Supplementary Data

Data S1: Additional information microbiome

S1A. Gut microbiome sample characteristics

The 16S marker gene 454 pyrosequencing run was performed in a total of 96 samples and yielded a total of 363,461 high quality reads with good sequencing depth; roughly 80% of the reads could be traced back to a specific bacterial genus. No apparent differences in either alpha or beta diversity between the microbiome of ADHD cases and controls could be observed (see main Table 1 for Shannon and Chao index; beta diversity not reported). For further details, see Table S1: Microbiome descriptives**.**

S1B. Age-matched sub-sample of ADHD cases and controls

***Table*.** Age-matched sub-sample

Pair *	Sample ID	Age (years)	Gender	Diagnosis
Pair 1	ADHD014	15	female	CONTROL
	ADHD026	15	female	CASE
Pair 2	ADHD041	16	male	CONTROL
	ADHD020	16	female	CASE
Pair 3	ADHD008	17	female	CONTROL
	ADHD037	17	female	CASE
Pair 4	ADHD002	18	female	CONTROL
	ADHD092	18	female	CASE
Pair 5	ADHD096	19	male	CONTROL
	ADHD018	19	female	CASE
Pair 6	ADHD003	20	male	CONTROL
	ADHD029	20	male	CASE
Pair 7	ADHD090	20	male	CONTROL
	ADHD099	20	male	CASE
Pair 8	ADHD093	20	female	CONTROL
	ADHD048	20	male	CASE
Pair 9	ADHD052	20	female	CONTROL
	ADHD044	20	female	CASE
Pair 10	ADHD024	21	male	CONTROL
	ADHD038	21	male	CASE
Pair 11	ADHD021	21	male	CONTROL
	ADHD035	21	male	CASE
Pair 12	ADHD040	21	female	CONTROL
	ADHD034	21	male	CASE
Pair 13	ADHD071	22	male	CONTROL
	ADHD045	22	male	CASE
Pair 14	ADHD028	23	male	CONTROL
	ADHD019	23	male	CASE
Pair 15	ADHD047	25	female	CONTROL
	ADHD095	25	male	CASE

* age-matched pairs; four ADHD subjects did not have a matching control subject

S1C. Gut microbiome metabolic potential: logistic regression including age

A logistic regression analysis on all 15 present candidate reactions was performed to assess the effect of age in our cohort. Age was found to have a significant effect in all candidate reactions when including all controls; $p < 0.05$). The logistic regression model used case control status as dependent variable and age, age^2, gender and the relative abundance of the candidate reactions as covariates (age^2 was added in order to account better for the age effect). For the analysis with the complete sample, reaction K01713 showed suggestive evidence for association ($p = 0.070$, OR=1.921, 95% confidence interval (CI) (0.948-3.895); n = 96). After removing the BIG controls the association between CDT (K01713) and ADHD risk was significant ($p = 0.024$, OR=2.786, 95% CI (1.144-6.785); n = 57).

S1D. Relation between taxa and reactions

To investigate which taxa from the above analysis contributed the most to the phenylalanine biosynthesis reaction that differed significantly between ADHD cases and controls, we repeated the functional (PICRUSt) analysis only on the five candidate taxa which differed the most between ADHD and controls: Clostridiales (order), Rikenellaceae (family), Porphyromonadaceae (family), *Bifidobacterium* (genus) and *Eggerthella* (genus) (main Figure 3). For this selection of taxa, we again observed a similar increase in ADHD of cyclohexadienyl dehydratase (K01713) as compared to the total microbiome (on average, 150% of control relative abundance; $p = 0.038$, BF corrected). Strikingly, we found that the cyclohexadienyl dehydratase ortholog (K01713) was solely present in Actinobacteria and its descending genus *Bifidobacterium*, not in any of the four candidate taxa. Specifically, in an additional PICRUSt functional analysis selecting only the genus *Bifidobacterium*, we observed that the total number of absolute counts for K01713 for the control and ADHD group combined was the same number when examining the analysis for the above described combined selection of five candidate taxa (16,329; representing an average

relative abundance of 0.0044 %); and more strikingly, accounting for 99.9% of the PICRUSt predicted CDT counts in the entire microbiome (Table S2).

Consequently, in pursue of the question which taxon or taxa were contributing the most to the presence of cyclohexadienyl dehydratase in the gut microbiomes, we did a Spearman correlation analysis on K01713 with all separate taxa on all hierarchical levels. Interestingly, a very strong positive correlation was observed for Actinobacteria (phylum/class; rho = 0.97) and *Bifidobacterium* (see *Figure* below) (genus, but also for its family and order; rho = 0.98); all surviving multiple testing correction. Furthermore, a strong negative correlation was observed for Firmicutes (phylum; rho = -0.80) and Clostridiales (order; rho = -0.75); also surviving multiple testing correction (data not shown). However, these latter effects can probably best be explained by the fact that a large increase in relative abundance for Actinobacteria and its descending taxonomies consequently lowers all other phyla in relative abundance, which reasonably affects Firmicutes and its descending taxonomies the most, as Firmicutes are by far the most represented bacteria in the gut microbiomes of our subjects (on average, 79.80% in the control group; main Figure 3 and Table S2).

The idea that the negative correlations between Firmicutes/Clostridiales and CDT (K01713) were not independent from the positive correlation between CDT and *Bifidobacterium* was confirmed by a multiple regression analysis. We took relative abundance of the cyclohexadienyl dehydratase reaction (K01713) as dependent and used the relative abundances of the same five candidate taxa as above (Clostridiales (order), Rikenellaceae (family), Porphyromonadaceae (family), *Bifidobacterium* (genus) and *Eggerthella* (genus)) as predictors. The overall model was significant ($F(4,91)=298.7$, $p < 0.001$, adjusted R square = 0.943), which is in line with the fact that the differences in reactions were determined based on the differences in taxa. However, only the genus *Bifidobacterium* showed a strong positive association with the phenylalanine biosynthesis reaction (standardized beta = 0.94, $t=25.3$, $p < 0.001$). The other candidate taxa, including Clostridiales (phylum Firmicutes), did not show significant associations with the

cyclohexadienyl dehydratase reaction (all standardized betas < 0.05 and > -0.05 and $p > 0.1$) in the multiple regression model. Therefore, differences in relative abundance of genus *Bifidobacterium* exclusively contributed to the observed differences in the phenylalanine biosynthesis pathway.

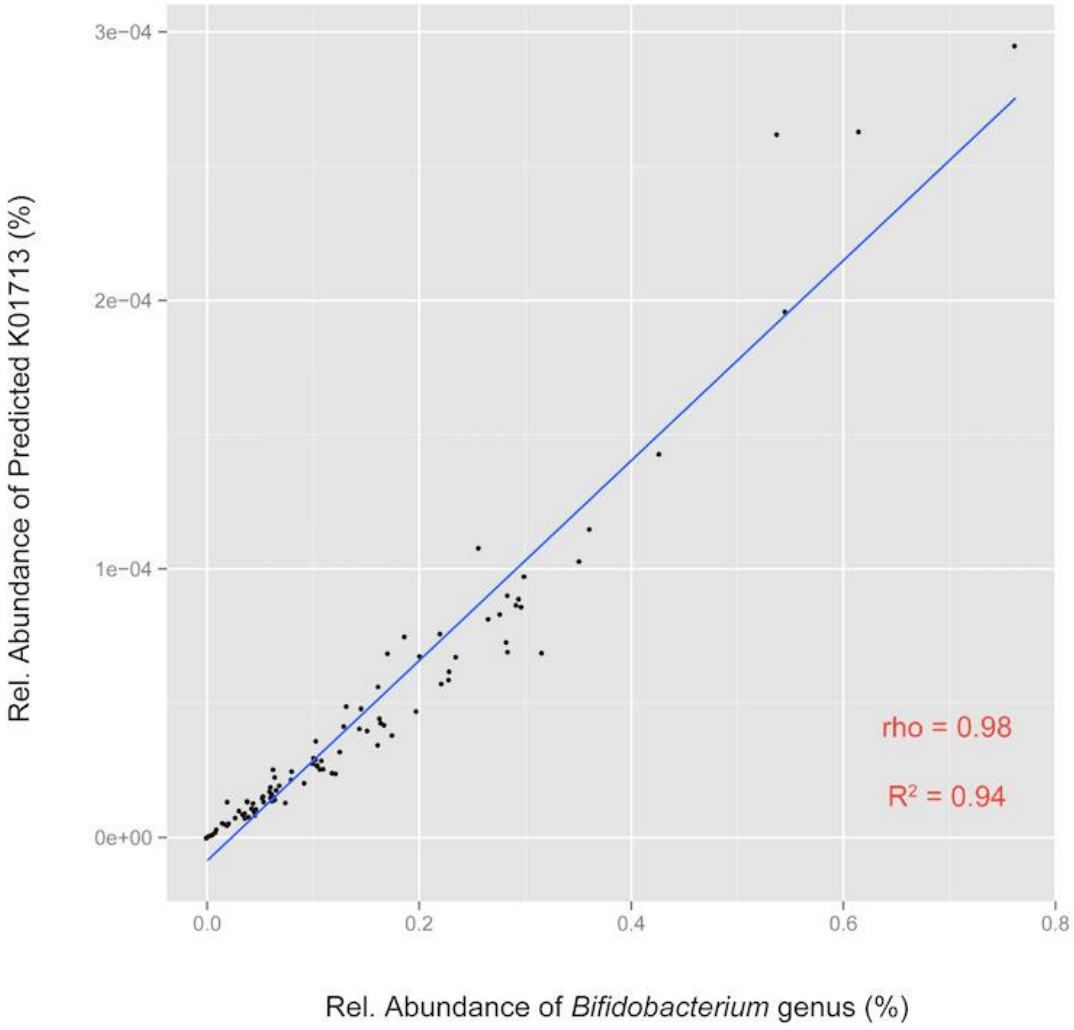

Figure. Spearman correlation of the relative abundance levels of predicted cyclohexadienyl dehydratase (CDT; KEGG Ortholog K01713; EC:4.2.1.9151) to the relative levels of *Bifidobacterium* genus according to 454 pyrosequencing, calculated over all 96 ADHD and control samples in the cohort ($p < 0.001$, Bonferroni corrected).

Data S2: fMRI main effects of reward anticipation

***Table*.** MNI stereotactic coordinates of local BOLD maxima (max. 3) in clusters significant for the main effect of reward anticipation (15 > 1 cent cues) at *p*FWE < 0.05 (cluster level).

		cluster statistics		local maxima	
region	**side**	**p(FWE)**	**size**	**T-value**	**x, y, z**
n = 87					
superior occipital gyrus	R	< 0.001	39442	10.72	15, -91, 3
calcarine sulcus	L			10.27	-12, -93, -0
precentral gyrus	L	< 0.001	30729	7.83	-26, -9, 46
supplementary motor area	L			7.58	-6, -1, 49
supplementary motor area	R			7.19	9, 6, 46
putamen	R	< 0.001	8824	7.34	18, 12, -0
caudate	L			6.88	-9, 11, -3
middle frontal gyrus	L	< 0.001	1802	6.25	-38, 44, 19
middle frontal gyrus	R	< 0.001	735	5.28	36, 44, 25
superior temporal gyrus	R	0.016	320	4.18	57, 8, -3
n = 28					
superior occipital gyrus	R	< 0.001	3658	7.53	20, -93, 7
inferior occipital gyrus	R			6.66	33, -81, -6
lingual gyrus	R			6.37	20, -94, -6
calcarine sulcus	L	< 0.001	4474	6.84	-12, -88, -0
middle occipital gyrus	L			6.71	-27, -93, 3
inferior occipital gyrus	L			6.06	-39, -85, -8
precentral gyrus	L	< 0.001	9571	6.71	-26, -7, 48
inferior parietal lobe	L			6.16	-45, -31, 36
putamen	L	< 0.001	1027	6.65	-12, 9, -9
caudate	L			5.46	-8, 0, -2
precentral gyrus	R	< 0.001	2063	6.02	35, -16, 69
pallidum	R	< 0.001	1013	5.68	18, 15, 1
putamen	R			4.73	20, 20, -6
cerebellum	R	0.004	385	5.65	14, -66, -21
superior occipital gyrus	L	< 0.001	1166	5.62	-24, -70, 30
precuneus	L			4.74	-11, -72, 48
superior parietal lobe	L			4.57	-29, -64, 49
cerebellum	L	0.007	352	5.38	-27, -51, -29
fusiform gyrus	L			4.25	-35, -52, -20
cerebellum	R	< 0.001	941	5.31	36, -48, -30
middle cingulate gyrus	L	0.002	443	5.01	-3, -27, 30
middle cingulate gyrus	R			4.53	5, -24, 34
posterior cingulate gyrus	L			4.36	-3, -39, 25
cerebellum	R	< 0.001	671	4.96	18, -68, -50
superior parietal lobe	R	0.001	465	4.92	23, -49, 52
middle frontal gyrus	L	0.004	388	4.89	-32, 35, 36
middle frontal gyrus	R	0.049	231	4.61	33, 54, 4
superior orbitofrontal gyrus	R			4.47	24, 54, -2