Supporting Information

Fig S1. Bar plots of bacterial composition in the intestinal samples. Samples were grouped by HIV-negative (n=44) and positive (n=18). (A) at the phylum level. (B) at the genus level.

Figure S2. Predictive functional analysis of colon wash samples. (A) BugBase predicted the relative abundance of biofilm-forming bacteria. Samples grouped by HIV-negative (n=12) and positive (n=5), p-value = 0.16. (B) BugBase predicted relative abundance of potentially pathogenic bacteria. Samples grouped by HIV-negative (n=12) and positive (n=5), p-value = 0.06. (C) BugBase predicted relative abundance of Gram-positive bacteria. Samples grouped by HIV-negative (n=12) and positive (n=5), p-value = 0.02. (D) BugBase predicted relative abundance of Gram-negative bacteria. Samples grouped by HIV-negative (n=12) and positive (n=5), p-value = 0.02. (E) The KEGG pathway of gut microbiota was predicted using PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States). Data are presented in a bar plot with 95% confidence intervals and p-values between gut samples from HIV-positive and negative patients.

Figure S3. Predictive functional analysis of colon brush samples. (A) BugBase predicted the relative abundance of biofilm-forming bacteria. Samples grouped by HIV-negative (n=11) and positive (n=5), p-value = 0.11. (B) BugBase predicted relative abundance of potentially pathogenic bacteria. Samples grouped by HIV-negative (n=11) and positive (n=5), p-value = 0.008. (C) BugBase predicted relative abundance of Gram-positive bacteria. Samples grouped by HIV-negative (n=11) and positive (n=5), p-value = 0.44. (D) BugBase predicted relative abundance of Gram-negative bacteria. Samples grouped by HIV-negative (n=11) and positive (n=5), p-value = 0.44. (E) The KEGG pathway of gut microbiota was predicted using PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States). Data are presented in a bar plot with 95% confidence intervals and p-values between gut samples from HIV-positive and negative patients.

Figure S4. Predictive functional analysis of TI wash samples. (A) BugBase predicted the relative abundance of biofilm-forming bacteria. Samples grouped by HIV-negative

1

31   (n=10) and positive (n=4), p-value = 0.14. (B) BugBase predicted relative abundance of

32   potentially pathogenic bacteria. Samples grouped by HIV-negative (n=10) and positive

33   (n=4), p-value = 0.04. (C) BugBase predicted relative abundance of Gram-positive

34   bacteria. Samples grouped by HIV-negative (n=10) and positive (n=4), p-value = 0.04.

35   (D) BugBase predicted relative abundance of Gram-negative bacteria. Samples grouped

36   by HIV-negative (n=10) and positive (n=4), p-value = 0.04. (E) The KEGG pathway of

37   gut microbiota was predicted using PICRUSt (Phylogenetic Investigation of

38   Communities by Reconstruction of Unobserved States). Data are presented in a bar plot

39   with 95% confidence intervals and p-values between gut samples from HIV-positive

40   and negative patients.

41   **Figure S5. Predictive functional analysis of TI brush samples.** (A) BugBase

42   predicted the relative abundance of biofilm-forming bacteria. Samples grouped by HIV-

43   negative (n=11) and positive (n=4), p-value = 0.75. (B) BugBase predicted relative

44   abundance of potentially pathogenic bacteria. Samples grouped by HIV-negative (n=11)

45   and positive (n=4), p-value = 0.22. (C) BugBase predicted relative abundance of Gram-

46   positive bacteria. Samples grouped by HIV-negative (n=11) and positive (n=4), p-value

47   = 0.41. (D) BugBase predicted relative abundance of Gram-negative bacteria. Samples

48   grouped by HIV-negative (n=11) and positive (n=4), p-value = 0.41. (E) The KEGG

49   pathway of gut microbiota was predicted using PICRUSt (Phylogenetic Investigation of

50   Communities by Reconstruction of Unobserved States). Data are presented in a bar plot

51   with 95% confidence intervals and p-values between gut samples from HIV-positive

52   and negative patients.

53

54   **Figure S6. Bar plots of bacterial composition in the saliva samples.** Samples were

55   grouped by HIV-negative (n=12) and positive (n=5). (A) at the phylum level. (B) at the

56   genus level.

57

58   **Figure S7. Diversity analysis of the HIV-negative samples by sampling site.** (A)

59   Principal coordinates analysis (PCoA) plot of weighted UniFrac distances (metrics of β-

60   diversity). Samples grouped by TI wash (n=10), TI brush (n=11), colon wash (n=12),

61   colon brush (n=11), and saliva (n=12). False discovery rate corrected q-value < 0.001

62   between saliva and every other group. q-value = 0.001 between TI brush and other

63   intestinal samples (TI wash, colon wash, colon brush). q-value > 0.2 between colon,

64   colon brush, and TI wash. (B) Faith's Phylogenic Diversity (metrics of α-diversity) at

65   sequencing depth 80,000. Samples grouped by TI wash (n=10), TI brush (n=11), colon

66   wash (n=12), colon brush (n=11), and saliva (n=12).

67

68   **Figure S8. Diversity analysis of the HIV-positive samples sampling site.** (A).

69   Principal coordinates analysis (PCoA) plot of weighted UniFrac distances (metrics of β-

70   diversity). Samples grouped by TI wash (n=4), TI brush (n=4), colon wash (n=5), colon

71   brush (n=5), and saliva (n=5). False discovery rate corrected q-value < 0.001 between

72   saliva and every other group. q-value > 0.05 between colon, colon brush, TI brush, and

73   TI wash samples. (B). Faith's Phylogenic Diversity (metrics of α-diversity) at

74   sequencing depth 80,000. Samples grouped by TI wash (n=4), TI brush (n=4), colon

75   wash (n=5), colon brush (n=5), and saliva (n=5). There was no significance between

76   any groups (q>0.05).

77

78   **Figure S9. Diversity analysis of all subjects by individual patients**. Principal

79   coordinates analysis (PCoA) plot of weighted UniFrac distances (metrics of β-

80   diversity). There was no significance between any groups (q>0.05).

81

82   **Table S1. Patients' clinical characteristics.**

83

84   **Table S2. Sequencing count of samples through each step in the DADA2 pipeline.**
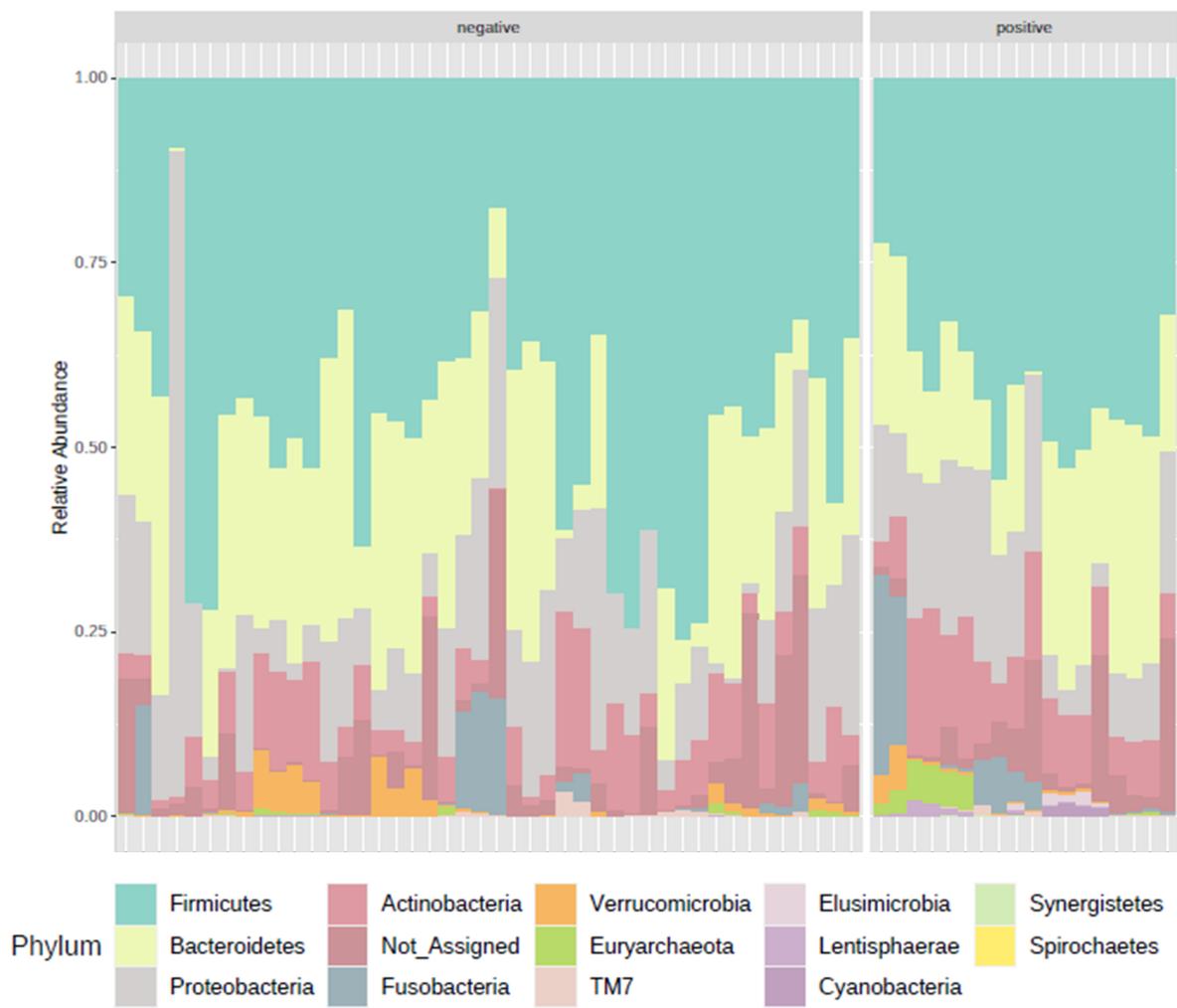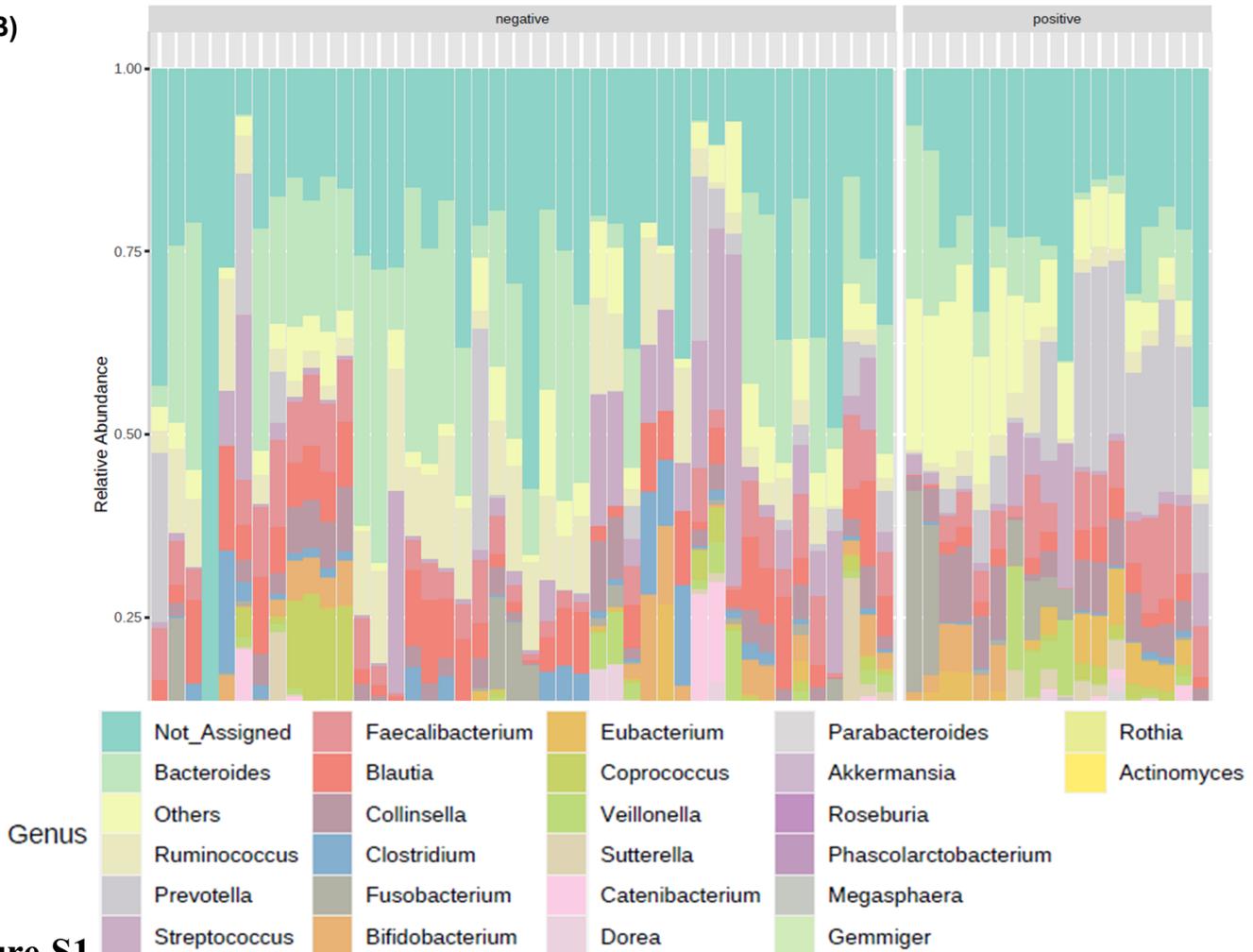
85

86   **Table S3. Sequencing count of amplicon sequence variants (ASVs) in all sample**

87   **types.**

88

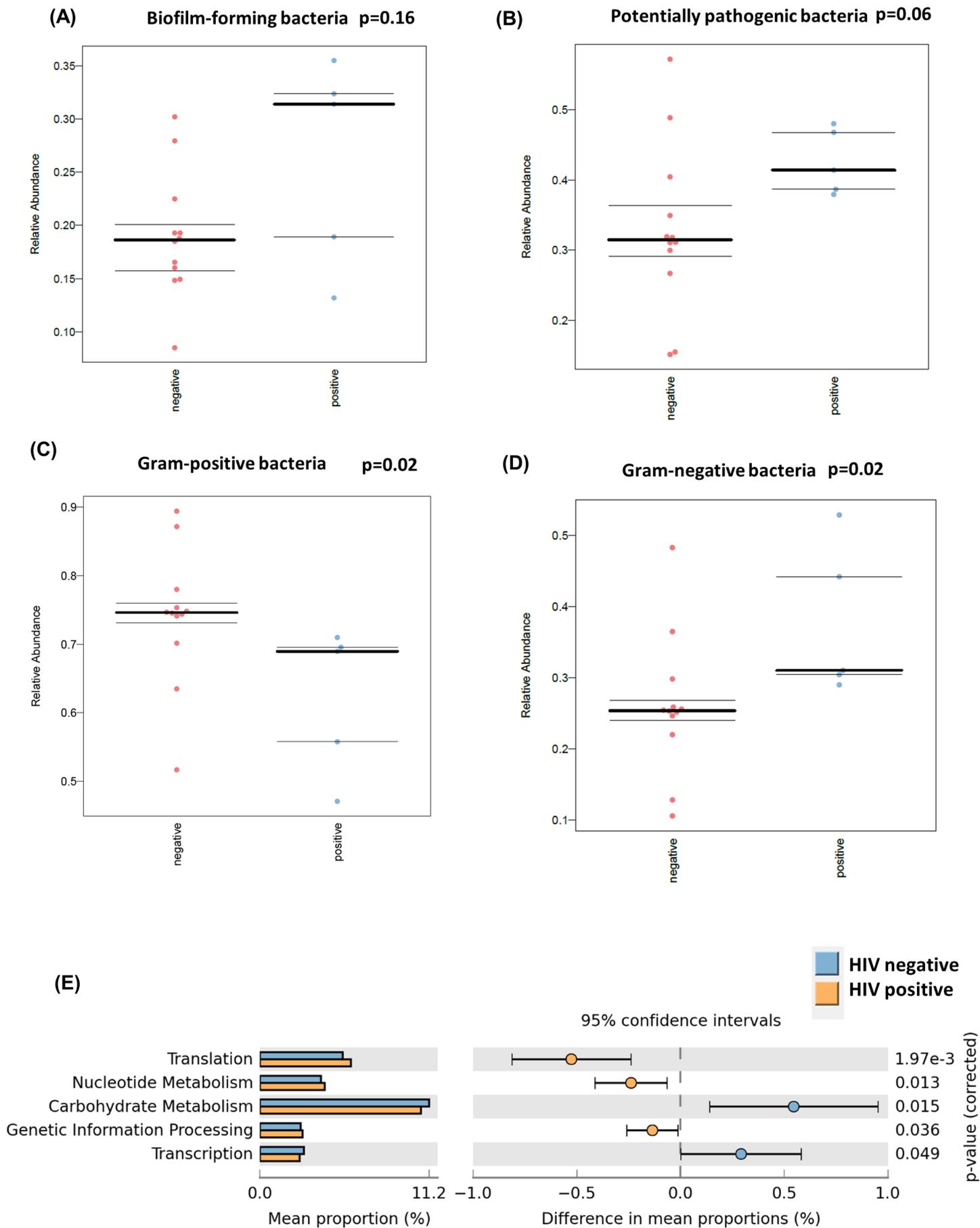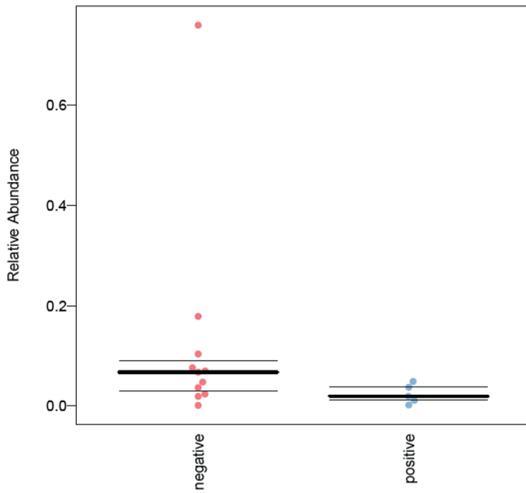89   **Table S4. Predicted pathway by PICRUSt.**

90

3

(A)

Phylum				
Firmicutes	Actinobacteria	Verrucomicrobia	Elusimicrobia	Synergistetes
Bacteroidetes	Not_Assigned	Euryarchaeota	Lentisphaerae	Spirochaetes
Proteobacteria	Fusobacteria	TM7	Cyanobacteria	

(B)

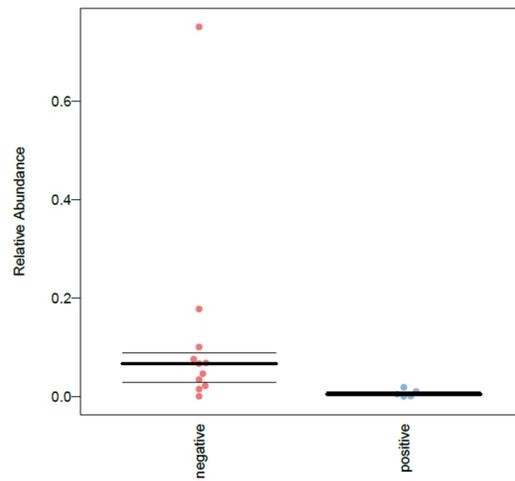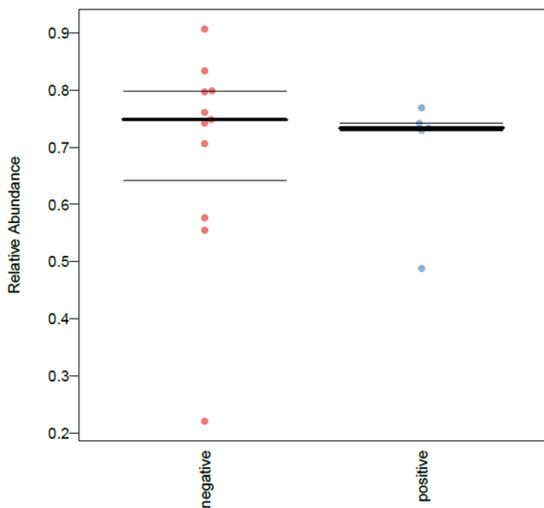| Genus				
Not_Assigned	Faecalibacterium	Eubacterium	Parabacteroides	Rothia
Bacteroides	Blautia	Coprococcus	Akkermansia	Actinomyces
Others	Collinsella	Veillonella	Roseburia	
Ruminococcus	Clostridium	Sutterella	Phascolarctobacterium	
Prevotella	Fusobacterium	Catenibacterium	Megasphaera	
Streptococcus	Bifidobacterium	Dorea	Gemmiger	

Figure S1

Figure S2

(A) Biofilm-forming bacteria   p=0.11

(B) Potentially pathogenic bacteria   p=0.008

(C) Gram-positive bacteria   p=0.44

(D) Gram-negative bacteria   p=0.44

(E)

95% confidence intervals

HIV negative
HIV positive

	Mean proportion (%)	Difference in mean proportions (%)	p-value (corrected)
Translation			2.02e-3
Nucleotide Metabolism			3.93e-3
"Folding, Sorting and Degradation"			6.84e-3
Cellular Processes and Signaling			9.10e-3
Carbohydrate Metabolism			9.20e-3
Membrane Transport			0.018
Transcription			0.019
Environmental Adaptation			0.022
Metabolism of Cofactors and Vitamins			0.025
Metabolism of Terpenoids and Polyketides			0.029
Replication and Repair			0.029
Cancers			0.038

Figure S3
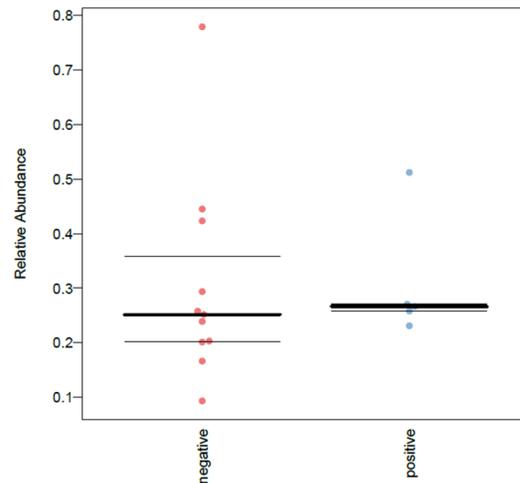
(A) Biofilm-forming bacteria p=0.14

(B) Potentially pathogenic bacteria p=0.004

(C) Gram-positive bacteria p=0.04
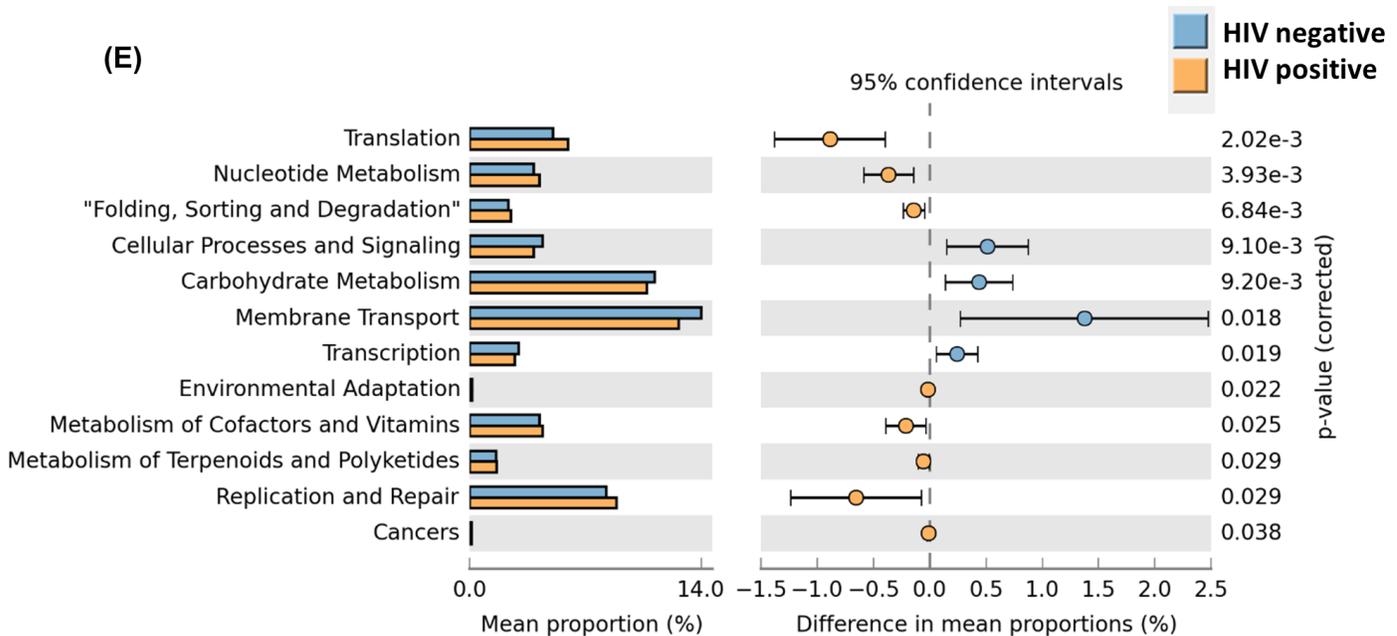
(D) Gram-negative bacteria p=0.04
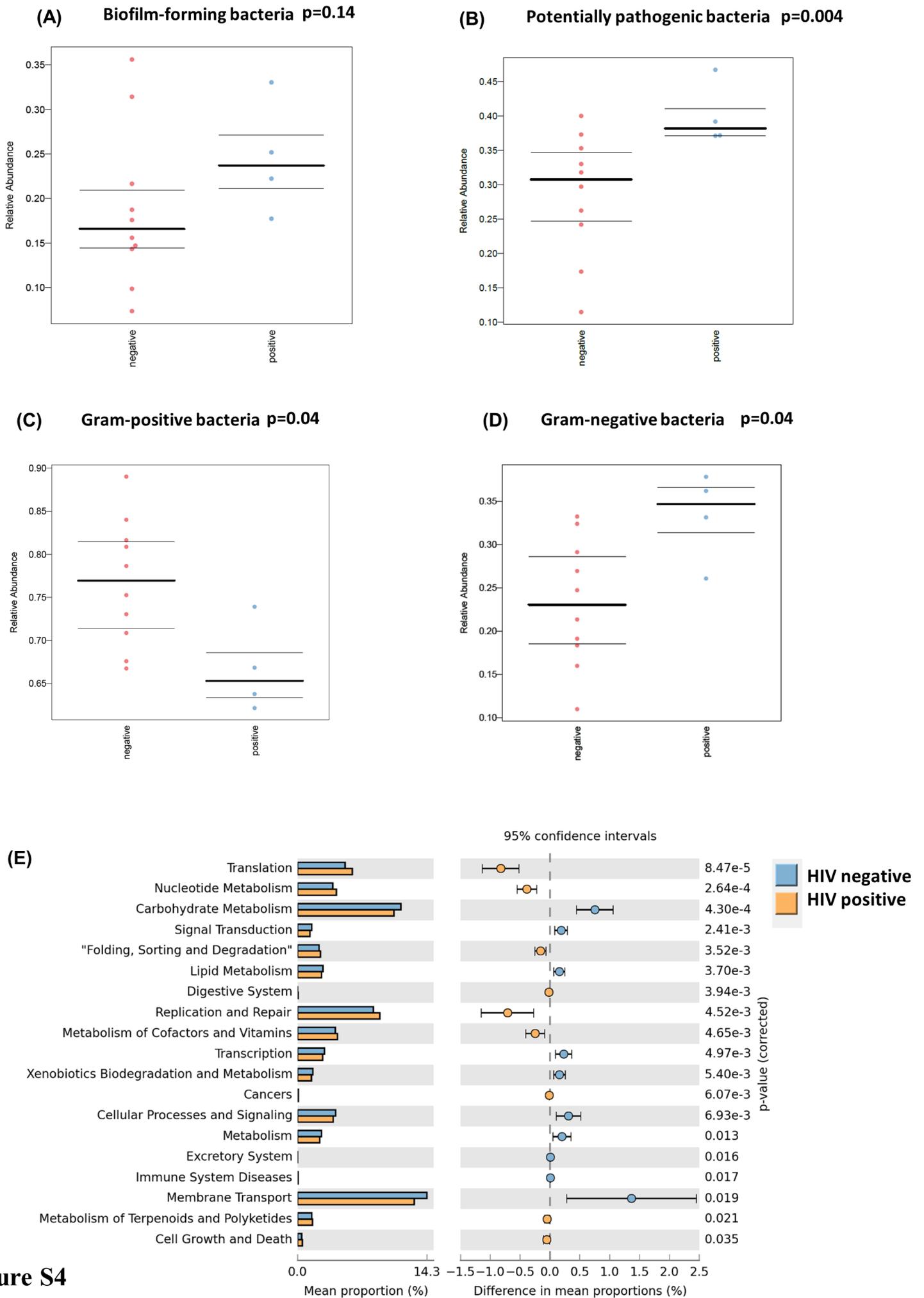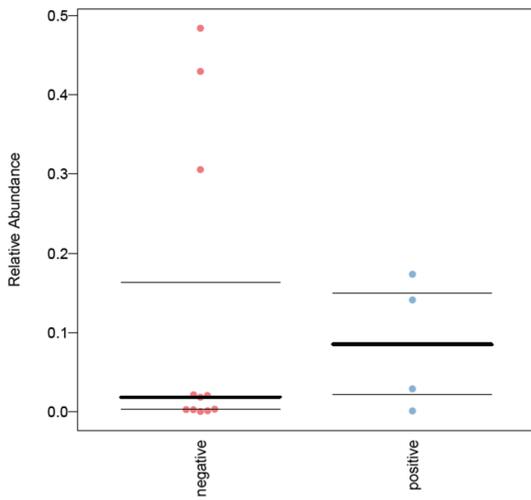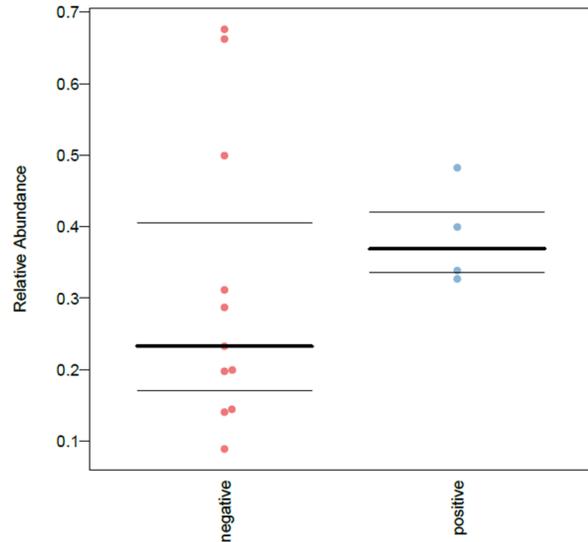
(E)

95% confidence intervals

	Mean proportion (%)	Difference in mean proportions (%)	p-value (corrected)
Translation			8.47e-5
Nucleotide Metabolism			2.64e-4
Carbohydrate Metabolism			4.30e-4
Signal Transduction			2.41e-3
"Folding, Sorting and Degradation"			3.52e-3
Lipid Metabolism			3.70e-3
Digestive System			3.94e-3
Replication and Repair			4.52e-3
Metabolism of Cofactors and Vitamins			4.65e-3
Transcription			4.97e-3
Xenobiotics Biodegradation and Metabolism			5.40e-3
Cancers			6.07e-3
Cellular Processes and Signaling			6.93e-3
Metabolism			0.013
Excretory System			0.016
Immune System Diseases			0.017
Membrane Transport			0.019
Metabolism of Terpenoids and Polyketides			0.021
Cell Growth and Death			0.035

HIV negative
HIV positive
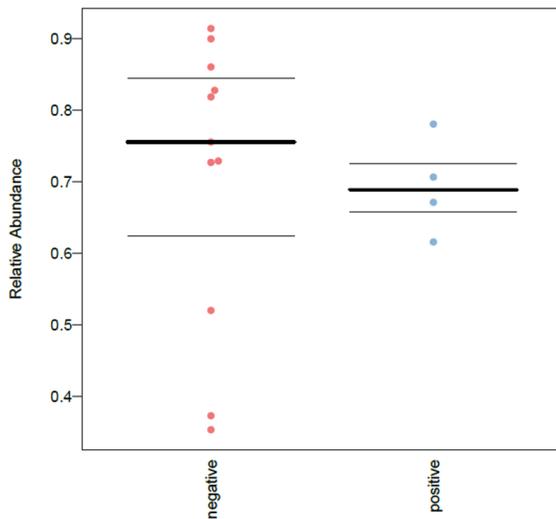
Figure S4

(A) Biofilm-forming bacteria   p=0.75

(B) Potentially pathogenic bacteria   p=0.22

(C) Gram-positive bacteria   p=0.41
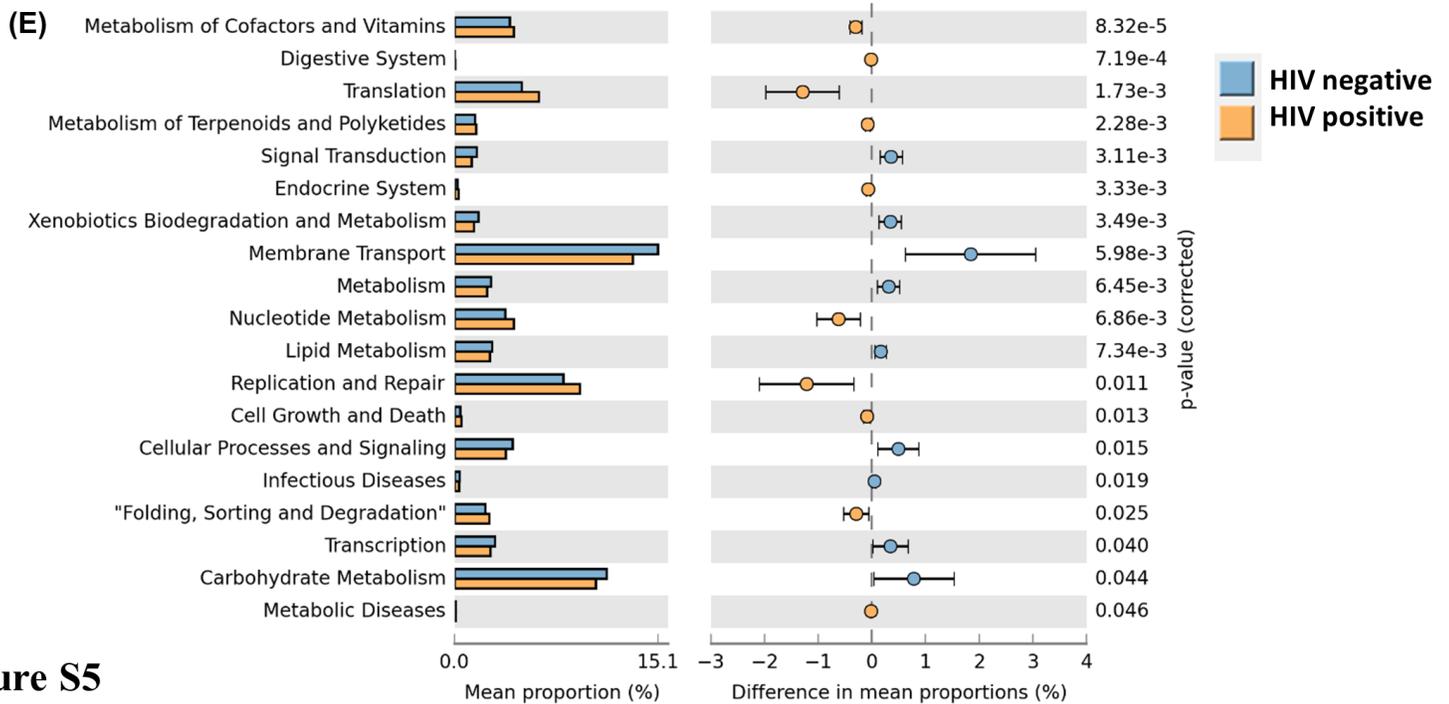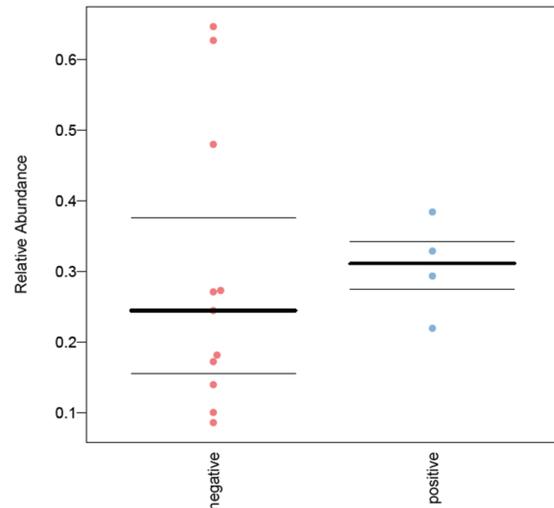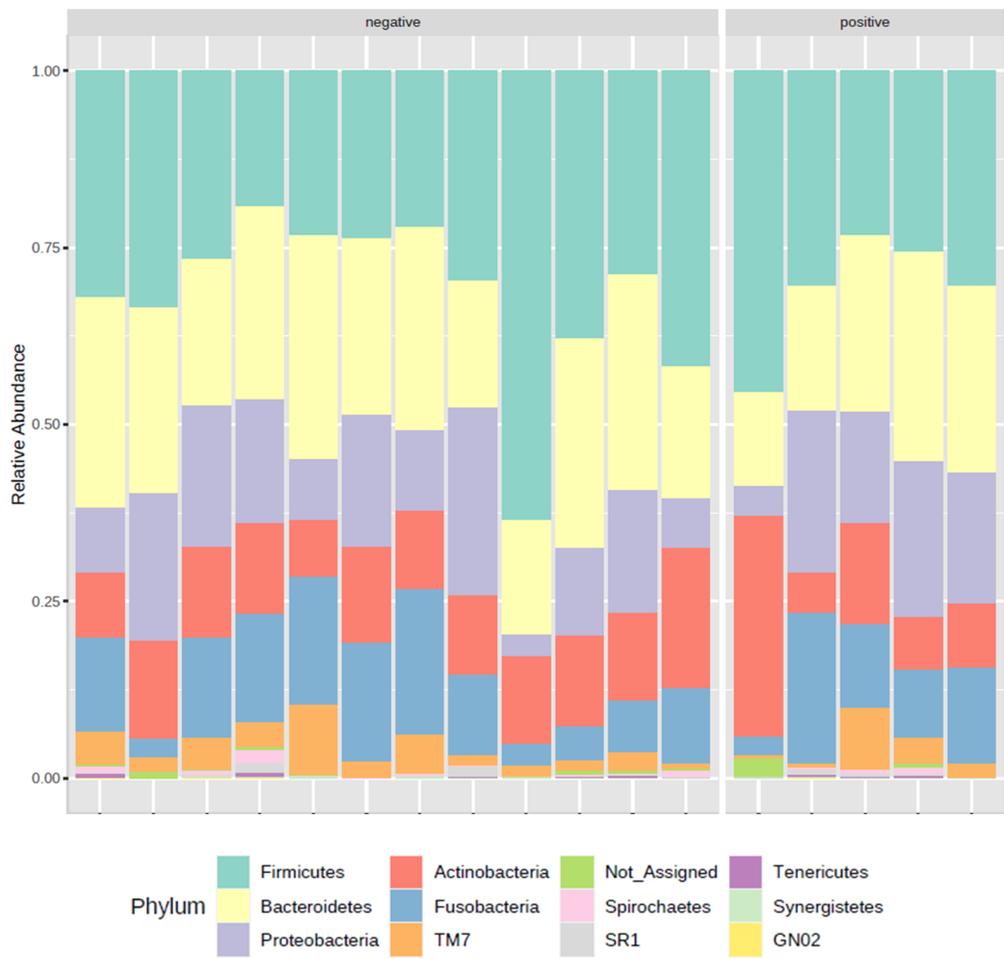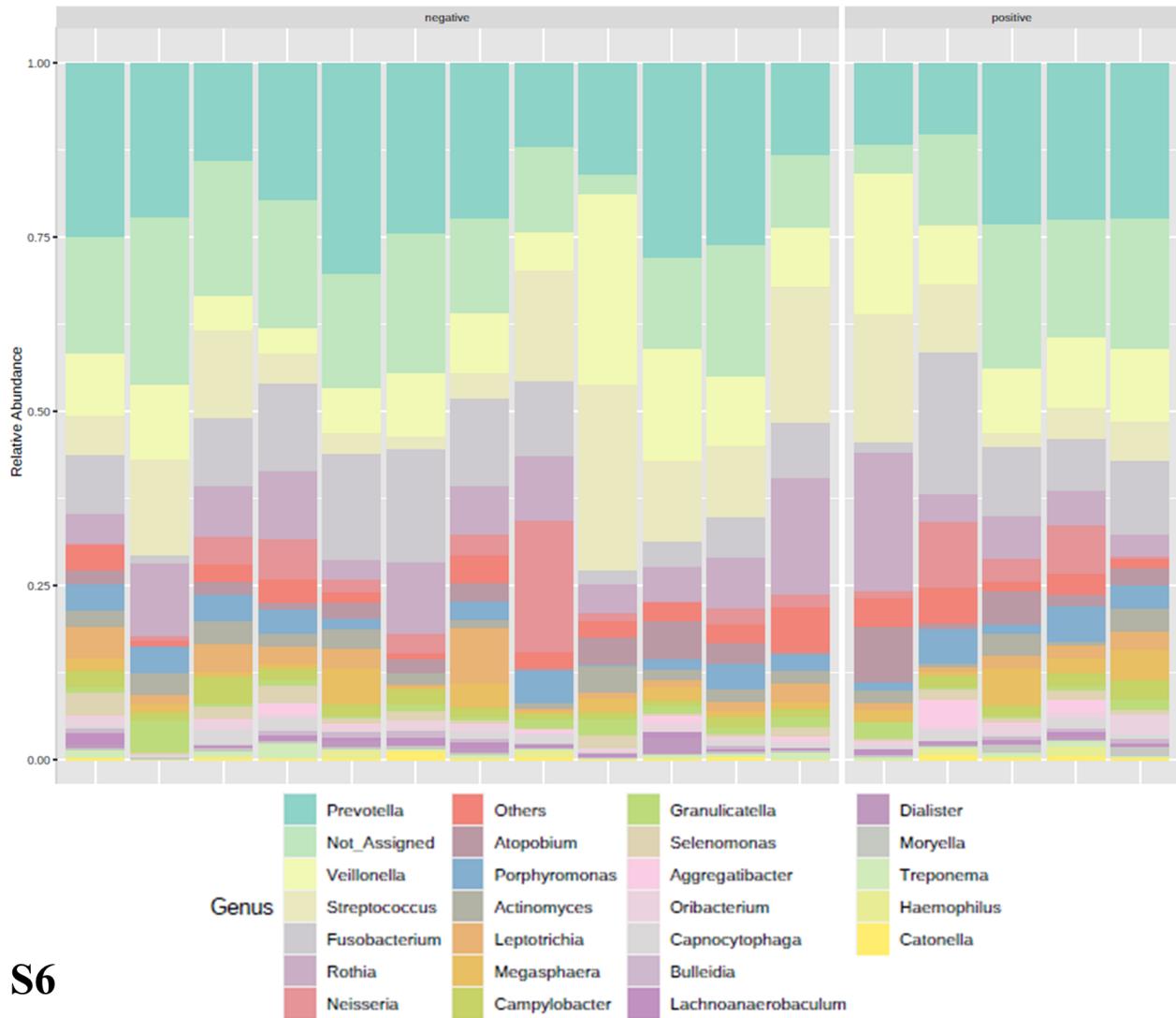
(D) Gram-negative bacteria   p=0.41

(E)

95% confidence intervals

	Mean proportion (%)	Difference in mean proportions (%)	p-value (corrected)
Metabolism of Cofactors and Vitamins			8.32e-5
Digestive System			7.19e-4
Translation			1.73e-3
Metabolism of Terpenoids and Polyketides			2.28e-3
Signal Transduction			3.11e-3
Endocrine System			3.33e-3
Xenobiotics Biodegradation and Metabolism			3.49e-3
Membrane Transport			5.98e-3
Metabolism			6.45e-3
Nucleotide Metabolism			6.86e-3
Lipid Metabolism			7.34e-3
Replication and Repair			0.011
Cell Growth and Death			0.013
Cellular Processes and Signaling			0.015
Infectious Diseases			0.019
"Folding, Sorting and Degradation"			0.025
Transcription			0.040
Carbohydrate Metabolism			0.044
Metabolic Diseases			0.046

HIV negative
HIV positive

Figure S5

(A)

(B)

Figure S6

(A)

(B)

Figure S7

(A)

Axis 2 (19.02 %)

Axis 1 (42.81 %)

	colon
	colonbrush
	saliva
	TI
	TIbrush

(B)

Figure S8

Axis 2 (19.02 %)

Axis 1 (42.81 %)

Patient 1
Patient 2
Patient 3
Patient 4
Patient 5
Patient 6
Patient 7
Patient 8
Patient 9
Patient 10
Patient 11
Patient 12
Patient 21
Patient 22
Patient 23
Patient 24
Patient 25

Figure S9