Supplemental Results:

Results of 16S rRNA analysis after the filtering procedure of taxonomic data

One sample was removed after NG-Tax pipeline, because of its low number of final total reads. The NG-Tax sequencing results were represented by a total of 1 700 925 reads corresponding to 2 186 OTUs which represented 176 genera (Figure S1). After filtering, the taxonomic data was represented by 1 659 853 reads corresponding to 1 965 OTUs within 77 genera. Removed genera (56.25%) were represented by 2.4% of the total reads and 0.18% of total relative abundance. One sample was removed due to having less than 10% of the total observed genera.

Sequencing depth comparison

After filtering of the taxonomic data, the total reads were compared between the groups. This comparison showed that the sequencing depth was evenly distributed among all groups; ADHD (mean $\text{rank}_{\text{ADHD}}$=52.51; $\text{Median}_{\text{ADHD}}$=12 888), controls (mean $\text{rank}_{\text{controls}}$=51.81; $\text{Median}_{\text{controls}}$=13 189) and subthreshold ADHD (mean $\text{rank}_{\text{subtreshold}}$=51.20; $\text{Median}_{\text{subtreshold}}$=10 503; H=0.25, P=0.988) (Table S1).

NCBI BLAST results

The *Ruminococacceae_UCG_004* genus was represented by two OTUs/sequences: OTU_65143327 and OTU_65143327 with the total reads of 1555 and 53 respectively. Since the *Ruminococacceae_UCG_004* genus was mainly represented by OTU_65143327, the sequences of this OTU were subjected to NCBI BLAST (https://blast.ncbi.nlm.nih.gov) to query for highly similar sequences (the blasted sequences can be found below this paragraph). BLAST indicated that *Ruminococacceae_UCG_004* could be *Evtepia gabavorous* (Accession:

MH636586.1), which we confirmed by re-running the NG-Tax pipeline increasing the read length to 100bp (of each forward and reverse), and we revealed that our targeted sequence is in 100% similar to N=11 uncultured bacteria and one cultured bacteria - *Evtepia gabavorous* (Query Cover: 100%; E-value: 2e-43; Percent Identity: 100%). Moreover, all the uncultured bacteria align with the sequence of *E. gabavorous* (N=10/11 have Percent Identity of 99-100% and N=1/11 has Percent Identity of 94%). This is probably due to the fact that these sequences were deposited in the NCBI before *E. gabavorous* was isolated hence named uncultured.

Blasted sequences:

>OTU 65143327; g__Ruminococcaceae_UCG-004 ; Forward sequence

GATGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGGAGGACCCTTGACGGAGTTTTCGGACAAC

GGATAGGAATCCTTAGTGGCGGACGGGTGA

>OTU 65143327; g__Ruminococcaceae_UCG-004 ; Reverse sequence

CTGGGCCGTGTCTCAGTCCCAATGTGGCCGGTCAACCTCTCAGTCCGGCTACTGATCGTCGCCTTGGTAG

Supplemental Tables and Figures

Figure S1. Filtering procedure applied in our study.

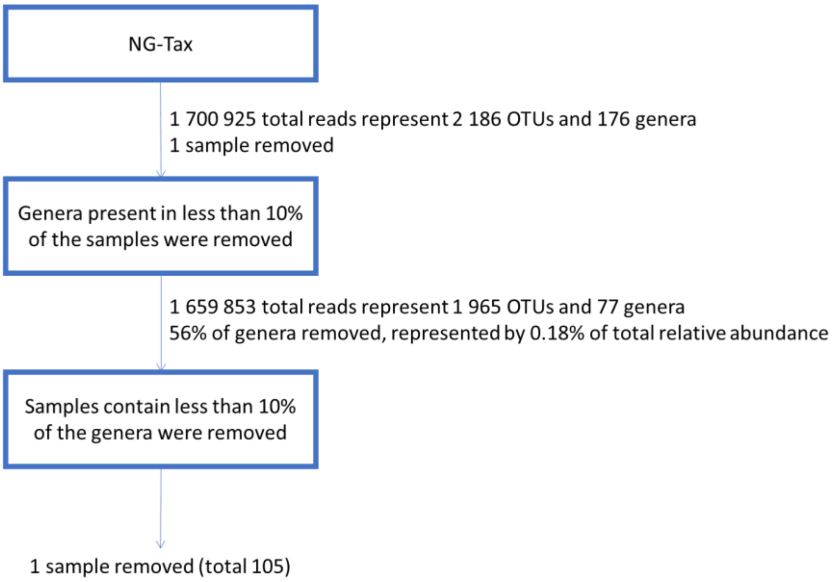


Figure S2. Alpha diversity comparison of the gut microbiome between participants with ADHD and controls. Alpha diversity was indicated by observed OTUs (P=0.08), Shannon index (P=0.16) and Phylogenetic diversity (P=0.15). Box plots represent median with whiskers on ±1.5 IQR.

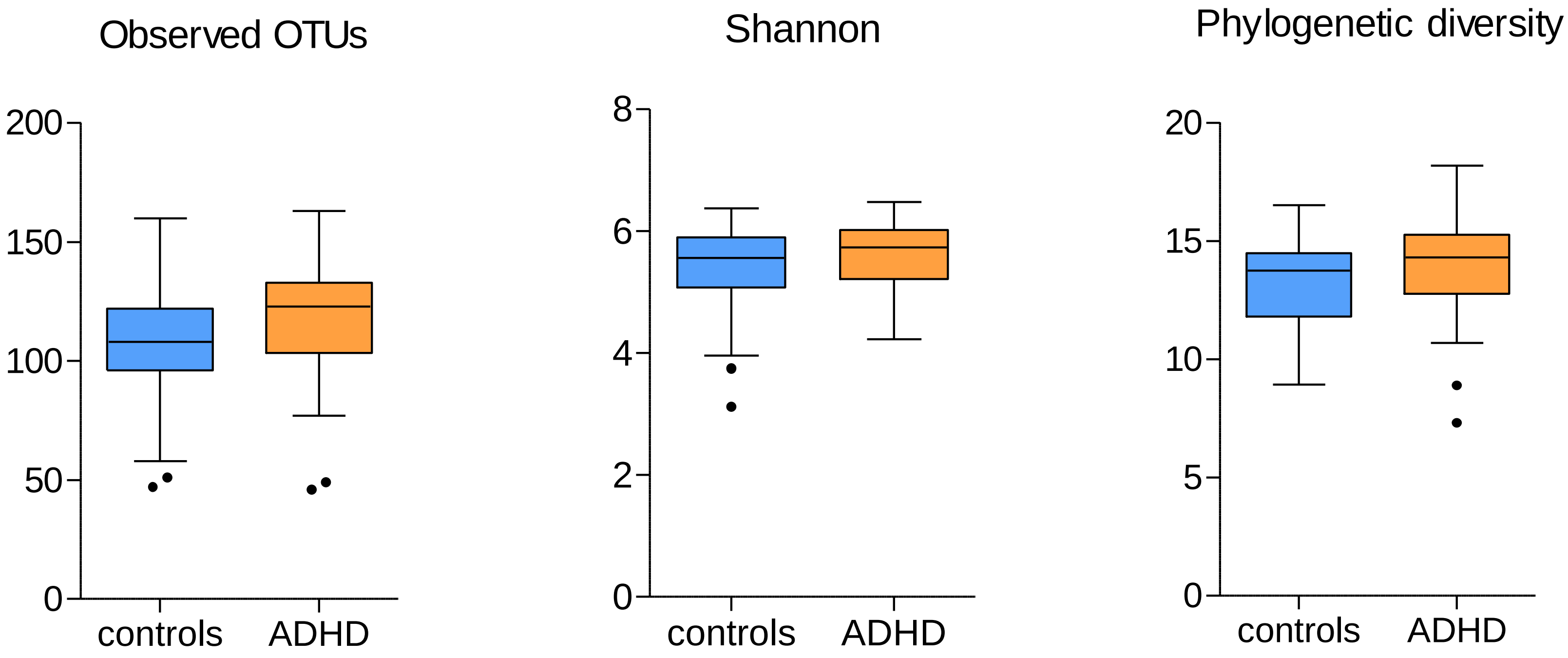


Figure S3. Principal coordinates analysis (PCoA) of weighted UniFrac distances representing microbial composition of participants with ADHD and controls. The first two components are plotted with the percentage of the variance explained by each principal component. Each point represents an individual sample.

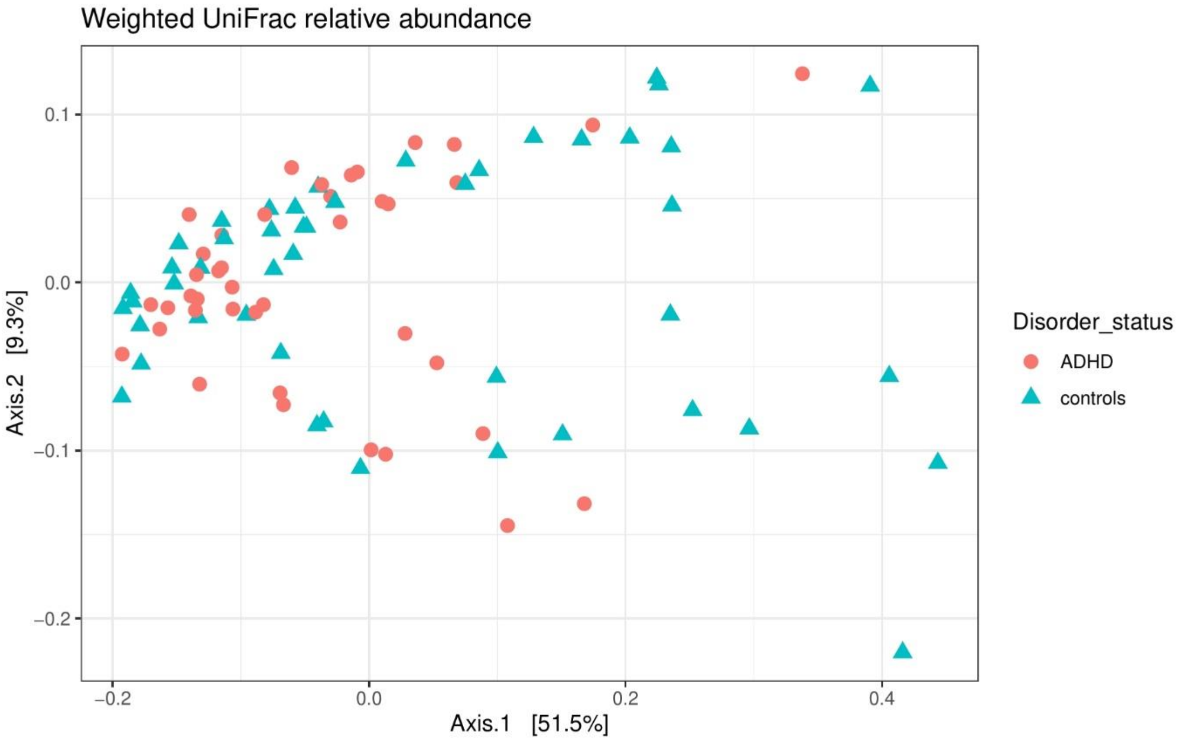
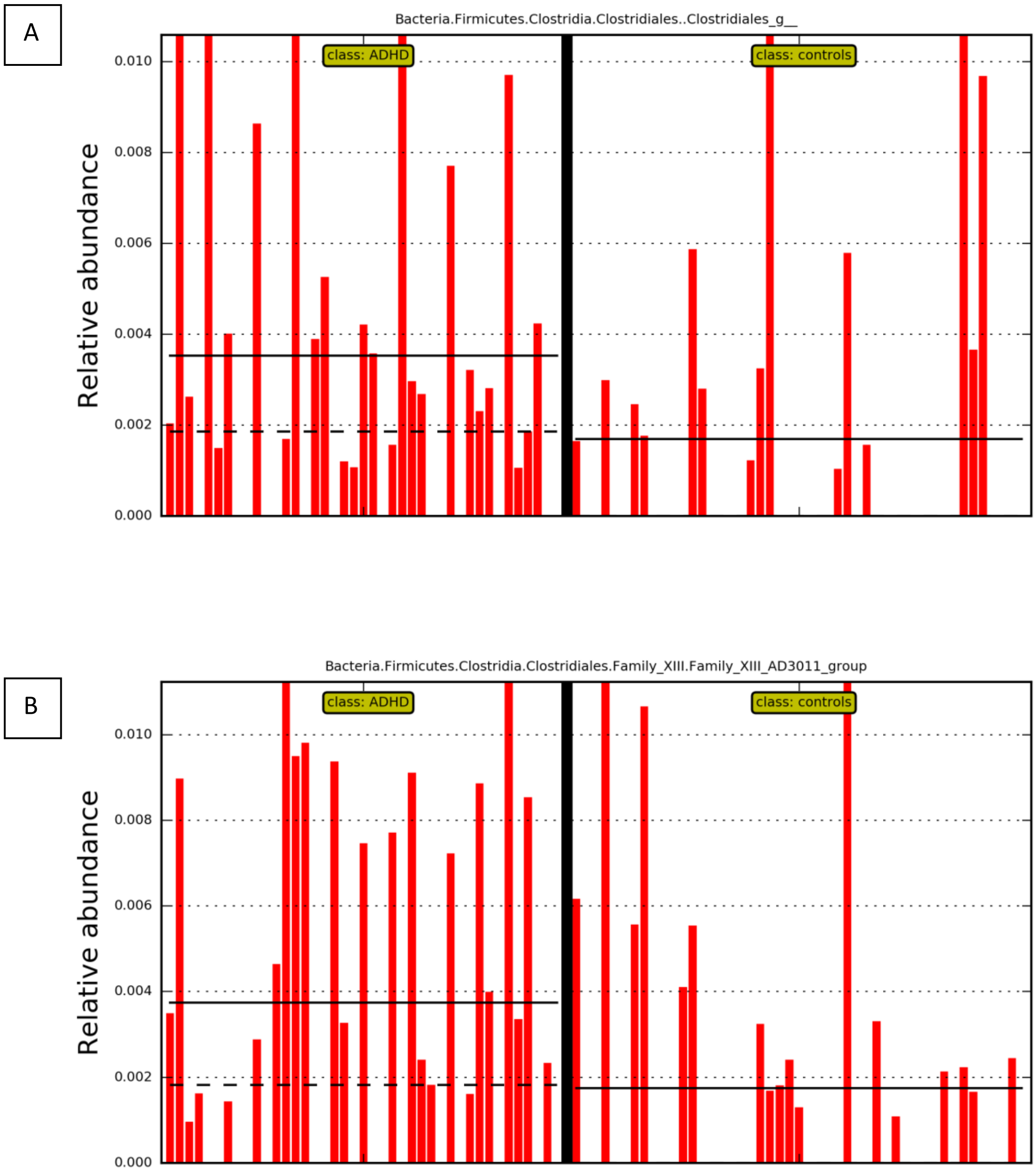
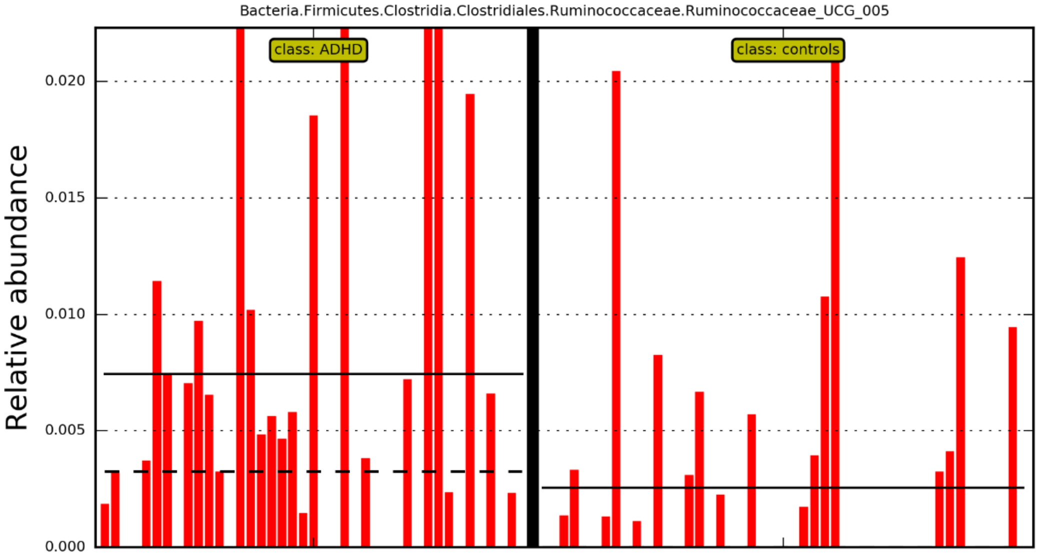


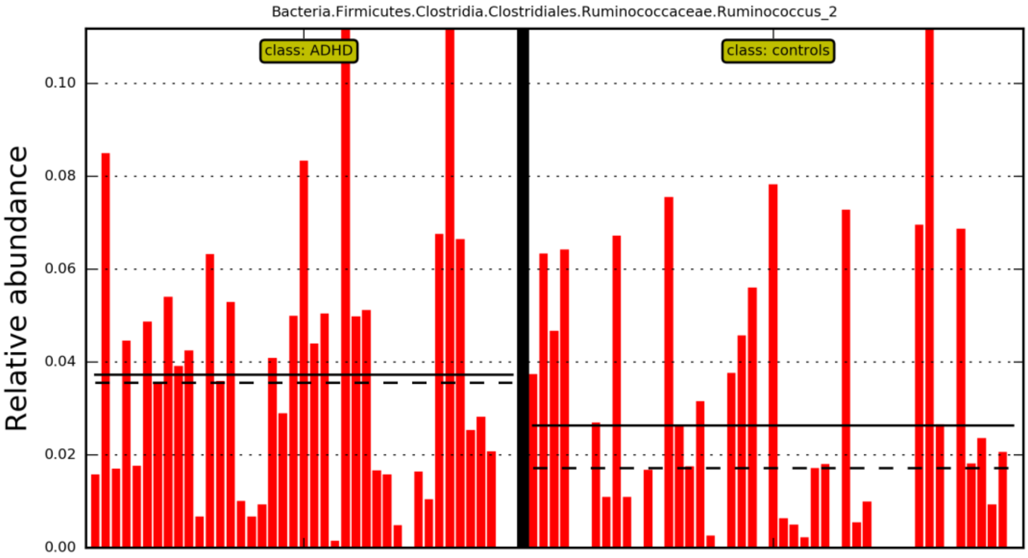
Figure S4. Distribution of the selected genera for behavioral analysis for each group (ADHD samples on the left and controls on the right). Barplots of *Clostridiales_g__* (A), *Family_XIII_AD3011_group* (B), *Ruminococcaceae_UCG_005* (C), *Ruminococcus_2* (D), *Ruminococcaceae_g__* (E), *Ruminococcaceae_NK4A214_group* (F) and *Ruminococcaceae_UCG_004* (G). Each bar represents the sum to 1 of the taxa relative abundance for an individual sample; straight line represents mean and dotted line represents median.



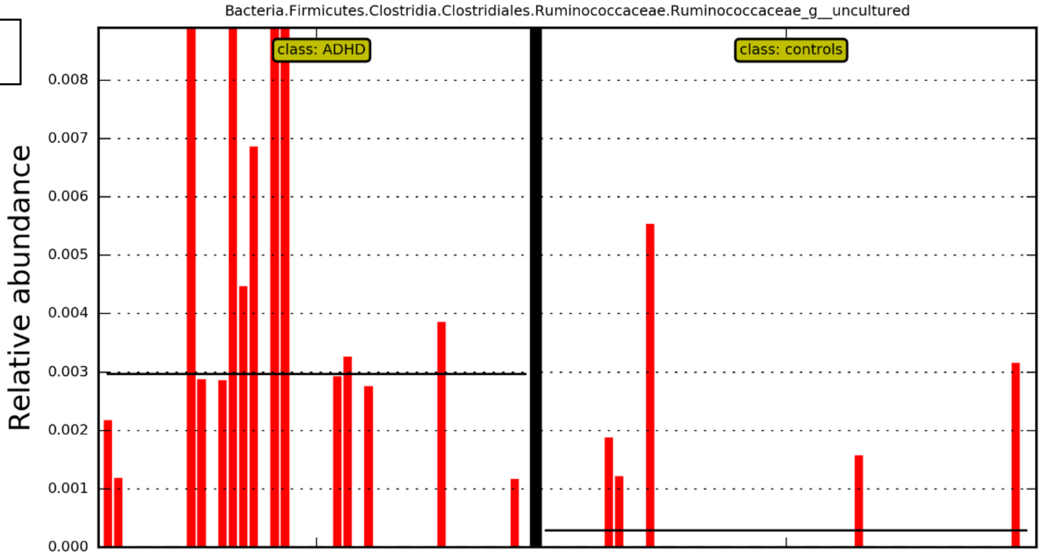
C


Bacteria.Firmicutes.Clostridia.Clostridiales.Ruminococcaceae.Ruminococcaceae_UCG_005
Relative abundance
0.000
0.005
0.010
0.015
0.020
class: ADHD
class: controls

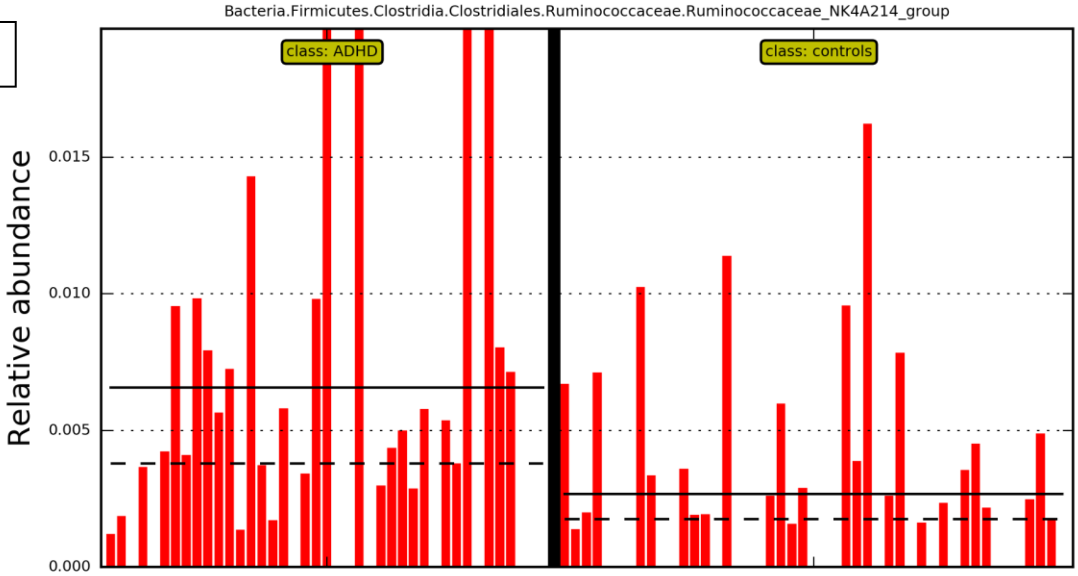
D


Bacteria.Firmicutes.Clostridia.Clostridiales.Ruminococcaceae.Ruminococcus_2
Relative abundance
0.00
0.02
0.04
0.06
0.08
0.10
class: ADHD
class: controls

E


Bacteria.Firmicutes.Clostridia.Clostridiales.Ruminococcaceae.Ruminococcaceae_g__uncultured
class: ADHD
class: controls
Relative abundance
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000

F


Bacteria.Firmicutes.Clostridia.Clostridiales.Ruminococcaceae.Ruminococcaceae_NK4A214_group
class: ADHD
class: controls
Relative abundance
0.015
0.010
0.005
0.000

G

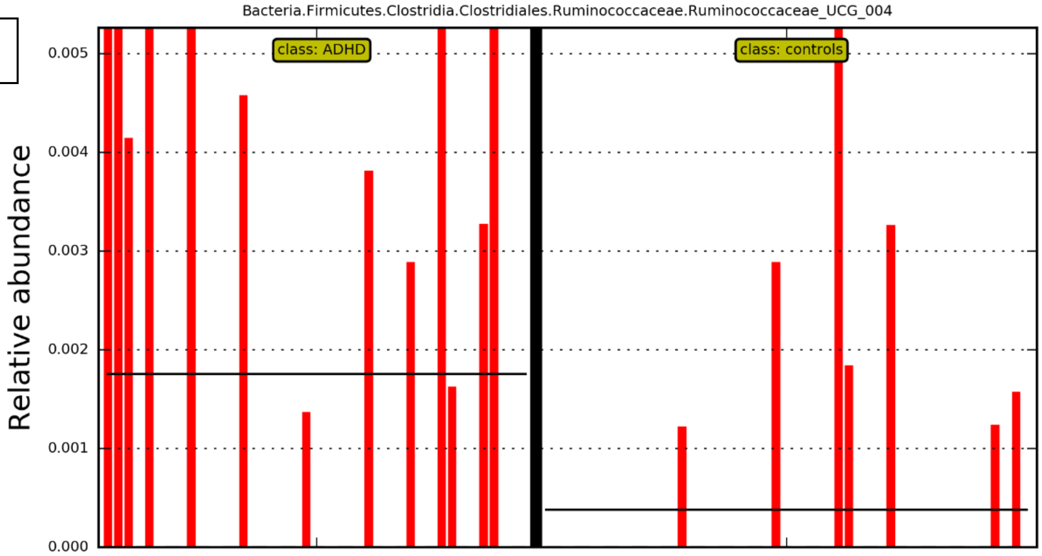
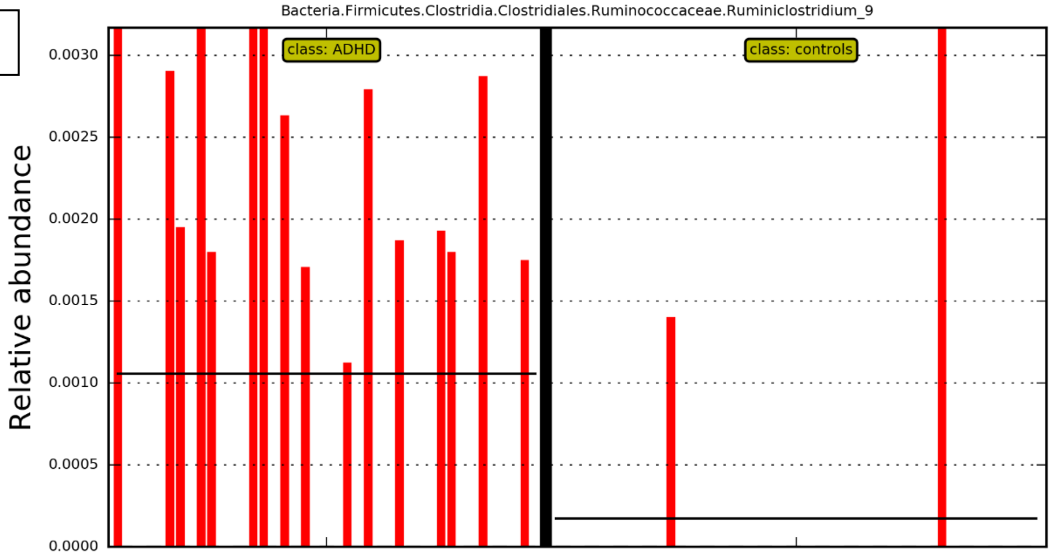
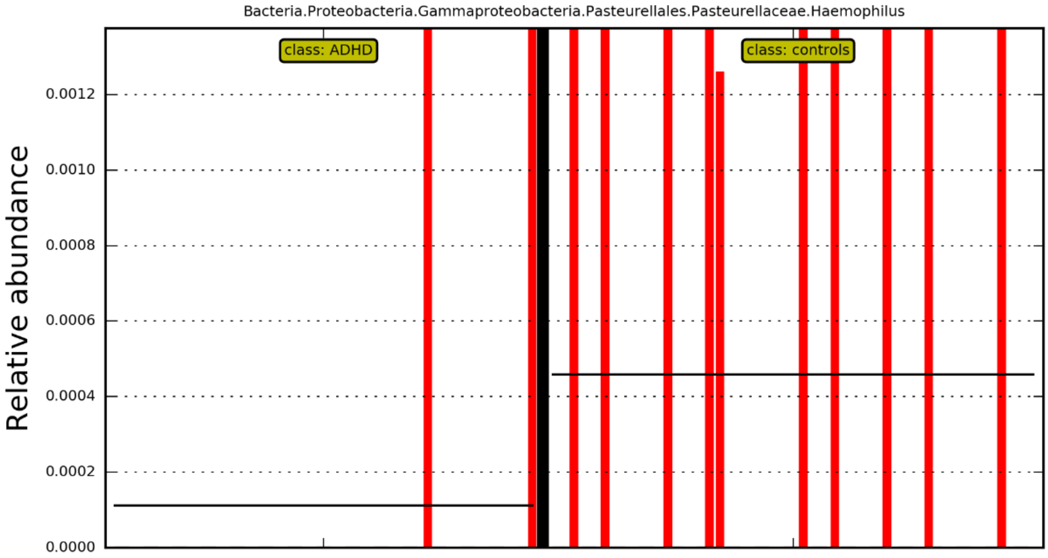

Bacteria.Firmicutes.Clostridia.Clostridiales.Ruminococcaceae.Ruminococcaceae_UCG_004
Relative abundance
0.005
0.004
0.003
0.002
0.001
0.000
class: ADHD
class: controls

Figure S5. Distribution of the non-selected genera for behavioral analysis for each group (ADHD samples on the left and controls on the right). Barplots of *Ruminiclostridium_9* (A), *Haemophilus* (B) and *Ruminococcaceae_UCG-003* (C). Each bar represents the sum to 1 of the taxa relative abundance for an individual sample; straight line represents mean and dotted line represents median.

A



B



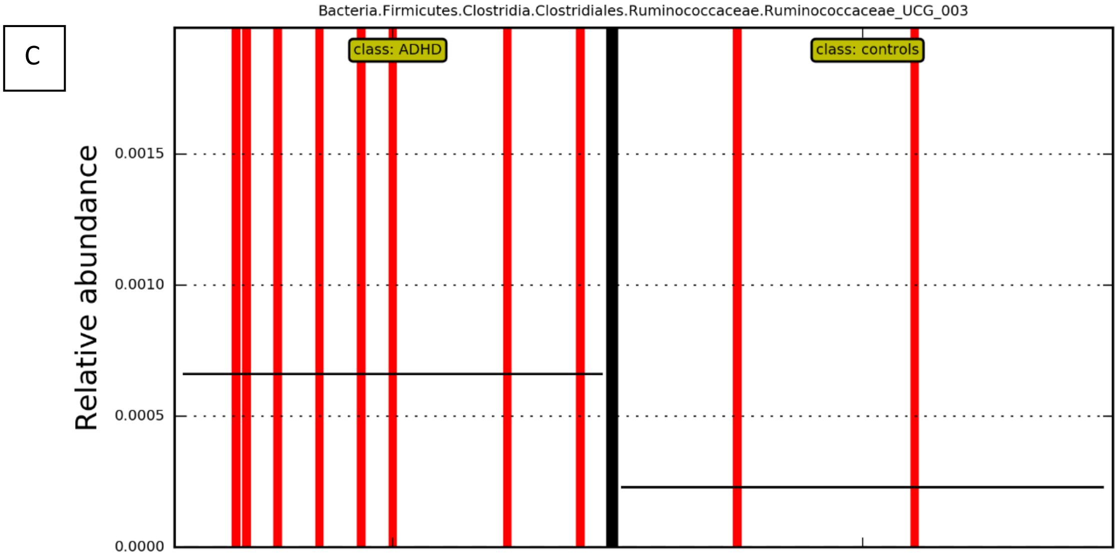

C
Bacteria.Firmicutes.Clostridia.Clostridiales.Ruminococcaceae.Ruminococcaceae_UCG_003
class: ADHD
class: controls
Relative abundance
0.0015
0.0010
0.0005
0.0000

Figure S6. Boxplot of two genera relative abundance being different between medicated and non-medicated participants with ADHD. Box plots represent median with whiskers on ±1.5 IQR.

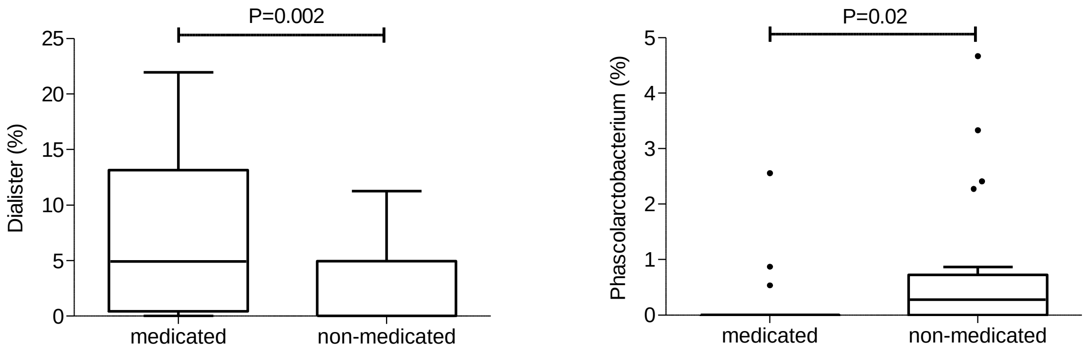
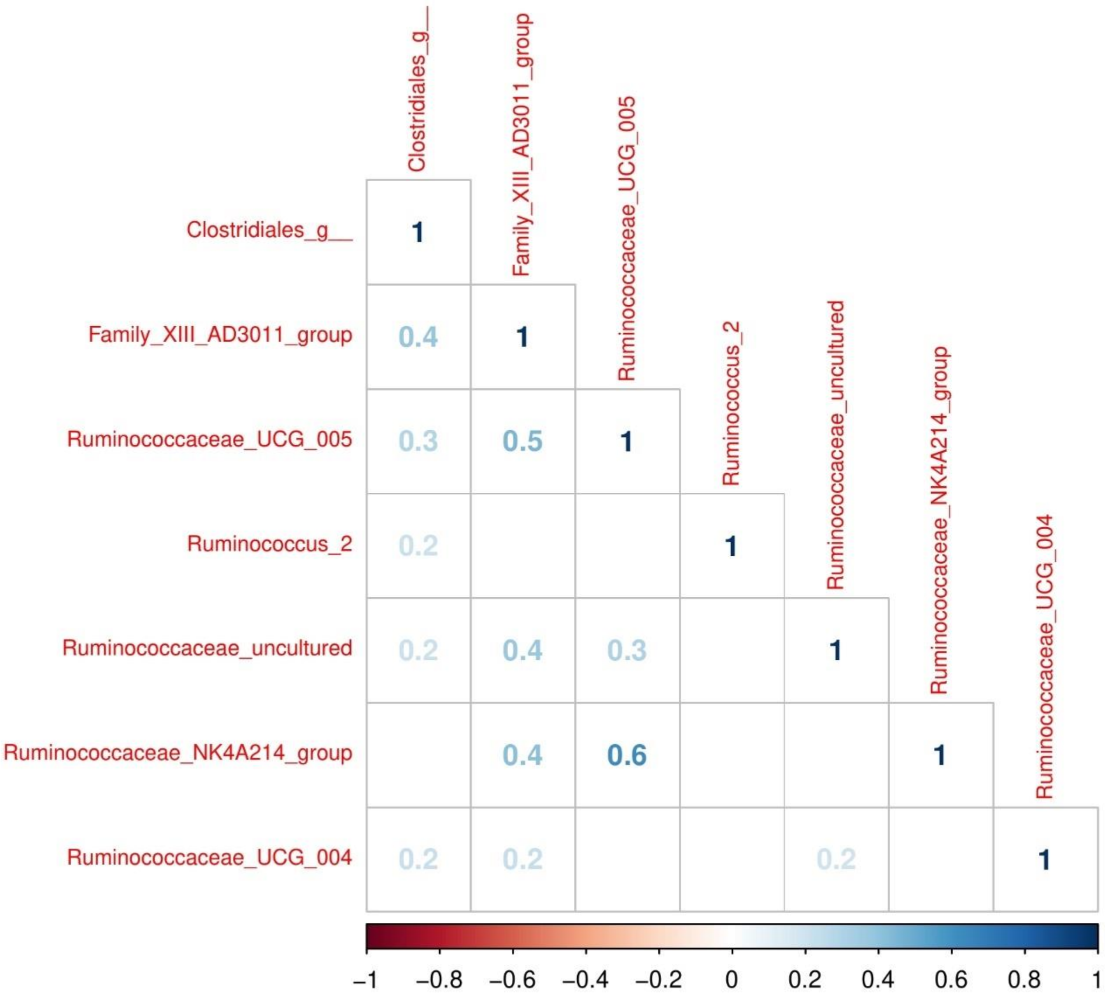


Figure S7. Spearman's rank correlation matrix of the selected bacterial taxa. The selection of the genera was done prior to correlation analysis and it was done based on their prevalence (see the method section). The matrix represents only nominal significant (P<0.05) correlations.



	Clostridiales_g__	Family_XIII_AD3011_group	Ruminococcaceae_UCG_005	Ruminococcus_2	Ruminococcaceae_uncultured	Ruminococcaceae_NK4A214_group	Ruminococcaceae_UCG_004
Clostridiales_g__	1						
Family_XIII_AD3011_group	0.4	1					
Ruminococcaceae_UCG_005	0.3	0.5	1				
Ruminococcus_2	0.2			1			
Ruminococcaceae_uncultured	0.2	0.4	0.3		1		
Ruminococcaceae_NK4A214_group		0.4	0.6			1	
Ruminococcaceae_UCG_004	0.2	0.2			0.2		1

Table S1. Total number of reads of participants with ADHD, controls and subthreshold ADHD.

	Total number of reads			
	N	Median (IQR)	Mean rank[a]	P-value[a]
ADHD	41	12888 (6041 – 21569)	52.51	.988
controls	47	13819 (7605 – 21413)	51.81	
subthreshold ADHD	15	10503 (6280 -21290)	51.20	

[a] calculated based on the Kruskal-Wallis Test

Table S2. Relative abundance of bacterial phyla of all participants in our study.

Phylum's relative abundance and FBR					
	ADHD median (IQR)	ADHD Mean rank[a]	controls median (IQR)	controls Mean rank[a]	P-values[a]
Firmicutes	77.25 (71.63 – 87.06)	49.63	78.85 (52.19 – 83.04)	40.02	0.078
Bacteroidetes	16.91 (8.56 – 25.63)	40.37	18.21 (10.93 – 47.49)	48.11	0.156
Actinobacteria	0.83 (.25 – 2.35)	45.98	0.71 (0.12 – 2.61)	43.21	0.612
Proteobacteria	0.14 (0 – 0.64)	46.06	0.18 (0 – .43)	43.14	0.579
Verrucomicrobia	0.11 (0 – 0.33)	44.26	0.11 (0 – .51)	44.71	0.929
FBR	4.61 (2.84 – 9.78)	47.70	4.44 (1.10 – 7.89)	40.85	0.208

N=88; [a] calculated based on the Mann-Whitney U test; FBR = *Firmicutes* to *Bacteroidetes* ratio; IQR = interquartile range

Table S3. Multiple regression model with the selected genera on inattention.

	Inattention			Results taken from Table3^b	
	B (S.E.)	95% CI	P-value	B (S.E.)	P-value
Clostridiales_g__	-4.312 (3.571)	-11.143 - 3.037	0.231	-1.467 (3.077)	0.634
Family_XIII_AD3011_group	-0.269 (4.663)	-8.888 – 8.423	0.954	5.323 (2.779)	0.059
Ruminococcaceae_UCG_005	-1.509 (3.053)	-7.363 - 4.205	0.622	1.495 (1.647)	0.367
Ruminococcus_2	1.525 (0.454)	0.675 - 2.362	0.001	1.098 (0.445)	0.016
Ruminococcacea_g__uncultured	6.081 (6.305)	-5.601 - 18.013	0.338	12.241 (5.011)	0.017
Ruminococcaceae_NK4A214_group	4.279 (2.651)	-0.805 - 9.481	0.111	3.392 (1.860)	0.072
Ruminococcaceae_UCG_004	43.920 (13.348)	19.372 - 69.095	0.001	39.291 (12.296)	0.002

The model includes all selected genera in one model; The identified outliers where removed from this analysis, N=89; The model was adjusted for age, sex, BMI, diff_days and a random factor for family relatedness; ^{b} Results from simple regression models (taken from Table 3 for the purpose of easier comparison).