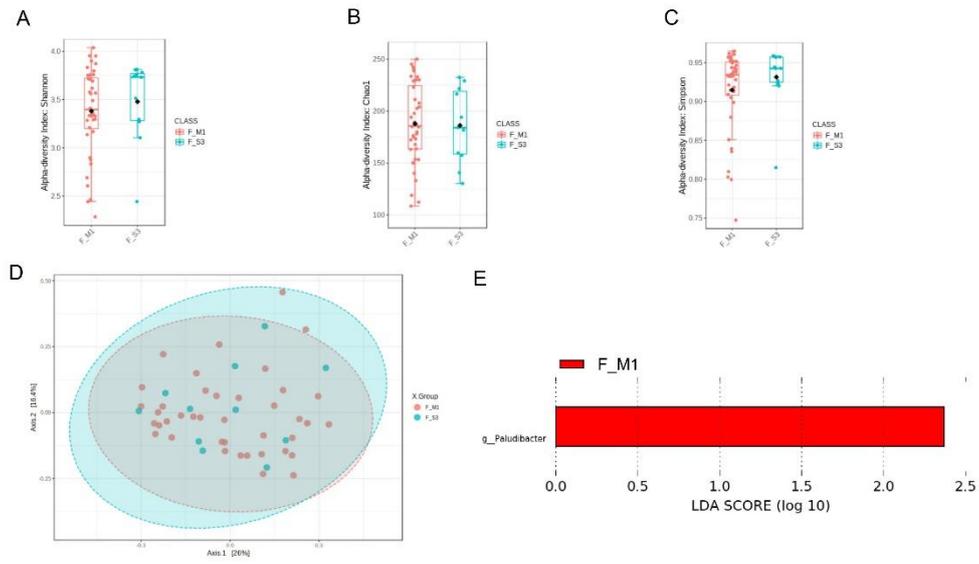
Supplementary Figure S3 Difference of sputum microbiota between NSCLC patients in stage III and stage IV. **(A)** Shannon index; **(B)** Chao1 index; **(C)** Simpson index among NSCLC patients in stage III and stage IV; **(D)** PCOA plot based on Bray-Curtis distance of sputum genus among NSCLC patients in stage III and stage IV; **(E)** Differentially abundant taxonomy between NSCLC patients in stage III and stage IV identified by LEFse.

Supplementary Figure S4 Difference of sputum microbiota between NSCLC patients in Intra group and Non_M group. **(A)** Chao 1 index, **(B)** Simpson index, **(C)** Shannon index among NSCLC patients in Intra group and Non_M group; **(D)** PCOA plot based on Bray-Curtis distance of sputum genus among NSCLC patients in Intra group and Non_M group.

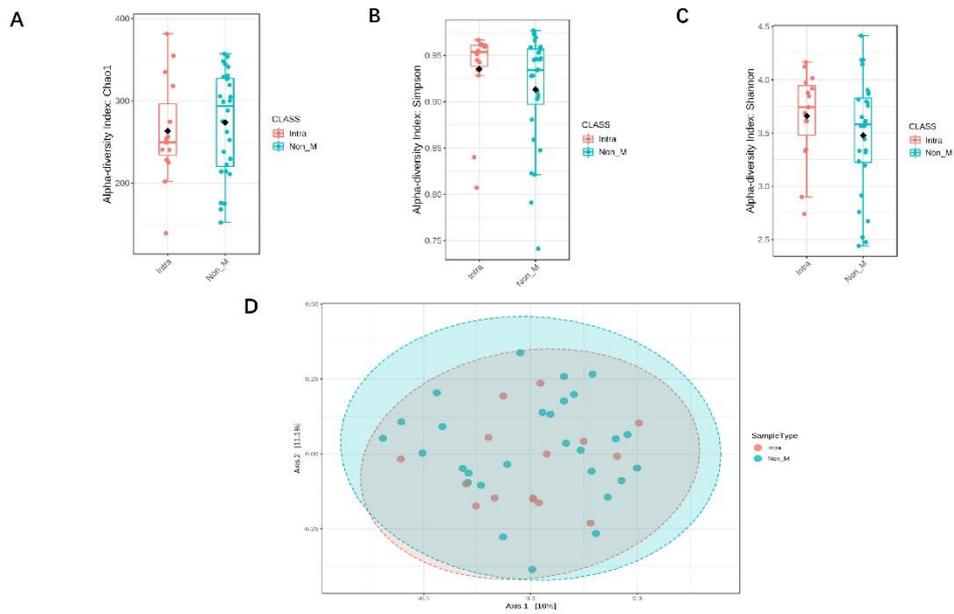
Supplementary Figure S5 Differentially abundant taxonomy and predicted metabolic function of sputum microbiota between NSCLC patients in Intra and Non_M group. **(A)** Differentially abundant taxonomy between patients in Intra and Non_M group identified by LEFse; **(B)** Differentially abundant of phylum *Firmicutes* between Intra and Non_M group; **(C)** Differentially abundant of genera *Peptostreptococcus* and *Streptococcus* between Intra and Non_M; **(D)** Differential predicted metabolic function based on MetaCyc database between patients in Intra and Non_M group. * $P < 0.05$, P was calculated using Mann-Whitney U test.

Supplementary Figure S6 Difference of sputum microbiota between NSCLC patients in LNM_N and LNM_Y group. **(A)** Chao 1 index, **(B)** Simpson index, **(C)** Shannon index among NSCLC patients in LNM_N and LNM_Y group; **(D)** PCOA plot based on Bray-Curtis distance of sputum genus among NSCLC patients in LNM_N and LNM_Y group.

Supplementary Figure S7 Differentially abundant taxonomy and predicted metabolic function of sputum microbiota between NSCLC patients in LNM_Y and LNM_M group. **(A)** Differentially abundant taxonomy between patients in LNM_Y and LNM_M group identified by LEFse; **(B)** Differentially abundant of phylum *Proteobacteria* between LNM_Y and LNM_M group; **(C)** Differentially abundant of genera *Actinobacillus*, *Neisseria*, *Parvimonas* and *Pseudomonas* between LNM_Y and LNM_M; **(D)** Differential predicted metabolic function based on MetaCyc database between patients in LNM_Y and LNM_M group. * $P < 0.05$, P was

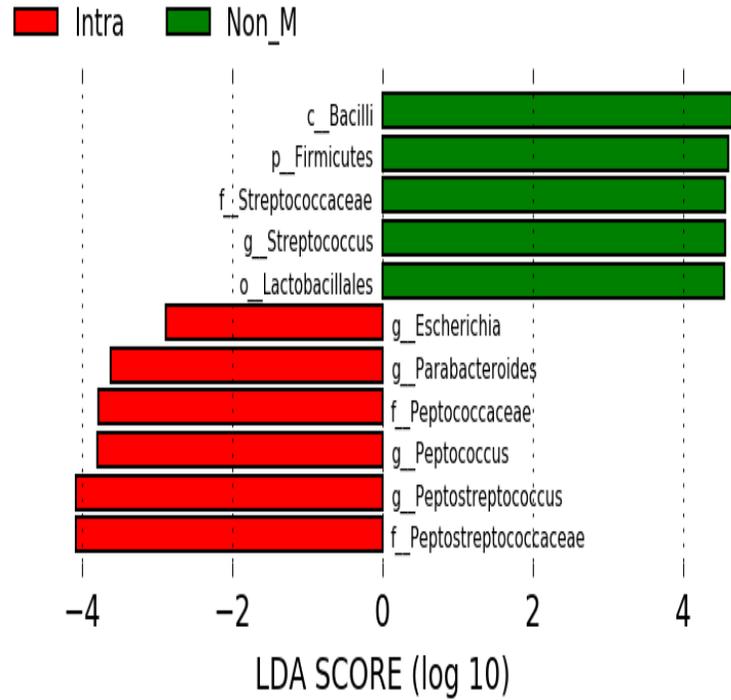


Supplementary Figure S3

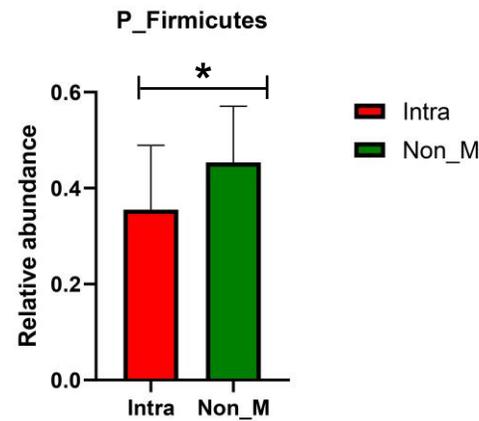


Supplementary Figure S4

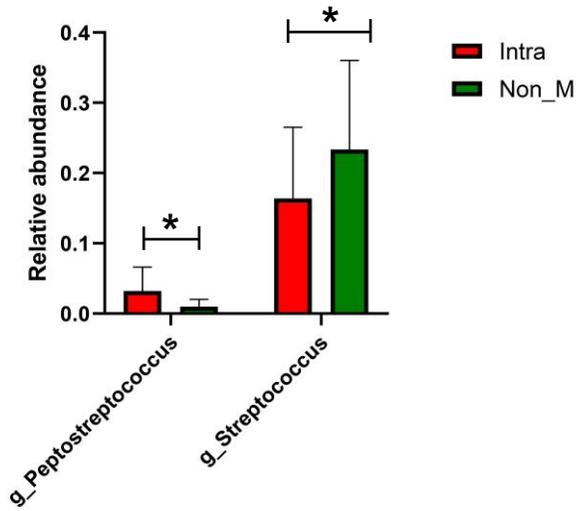
A



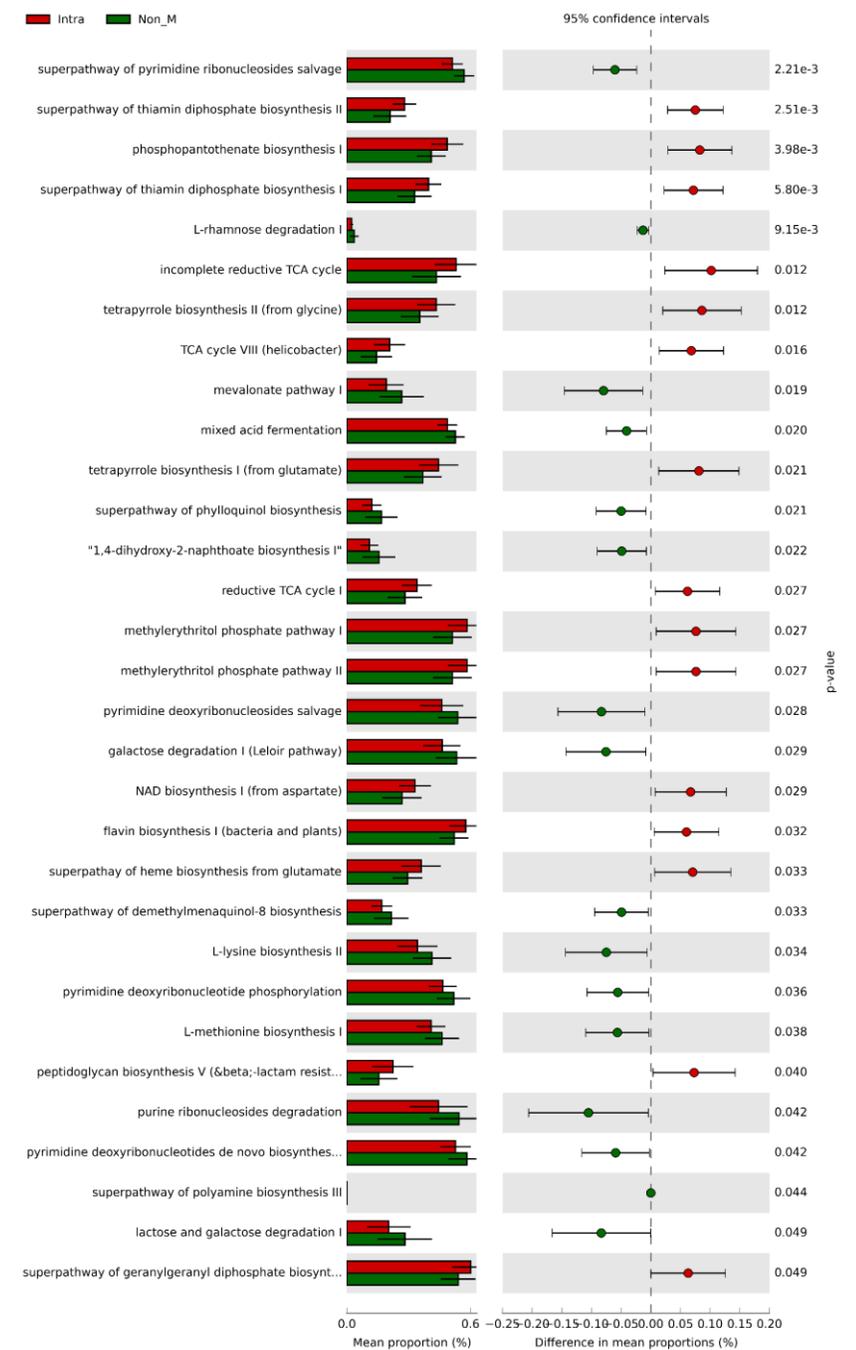
B



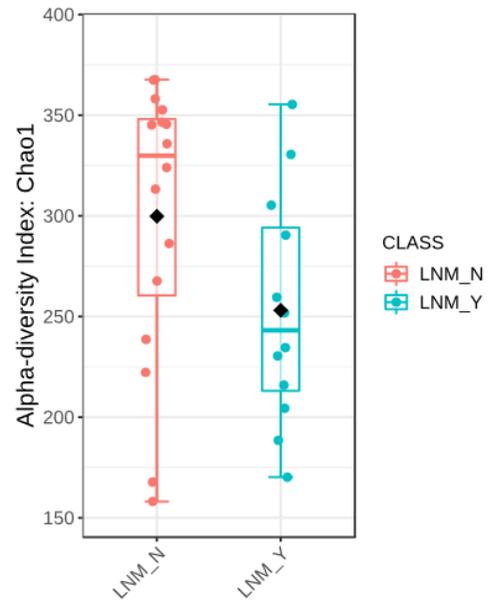
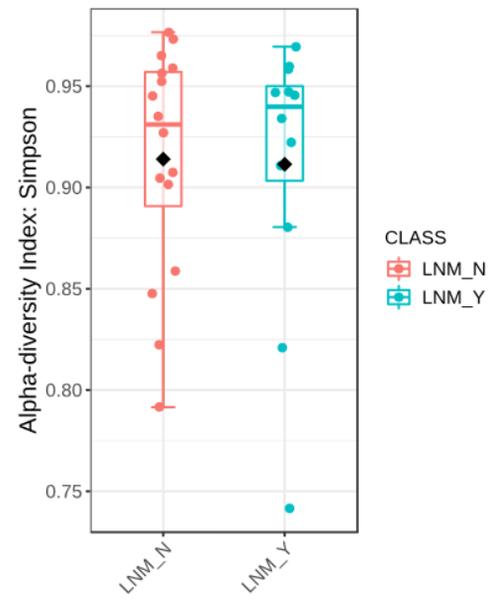
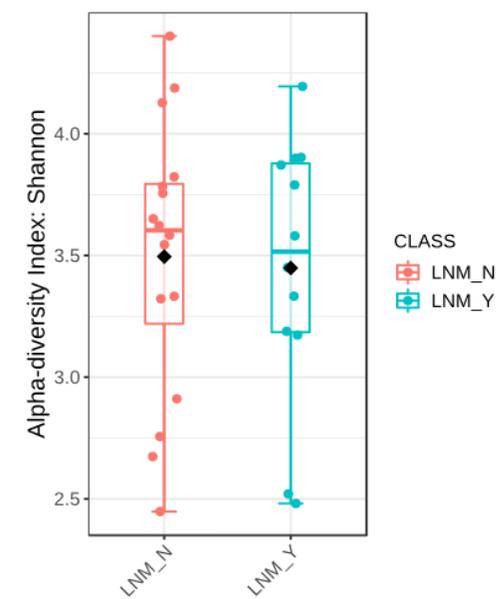
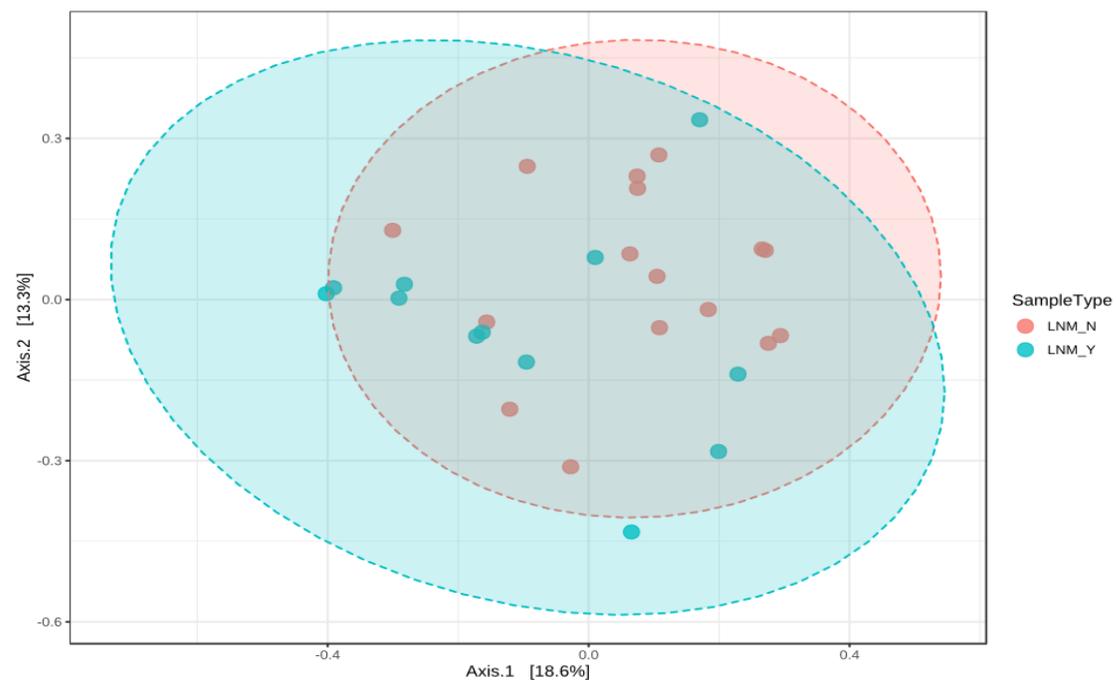
C



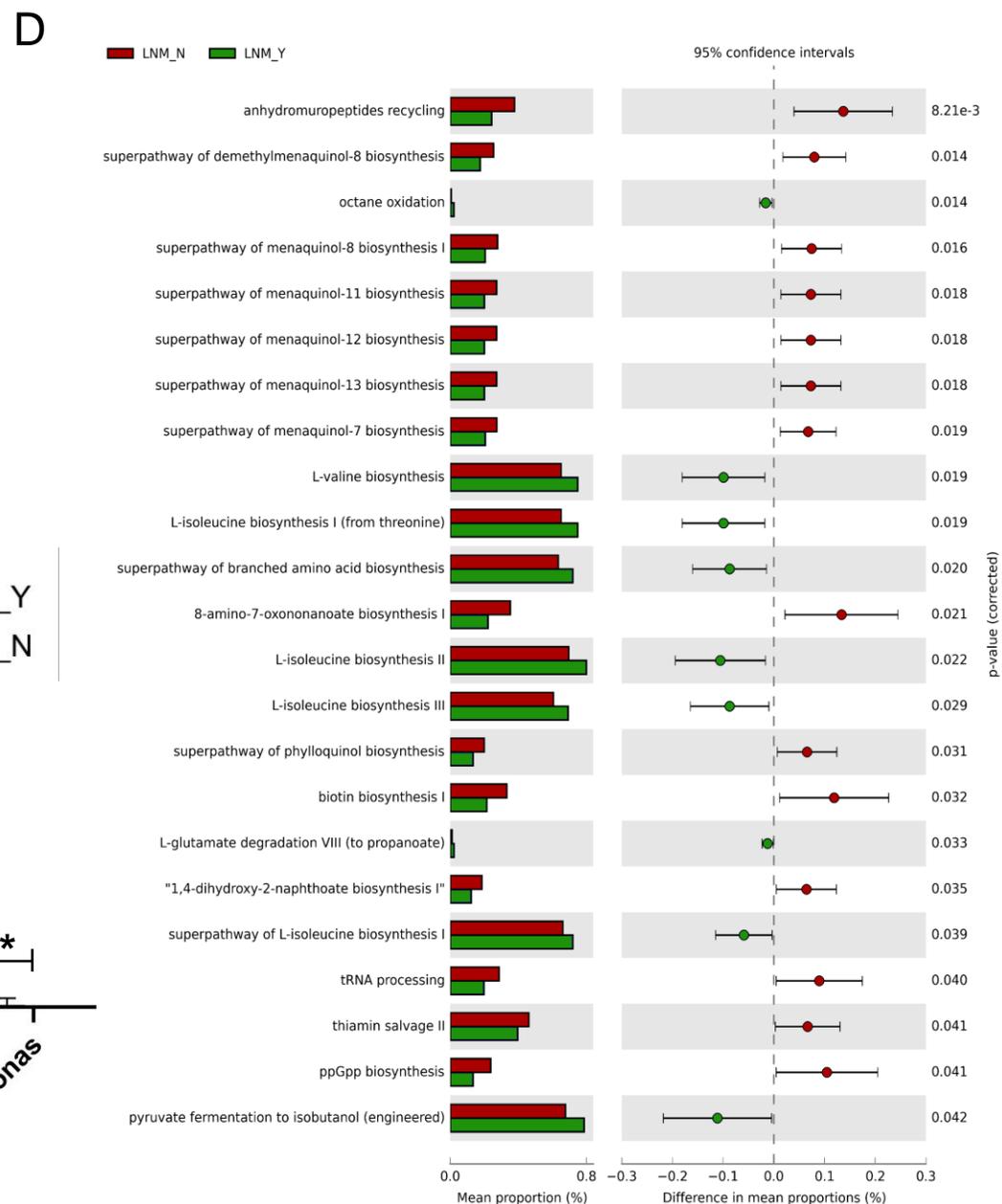
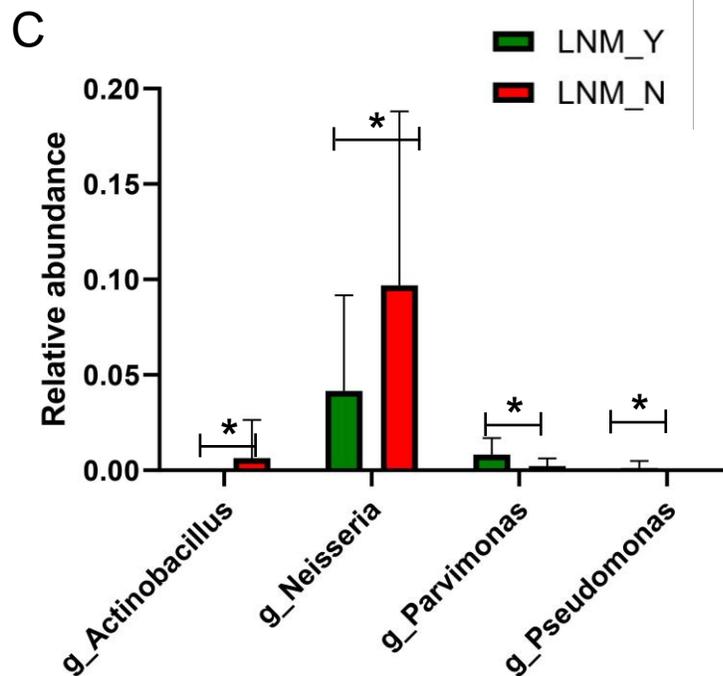
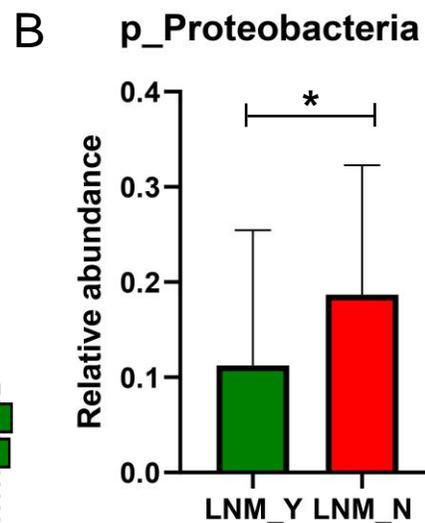
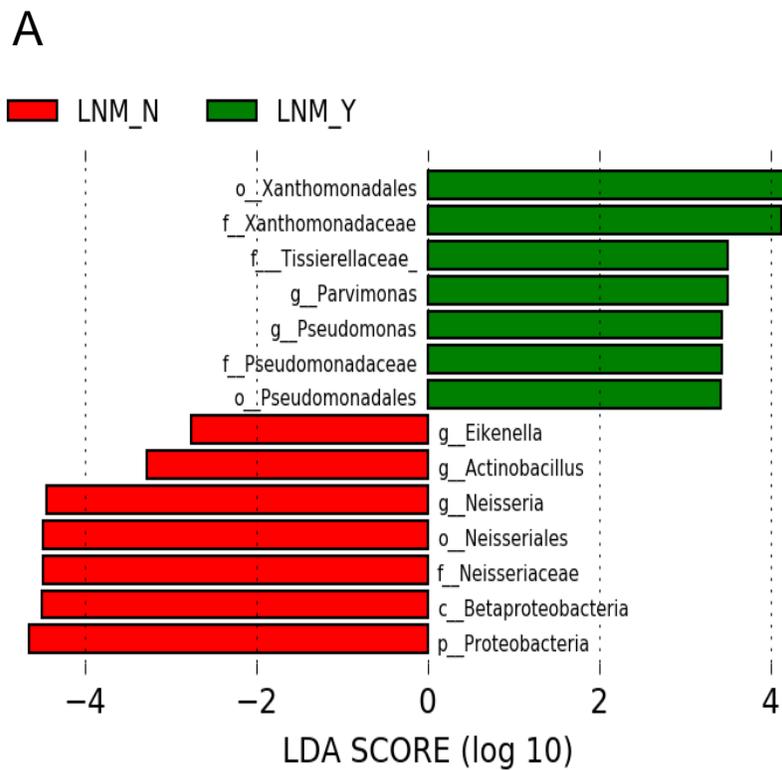
D



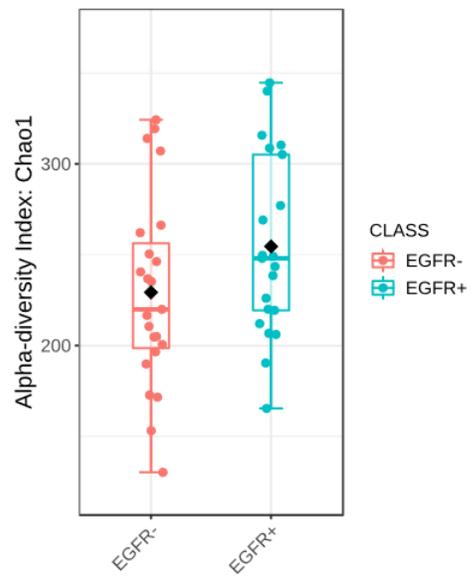
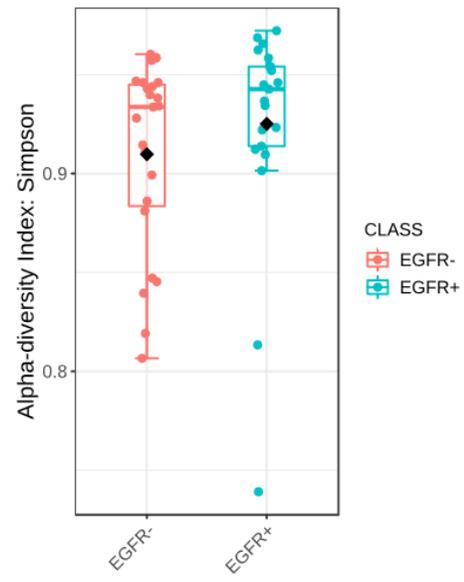
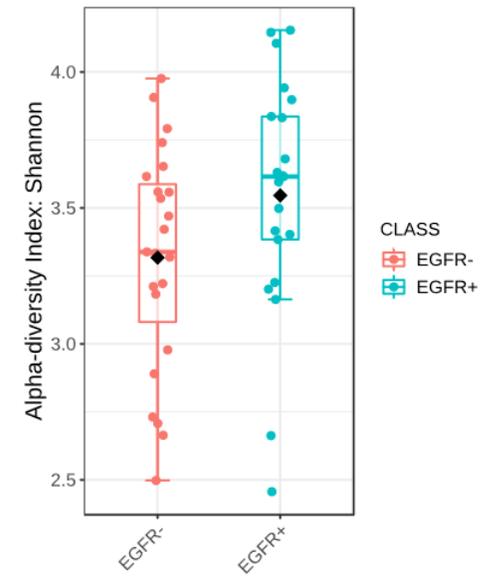
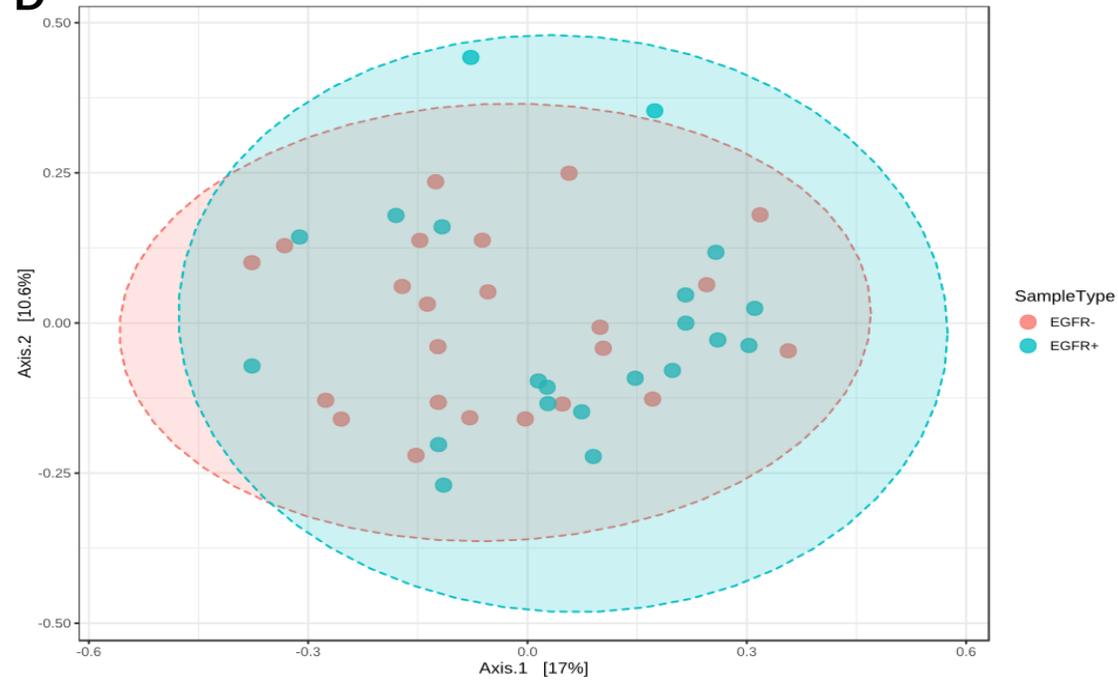
Supplementary Figure S5

A**B****C****D**

Supplementary Figure S6



Supplementary Figure S7

A**B****C****D**

Supplementary Figure S8