A prospective investigation into the association between the gut microbiome composition and cognitive performance among healthy young adults

Kolade Oluwagbemigun1*, Maike E. Schnermann1, Matthias Schmid2, John F. Cryan3 and Ute Nöthlings1

Abstract

Background: There is emerging evidence that the gut microbiome composition is associated with several human health outcomes, which include cognitive performance. However, only a few prospective epidemiological studies exist and none among young adults. Here we address the gap in the literature by investigating whether the gut microbiome composition is prospectively linked to fluid intelligence among healthy young adults.

Methods: Forty individuals (65% females, 26 years) from the DOrtmund Nutritional and Anthropometric Longitudinally Designed (DONALD) study provided a fecal sample for gut microbiome composition and subsequently (average of 166 days) completed a cognitive functioning test using the Cattell’s Culture Fair Intelligence Test, revised German version (CFT 20-R). The assessment of the gut microbiome at the genera level was by 16S rRNA V3-V4 Illumina sequencing. The relative abundance of 158 genera was summarized into bacterial communities using a novel data-driven dimension reduction, amalgamation. The fluid intelligence score was regressed on the relative abundance of the bacterial communities and adjusted for selected covariates.

Results: The 158 genera were amalgamated into 12 amalgams (bacterial communities), which were composed of 18, 6, 10, 14, 8, 10, 16, 13, 12, 12, 3, and 11 genera. Only the 14-genera bacterial community, named the “Ruminococcaceae- and Coriobacteriaceae-dominant community” was positively associated with fluid intelligence score (β = 7.8; 95% CI: 0.62, 15.65, P = 0.04).

Conclusion: Among healthy young adults, the abundance of a gut bacterial community was associated with fluid intelligence score. This study suggests that cognitive performance may potentially benefit from gut microbiome-based intervention.

Keywords: Gut microbiome, Relative abundance, Ruminococcaceae, Coriobacteriaceae, Cognitive performance, Fluid intelligence, DONALD Study

Background

It has become increasingly recognized that the gut microbiome may play a substantial role in the occurrence of many human conditions [1, 2]. More recently, the association of the gut microbiome with cognitive neurodevelopment and brain functioning has attracted much attention [3–7]. This association is attributed to the...
microbiota-gut-brain-axis [8, 9]. Indeed, epidemiological investigations have demonstrated associations between the composition of the gut microbiome and prevalent neurodegenerative conditions such as Alzheimer’s disease [10–13] and Parkinson’s disease [12, 14].

In apparently healthy individuals, the association between the gut microbiome composition and cognitive functioning is also documented [15–24]. However, these studies have yielded inconsistent results. These inconsistent results may reflect limitations of the design of these studies, which is mainly cross-sectional [15–17, 19–22, 24]. Therefore, more prospective studies where the gut microbiome composition is profiled before the assessment of cognitive performance will help to draw a more definite conclusion and contribute to improving our knowledge on the influence of the gut microbiome on cognitive functioning. Additionally, previous studies have been among infants [18, 23], children [24], middle-aged adults [20], and older adults [15–17, 19–22]. Considering that the cognitive health in young adulthood positively correlates with memory and brain functioning in later life [25], studies among young adults would be necessary.

Crystallized and fluid cognition (or intelligence) are the two cognitive domains [26]. The fluid intelligence is crucial because it is a person’s innate ability to process and manipulate one’s environment [27]. In fact, it is positively associated with better psychological and health outcomes throughout adulthood and into old age [28]. Thus, the relationship between the gut microbiome and fluid intelligence should be of research interest.

To this end, the present prospective epidemiological study sought to investigate whether the gut microbiome composition is independently associated with fluid intelligence among young adults.

Results
Description of study population
Table 1 presents the basic characteristics of the 40 individuals in the current analysis. The median age at the time of fecal sampling was 26 years. About two-third (65%) were females. The median time between fecal sampling and the assessment of cognitive performance was 166 days. Furthermore, the birth weight was approximately 3.5 kg. Their BMI of 23.4 kg/m² is within the normal weight range. They had a moderate physical activity of approximately 33 MET-hour/week. The median carbohydrate intake and alcohol consumption were 199 g/day and 0.39 g/day, respectively. The median Shannon alpha diversity of the gut microbiome was 6.1. Finally, the median fluid intelligence score was 110.

Table 2 shows the 133 assigned genera and the 12 mutually exclusive amalgams to which they were assigned. These 12 amalgams were composed of 18, 6, 10, 14, 8, 10, 16, 13, 12, 12, 3, and 11 genera, respectively. They include well-known genera such as *Fusobacterium*, *Lachnoclostridium*, *Staphylococcus*, *Akkrermansia*, *Bacteroides*, *Streptococcus*, *Phascolarcobacterium* and *Desulfovibrio*, *Enterococcus*, *Paraprevotella*, *Blautia*, *Pseudobutyrivibrio*, and *Dialister*, respectively. Twenty-five genera were not assigned.

The boxplots for the average center log-ratio transformed relative abundance (RA) of all genera in each amalgam (Fig. 1) shows that the lowly abundant genera were grouped together in one amalgam. The inter-quartile range of ten amalgams being lower than the unassigned group also suggests that lowly abundant genera were not systematically excluded by amalgamation. Further, the fact that the 12 amalgams and the unassigned group overlapped with at least one other group suggests that none of the groups was different from all others. The 12 amalgams were named as follows: “Lachnospiraceae-dominated community I”, “Lachnospiraceae- and Ruminococcaceae-dominated community”, “Lachnospiraceae- and Christensenellaceae-dominated community”, “Ruminococcaceae- and Coriobacteriaceae-dominated community”, “Erysipelotrichaceae-dominated community I”, “Lachnospiraceae-dominated community II”, “Erysipelotrichaceae-dominated community II”, “Erysipelotrichaceae-dominated community III”, “Erysipelotrichaceae-dominated community IV”, “Erysipelotrichaceae-dominated community V”, “Erysipelotrichaceae-dominated community VI”, “Erysipelotrichaceae-dominated community VII”, “Erysipelotrichaceae-dominated community VIII”, “Erysipelotrichaceae-dominated community IX”, “Erysipelotrichaceae-dominated community X”, “Erysipelotrichaceae-dominated community XI”, “Erysipelotrichaceae-dominated community XII”, “Erysipelotrichaceae-dominated community XIII”, “Erysipelotrichaceae-dominated community XIV”, “Erysipelotrichaceae-dominated community XV”, and “Erysipelotrichaceae-dominated community XVI”.

| Table 1 | Basic characteristic of the study population (N = 40) |
|---------|-----------------------------------------------------|
| N       |                                                     |
| Sex, females | 40 26 (65)                                      |
| Age, years | 40 26 (22, 30)                                    |
| Birth weight | 40 3480 (3215, 3665)                              |
| Body mass index, kg/m² | 40 23.43 (21.08, 25.06) |
| Physical activity, MET-hour/week | 40 33.14 (21.72, 51.12) |
| Energy intake, kcal/day | 40 1551.3 (1418.96, 1746.72) |
| Carbohydrate intake, g/day | 40 198.62 (173.97, 230.34) |
| Fiber intake, g/day | 40 14.69 (13.10, 16.63) |
| Protein intake, g/day | 40 51.02 (46.87, 56.87) |
| Fat intake, g/day | 40 61.17 (55.61, 67.5) |
| Alcohol consumption, g/day | 40 0.39 (0.21, 1.1) |
| Smoking status, current smokers | 40 37 6 (16) |
| Education, ≥ 12 years of education | 40 22 (55) |
| Antibiotics intake | 40 13 (39) |
| Probiotics intake | 40 17 (52) |
| Time between fecal sampling microbiome and cognition assessment, days | 40 166 (130, 194.5) |
| Shannon alpha diversity index | 40 6.1 (5.86, 6.3) |
| Fluid intelligence score | 40 110 (100, 119) |

a n(%)  
b Median (25th, 75th percentile)  
c N = count, % = percentage
## Table 2  The twelve amalgams of the gut microbiome and their description

| Amalgams | Numbers of genera | Genera | Description |
|----------|-------------------|--------|-------------|
| V1       | 18                | Tyzzerella, Lachnospiraceae FCS020 group, Howardella, Prevotellaceae NK3B31 group, Alloprevotella, Coprobacillus, Intestinibacter, NB1-1 uncultured bacterium, Family XIII AD3011 group, Prevotella 7, Uncultured Thermoaerobacterales bacterium, Haemophilus, Eubacterium, Bacteroides S24-7 group uncultured organism, Lactobacillus, Fusobacterium, Veillonellaceae uncultured bacterium, Peptococcus | Lachnospiraceae-dominant community I |
| V2       | 6                 | Lachnospiraceae FE2018 group, Lachnoclostridium, Ruminococcaceae UCG-003, Erysipelotrichaceae UCG-004, Defluviitaleaceae UCG-011 | Lachnospiraceae- and Ruminococcaceae-dominant community |
| V3       | 10                | Lachnospiraceae NC2004 group, Tyzzerella 4, R-7 group, Christensenellaceae uncultured bacterium, Clostridium sensu stricto 1, Veillonella, Anaerotruncus, Staphylococcus, Moraxellaceae uncultured organism, Prevotella 9, Eubacterium, Eubacterium hallii, Comamonas, Leucobacter, Lactobacillus, Fusobacterium, Veillonellaceae NK3B31 group, Alloprevotella, 
|           |                   |        | Lachnospiraceae- and Christensenellaceae-dominant community |
| V4       | 14                | Ruminiclostridium 5, Ruminococcaceae UCG-010, Coriobacteriaceae uncultured bacterium, Slackia, [Eubacterium] hallii group, Peptoclostridium, Akkermansia, Lactococcus, Erysipelotrichaceae Incertae Sedis, [Eubacterium] nodatum group, Prevotellaceae uncultured bacterium, Robiginitalea, Peudomonas, Bacteroides S24-7 group uncultured bacterium, Porphyromonadaceae-dominant community | Ruminococcaceae- and Coriobacteriaceae-dominant community |
| V5       | 8                 | Erysipelotrichaceae bacterium 21_3, Erysipelotrichaceae UCG-004, Bacteroides, Anaerostipes, Clostridiales vadinBB60 group uncultured organism, Coprobacter, Megaphaera, boneC3G7 uncultured bacterium | Erysipelotrichaceae-dominant community I |
| V6       | 10                | Roseburia, Lachnoclostridium 5, Hungatella, Streptococcus, Terrisporobacter, Oscillospira, Uncultured Mollicutes bacterium, Acidaminococcus, Haemophilus, Acinetobacter | Lachnospiraceae-dominant community II |
| V7       | 16                | Tunicibacter, Holdemanella, Catenibacterium, Ruminococcaceae UCG-002, Flavonifractor, Allisonella, Megamonas, Gastranerophilale uncultured bacterium, Escherichia-Shigella, Eggerthella, Phascolarctobacterium, Clostridiales vadinBB60 group uncultured bacterium, Lachnospiraceae uncultured bacterium, Prevotellaceae UCG-003, Desulfovibrio, Anaeroplasma | Erysipelotrichaceae-dominant community II |
| V8       | 13                | Fuscatenibacter, Tyzzerella 3, Eisenbergiella, Ruminococcaceae UCG-013, Butyrinicoccus, Parasutterella, Sutterella, Erysipelotrichaceae Clostridium, Solobacterium, Prevotellaceae UCG-001, Enterococcus, Thalassospira, Enterorhabdus | Lachnospiraceae-dominant community III |
| V9       | 12                | Intestimonas, Ruminiclostridium 6, Ruminiclostridium 1, Lachnospiraceae NK4A136 group, Lachnospiraceae UCG-001, Butyrimonas, Rikenellaceae RC9 gut group, Selenellimassilia, Raoultella, NB1-1 uncultured organism, Mitsuokella, Paraprevotella | Ruminococcaceae-dominant community I |
| V10      | 12                | Faecalibacterium, Ruminococcaceae UCG-004, Ruminiclostridium, Ruminococcaceae UCG-009, Ruminiclostridium 9, Buturia, [Eubacterium] ventriosum group, Bilaphila, Erysipelotrichaceae UCG-003, Defluviitaleaceae UCG-011, Prevotella 9, Collinsella | Ruminococcaceae-dominant community II |
| V11      | 3                 | Pseudobutyrivibrio, [Ruminococcus] gauveaeau group, [Eubacterium] ruminantium group | Lachnospiraceae community |
| V12      | 11                | Odontobacter, Porphyromonadaceae uncultured, Barnesiella, Subdoligranulum, Ruminococcaceae UCG-003, Asteroleplasma, Methanobrevibacter, Opitecota vadinHA64 uncultured bacterium, Lachnospiraceae UCG-002, Dialister, Akctisps | Porphyromonadaceae-dominant community |
| V0       | 25                | Lachnospiraceae UCG-004, Lachnospiraceae UCG-005, Lachnospiraceae UCG-008, Dorea, Coprococcus 1, Coprococcus 2, Lachnospiraceae, uncultured, Lachnospira, Marvinbyantia, Shuttleworthia, Ruminococcaceae NK4A214 group, Ruminococcaceae UCG-005, Ruminococcaceae UCG-014, [Eubacterium] coprostanoligenes group, Ruminococcaceae uncultured bacterium, Ruminococcus 1, Rhodococcus, Bifidobacterium, Parabacteroides, Faecalibacterium, Victivallis, Cloacibacillus, Enterobacter, Prevotella 2, Succinicrobium | Group of unassigned genera |
“Lachnospiraceae-dominant community III”, “Ruminococcaceae-dominant community I”, “Ruminococcaceae-dominant community II”, “Lachnospiraceae community”, and “Porphyromonadaceae-dominant community”. Further, among the bacterial communities, there were no exceptionally high pairwise correlations with the maximum correlation being moderate ($|0.5| \leq r \leq |0.7|$, $P<0.05$) and appearing only in eight correlations. The Shapiro–Wilk W test for normality ($P=0.51$) on fluid intelligence score indicated that it has a normal distribution with constant variance. Thus, fluid intelligence score was modelled untransformed.

In the adaptive LASSO, the value of the lambda that gives the minimum mean cross-validated error was 84.48. There were two predictors of fluid intelligence score with non-zero coefficients at this value. Figure 2 displays these two predictors, the “Ruminococcaceae- and Coriobacteriaceae-dominant community” and sex, with their regularized regression coefficients of 3.86 and $-5.38$, respectively.

Additionally, the RF-RFE recommends that seven predictors of fluid intelligence score were optimal when the root-mean-square error, R-squared, and the mean absolute error reached their maximum levels at 13.18, 0.07, and 10.76, respectively. These seven predictors were “Lachnospiraceae-dominant community I”, time between fecal sampling and cognitive measurement, alcohol consumption, sex, carbohydrate intake, “Erysipelotrichaceae-dominant community II”, and “Ruminococcaceae- and Coriobacteriaceae-dominant community” with importance scores of 3.87, 2.7, 2.55, 2.18, 2.16, 1.63, and 1.48, respectively (Fig. 3). Thus, the true relevant predictors of fluid intelligence score predictors, shared by the adaptive
LASSO and RF-RFE, were Ruminococcaceae- and Coriobacteriaceae-dominant community and sex.

Regressing fluid intelligence score on “Ruminococcaceae- and Coriobacteriaceae-dominant community” and sex using the ordinary least squares regression showed that “Ruminococcaceae- and Coriobacteriaceae-dominant community” was positively associated with fluid intelligence score ($\beta = 7.8; 95\% CI: 0.62, 15.65, P=0.04$) and females had lower intelligence score when compared to males ($\beta = -9; 95\% CI: -17.36, -0.71, P=0.03$). The model F statistic was significant ($F=4.05, P=0.03$), indicating that the model accounts for a significant portion of the variation in the data. The adjusted R-squared indicates that these two predictors explained approximately 14% of the variation in fluid intelligence score.

**Discussion**

The current prospective epidemiological study among young adults investigated whether the gut microbiome composition was independently associated with fluid intelligence. Among the 12 bacterial communities retrieved from this study population, the “Ruminococcaceae- and Coriobacteriaceae-dominant community” was positive and independently associated with fluid intelligence. This 14-genera community comprises *Ruminiclostridium* sp, Ruminococcaceae UCG-010, Coriobacteriaceae uncultured bacterium, *Slackia*, *Eubacterium* hallii group, *Peptoclostridium*, *Akkermansia*, *Lactococcus*, *Erysipelotrichaceae incertae sedis*, *Eubacterium* nodatum group, *Prevotellaceae uncultured bacterium*, *Robiginitalea*, *Pseudomonas*, and, *Bacteroidales S24-7 group uncultured bacterium*.

Our finding echoes those of previous epidemiological investigations, which reported the relationship of Ruminococcaceae and Coriobacteriaceae with cognitive functioning. The abundance of Ruminococcaceae was positively associated with good cognition [15, 29]. A study with a probiotic supplementation showed that an increase in the abundance of a genus in Ruminococcaceae resulted in better protection against the negative effects of stress on working memory [30]. Furthermore, it was reported that the abundance of Coriobacteriaceae was positively associated with a better cognitive performance [20]. The abundance of Ruminococcaceae was reduced in Alzheimer disease [29], multiple sclerosis [31] and schizophrenia [32], and the abundance of many of its species was reduced in Parkinson’s disease [14]. Some of the genera in this community are also associated with cognition. The abundance of *Erysipelotrichaceae incertae sedis* and *Eubacterium* hallii group were positively associated with cognitive functioning [15] and social cognition [33], respectively. Besides, two intervention studies reported that the ingestion of multispecies probiotics, which includes *Lactococcus*, was associated with reduced cognitive...
reactivity to mood disorders [34, 35]. Diet-induced increase in the abundance of Akkermansia and Slackia was associated with improved Alzheimer’s disease biomarkers in individuals with mild cognitive impairment [36]. There was also a reduced abundance of Slackia in multiple sclerosis [37]. The fact that the majority of these aforementioned results are in a similar direction as ours suggests that our findings are reliable and biologically relevant.

The first potential biological mechanism by which the gut microbiome influence cognitive functioning is through its stimulation of the afferent neurons of the enteric nervous system that communicates with the central nervous system via the vagus nerve [9]. In addition, the gut microbiome possesses the ability to produce and modify various immune, metabolic, and neuroactive factors that affect the central nervous system [9]. Important neuroactive factors are produced from the gut microbiome’s modulation of the dietary protein and carbohydrate metabolism [9]. First, Lactococcus [38] and Pseudomonas [39] are able to modulate the serotonin signaling/metabolism. Akkermansia was predicted to be able to produce serotonin [40]. Further, Pseudomonas is one of the gamma aminobutyric acid-modifying genera [41] and gamma aminobutyric acid level is lower in individuals with Alzheimer’s disease as compared to healthy individuals, [42]. Lactococcus produces dopamine [43] and histamine [44, 45] that regulate cognitive functions. Akkermansia and [Eubacterium] haliit group produce short chain fatty acids (SCFA) from carbohydrate metabolism. Akkermansia produces acetate and propionate [46], while [Eubacterium] haliit group produces propionate [2] and butyrate [10]. The production of these SCFA generally has a beneficial influence on many neurodegenerative conditions [10]. Indeed, systemic acetate has the capability to cross the blood–brain barrier where it can activate acetyl-CoA carboxylase leading to the enhancement of the expression of neuropeptides that induces hypothalamic neuronal activation and suppresses appetite [47]. Akkermansia also tend to produce indole and indole acetic acid from tryptophan metabolism [39]. The indirect mechanism through inflammation implicates Akkermansia and Slackia [48–50]. Akkermansia plays a critical role in maintaining the integrity of the mucin layer and reducing inflammation [48]. Slackia is an equal producer [49] and equal is crucial in maintaining immune homeostasis because it induces anti-inflammatory response [50]. The SCFA also act as anti-inflammatory mediators [51]. Clearly, these mechanisms may critically interact with one another in complex ways. Considering that only a limited number of genera in this community have documented potential mechanism of action, further work is needed for insights into how the bacteria in this community work together to impact cognition.

The two prospective studies, which are both among infants, reported that Firmicutes-dominant and Bacteroidetes-dominant clusters [23], Bacteroides-dominant cluster [18], and Bacteroides [23] were positively associated with a favorable cognitive function. The Bacteroides-dominant cluster of Carlson et al. [18] and our “Erysipelotrichaceae-dominant community I” both have Bacteroides. Thus, they are comparable. The fact that the “Erysipelotrichaceae-dominant community I” and “Lachnospiraceae-dominant community I” with the highest and lowest variance respectively were not associated with fluid intelligence suggests that the absence of association of “Erysipelotrichaceae-dominant community I” with fluid intelligence is unlikely to be influenced by its variation. Therefore, our finding for the “Ruminococcaceae- and Coriobacteriaceae-dominant community” suggests that the relationship between the gut microbiome composition and cognitive function may be different between infants and young adults. Furthermore, between three to five genera in the “Ruminococcaceae- and Coriobacteriaceae-dominant community” consistently cluster together across all the tested number of amalgams. This suggests that the membership of this community is not arbitrary but highly reproducible and the community might indeed represent a relevant biomarker. Furthermore, the bacterial communities retrieved in this study are consistent with other studies using different dimension reduction methods. A study among older German adults was also able to recover a Ruminococcaceae-dominant bacterial community and a community including the Coriobacteriaceae [52]. The [Eubacterium] haliit group and Peptoclostridium in our “Ruminococcaceae- and Coriobacteriaceae-dominant community” were also among the genera in one of the bacterial communities of Leong et al. [53]. Besides, our “Erysipelotrichaceae-dominant community I” and “Ruminococcaceae-dominant community II” are somewhat comparable to the Bacteroides- and Faecalibacterium-dominant clusters of Carlson et al. [18].

A recent study in a different population showed that age, sex, education, average food intake, and tyrosine intake explained 6% of the variance of fluid intelligence [54]. This is lower than the 14% variance of fluid intelligence explained by the abundance of the “Ruminococcaceae- and Coriobacteriaceae-dominant community” and sex in our study. This finding underscores the importance of the gut microbiome composition in fluid intelligence. Sex disparity in different aspects of cognitive performance is well documented and much debated [55–58]. In consort with our findings, studies among young adults that assessed fluid intelligence by the CFT
inversely related to fluid intelligence score. This lends credence to the growing evidence that the gut microbiome may influence cognitive performance. Taken together, our study suggests that cognitive performance may potentially benefit from gut microbiome-based intervention and this group of bacteria may have a promising health-promoting role.
Methods

Study population
The DOrtmund Nutritional and Anthropometric Longitudinally Designed (DONALD) study is an ongoing, open prospective epidemiologic cohort study of individuals living in the German town of Dortmund and surrounding cities that commenced in 1985. The study was designed to investigate the relationships among dietary intake, metabolism, and growth from infancy into adulthood. Participants’ examinations included annually repeated anthropometric measurements and three-day weighed dietary records. Early-life factors such as birth weight of study participants were extracted from maternal delivery documents. The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the University of Bonn. Informed consent was obtained from the parents or legal guardians of the participants in childhood and later on from the participants themselves. Details of the recruitment and follow-up in the DONALD study are presented elsewhere [61].

Study design
This is a prospective study of adult (age ≥ 18 years) DONALD study participants who provided fecal samples for gut microbiome compositional analysis [59] and who subsequently attended cognitive testing few months after fecal sampling. These individuals were singletons, full term (36–42 weeks) and birth weight of ≥ 2500 g. There were 40 individuals in total. For these 40 individuals, we retrieved their gut microbiome RA data, cognition data, and other covariates.

Assessment of gut microbiome composition
Details of the fecal sampling, DNA extraction, 16S ribosomal RNA sequencing have been published [59]. Briefly, fecal samples were collected between 2017 and 2018 at the participants’ home into tubes containing RNAlater (Qiagen) and sent to the Biobank within 24 h of collection. Bacterial genomic DNA was extracted from 0.25 g of fecal sample using the repeat bead beating plus column protocol as in combination with the QIAamp Fast DNA Stool Mini Kit. For the 16S ribosomal RNA sequencing, the V3-V4 regions of the 16S rRNA gene were amplified through 30 cycles of PCR reactions according to the 16S Metagenomic Sequencing preparation protocol for Illumina MiSeq. The Quantitative Insights into Microbial Ecology (QIIME version 1.8.0) was used for quality filtering of pair-end reads, which is based on a quality score of > 25 and the removal of mismatching barcodes. Prior to the analysis as part of the quality control process, sequences which produced less than 40,000 reads were manually removed. USEARCH (version 7, 64-bit) was employed for deionization, chimera detection and clustering into operational taxonomic units (OTUs; 97% identity). Alignment of OTUs was carried out using the python nearest alignment space termination (PyNAST) and assignment of taxonomy of 97% similar identity against the SILVA SSURef database release v123. The R package Phyloseq was used to determine 341 genera and their relative abundances. The 158 abundant genera with RA ≥ 0.2% in at least 10% of the samples [59] were considered for this present study.

Assessment of cognitive performance, fluid intelligence
The cognitive performance was assessed between 2017 and 2018 using the Cattell’s Culture Fair Intelligence Test, revised German version (CFT 20-R) [62]. This computer-based CFT 20-R was done in the participants’ homes. The test comprises two parts, covering four domains: series, classifications, matrices, and topologies. The first and second parts comprised 56 and 45 figure-based questions and are to be complete within 14 min (4-4-3-3 min) and 12 min (3-3-3-3 min), respectively. The sum score of the domains was calculated from the correct answers, separately for each part. Age-standardized intelligence quotient (IQ) scores were derived from these sum scores to form IQ1 for the first part, IQ2 for the second part, and overall IQ–average of IQ1 and IQ2. If there was a difference ≥ 12 points between IQ1 and IQ2, the higher IQ was chosen as fluid intelligence score, otherwise the overall IQ was chosen as the fluid intelligence score.

Assessment of other covariates
Demographic information such as sex and birth weight were retrieved from maternal delivery records. Age at fecal sampling was calculated from the documented date of birth and the date of fecal sampling. Dietary intake was assessed annually using three-day weighed dietary records on three consecutive days. From all dietary records prior to fecal sampling (age ≤ 18 years), means of daily energy (kcal/day) and carbohydrate, fiber, protein, and fat (g/day) intakes were calculated. Anthropometric measurements were conducted annually at the study center. For the current study, we considered the weight and height measurement closest to fecal sampling to calculate the body mass index (BMI). Physical activity was self-reported in a validated questionnaire covering the frequency and duration of individuals’ participation in the home and leisure physical activities during the week and at weekends. This was converted to metabolic equivalent of task (MET)-hours/week and the average over the available records of age ≤ 18 years was calculated for each individual. Educational status and lifestyle factors of the participants such as alcohol consumption and smoking,
and self-reported intake of antibiotics and probiotics six months prior to fecal sampling, were obtained by questionnaires.

**Statistical analyses**

Continuous and categorical variables were summarized as median (25% and 75% percentile), and as count (percentage), respectively.

Microbiome datasets are typically compositional, high dimensional, and zero-inflated. Hence, modeling its individual variables (taxa), particularly as predictors in conventional regression models is suboptimal [63]. The conventional and appealing statistical methods that efficiently model microbiome data include the simple log-ratio, log-ratio between two geometric means (balances), and summed log-ratio with the summed log-ratio being arguably the most interpretable [64]. Recently, a data-driven summed log-ratio method called amalgamation for reducing the dimensionality of compositional data has been proposed. [64]. Amalgamation outperforms traditional dimension reduction approaches, especially in terms of interpretability [64]. The resulting components (amalgams) from amalgamation can be described as bacterial communities [64].

For this current analysis, we considered the previously reported 158 abundant genera in this study population [59]. These include the core genera, *Bacteroides*, *Lachnoclostridium*, and *Blautia*, and the diet-related genera, *Phascolarctobacterium*, *Dialister*, and *Desulfovibrio* [59]. The zero counts in the abundance of the genera were replaced by the Bayesian-multiplicative method. Afterwards, we applied the amalgamation method. We started with three amalgams, using the simplest amalgamation where each genus only contributes to one amalgam and an objective function that preserves the Aitchison distances between samples. The number of amalgams was increased by one until we achieved amalgams where *Bacteroides*, *Lachnoclostridium*, *Blautia*, *Phascolarctobacterium*, *Dialister*, and *Desulfovibrio* were loaded. This was achieved at 12 amalgams. The amalgams (bacterial communities) were named according to the dominating bacterial family. As recommended by Quinn et al. [64], we center log-ratio-transformed the resultantly assigned amalgams for further analysis.

A minimum of approximately two observations per variable is adequate for a valid linear regression model [65]. Therefore, for our 40 observations, we will require a maximum of 20 predictors. This implies that our model should comprise the 12 bacterial communities and a maximum of eight covariates. Additionally, we estimated the power of our multivariable linear regression with 40 observations (n = 40), 20 predictors, a significance level of 0.05, and 46% average variation in cognitive performance captured by models comprising age, sex, and RA of the gut microbiome reported by another study [20]. With these values, the power of our multivariable linear regression model would be 72%, which is close to the conventional threshold of 80% [66]. We selected covariates, such as age [67, 68], sex [68, 69], BMI [70, 71], carbohydrate intake [59, 72], and alcohol consumption [73, 74] for which the direction of association with gut microbiome composition and cognition is well substantiated in the literature. We used these covariates to draw directed acyclic graphs in order to determine the minimal sufficient adjustment sets for estimating the direct effect of the gut microbiome RA on cognition. From these covariates, we selected eight covariates in the study population that were measured and have no missing. These were sex (reference: female), age at fecal sampling, BMI at fecal sampling, average physical activity, Shannon alpha diversity index, average carbohydrate intake, average alcohol consumption, and time between fecal sampling and cognitive measurement. In addition, we assessed whether the response variable, fluid intelligence score has a normal distribution with constant variance using the Shapiro–Wilk W test for normality.

Furthermore, using two variable selection regression methods, the adaptive Least Absolute Shrinkage and Selection Operator (LASSO) regression and the random forest regression with recursive feature elimination (RF-RFE), we identified true relevant predictors of fluid intelligence score by regressing it on the center log-ratio-transformed amalgams and the selected covariates. For both methods, we used five-fold cross-validation. The coefficient-specific penalty level of the LASSO was the inverse of the absolute values of the best ridge coefficients. We extracted the predictors with non-zero coefficients at the value of the lambda that gives the minimum mean cross-validated error. In the RF-RFE, we selected the optimal set of predictors when the root-mean-square error, R-squared, and the mean absolute error reached the maximum level. The true relevant predictors were the predictors shared by adaptive LASSO and RF-RFE. Finally, in order to obtain unbiased estimates for the true relevant predictors, we regressed fluid intelligence score on them using ordinary least squares regression using a five-fold cross-validation.

All statistical analyses were performed using the statistical software R version 4.1.1. The replacement of zeros, amalgamation, the adaptive LASSO, and the RF-RFE were implemented using the cmultRepl function from the zCompositions package, the amalgam package, the glmnet package, and the caret package, respectively.
Abbreviations
BMI: Body mass index; CFT 20-R: Cattell’s Culture Fair Intelligence Test, revised German version; DONALD: DOrtmund Nutritional and Anthropometric Longitudinally Designed; IQ: Intelligence quotient; LASSO: Least Absolute Shrinkage and Selection Operator; MET: Metabolic equivalent of task; RF-RFE: Random forest regression with recursive feature elimination; RA: Relative abundance; SCFA: Short chain fatty acids.

Acknowledgements
We would like to thank the DONALD Study participants and their families for their interest, time, and commitment to the study. We would like to thank the DONALD Study center staffs for collecting and managing the data. We appreciate the support of Petra Pickert in fecal sample collection. We also thank Stefan Benda for formatting the figures in this manuscript.

Authors’ contributions
KO: study concept and design, data analysis and interpretation, and drafting of the manuscript; MES: reviewed and updated of the cognition data; MS: provided statistical advice; JFC: contributed to the interpretation of the results; UN: secured the funding and provided supervision; All authors: contributed to the critical revision of the manuscript. All authors read and approved the final manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL. The German Federal Ministry of Education and Research [grant number, 01EA1906D] supported this work. The funder had no role in the design of this study, the collection and analysis of data, the preparation of the manuscript, and the decision to publish this work.

Availability of data and materials
The datasets generated and/or analyzed during the current study are not publicly available due to ethical and privacy issues, but they are available from the corresponding author on reasonable request and the approval of the ethics committee. Further information and the request for data sharing are available from the corresponding author or from the ethics committee.

Declarations
Ethics approval and consent to participate
The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the University of Bonn. Informed consent was obtained from the parents or legal guardians of the participants in childhood and later on from the participants themselves.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1 Nutritional Epidemiology, Department of Nutrition and Food Sciences, University of Bonn, Bonn, Germany. 2Department of Medical Biometry, Informatics and Epidemiology, University Hospital Bonn, University of Bonn, Bonn, Germany. 3APC Microbiome Ireland, University College Cork, Cork, Ireland.

Received: 13 December 2021 Accepted: 17 March 2022
Published online: 19 April 2022

References
1. Wang B, Yao M, Lv L, Zeng Z, Li L. The human microbiota in health and disease. Engineering. 2017;3(1):71–82.
2. Durack J, Lynch SV. The gut microbiome: relationships with disease and opportunities for therapy. J Exp Med. 2019;216(1):20–40. https://doi.org/10.1084/jem.20180448.
3. Dinan TG, Cryan JF. Gut instincts: microbiota as a key regulator of brain development, ageing and neurodegeneration. J Physiol. 2017;595(2):489–503.
4. de la Fuente-Nunez C, Meneguetti BT, Franco OL, Lu TK. Neuroniocobiology: how microbes influence the brain. ACS Chem Neurosci. 2018;9(2):141–50.
5. Mohajeri MH, La Fata G, Steenert RE, Weber P. Relationship between the gut microbiome and brain function. Nutr Rev. 2018;76(7):481–96.
6. Warner BB. The contribution of the gut microbiome to neurodevelopment and neuropsychiatric disorders. Pediatr Res. 2019;85(2):216–24.
7. Zhu S, Jiang Y, Xu K, Cui M, Ye W, Zhao G, et al. The progress of gut microbiome research related to brain disorders. J Neuroinflammation. 2020;17(1):25.
8. Cryan JF, O’Riordan KL, Cowan CS, Sundhau KV, Bastaanssen TFS, Boehme M, et al. The microbiota-gut-brain axis. Physiol Rev 2019;99(4):1877–1913.
9. Needham BD, Kaddurah-Daouk R, Mazmanian SK. Gut microbial molecules in behavioural and neurodegenerative conditions. Nat Rev Neurosci. 2020;21(12):717–31.
10. Liu S, Gao J, Zhu M, Liu K, Zhang H-L. Gut microbiota and dysbiosis in Alzheimer’s disease. Implications for pathogenesis and treatment. Mol Neurobiol. 2020;57(12):5026–43.
11. Zhuang Z, Yang R, Wang W, Qi L, Huang T. Associations between gut microbiota and Alzheimer’s disease, major depressive disorder, and schizophrenia. J Neuroinflammation. 2020;17(1):288.
12. Chen Y, Xu J, Chen Y. Regulation of neurotransmitters by the gut microbiota and effects on cognition in neurological disorders. Nutrients. 2021;13(6):2099.
13. Brenchowitz WD, Xiang Y, McEvoy CT, Yang C, Yaffe K, Le W-D, et al. Current Alzheimer disease research highlights: evidence for novel risk factors. Chin Med J (Engl). 2021;134(18):2190–9.
14. Romano S, Savva GM, Bedarf JR, Charles IG, Hildebrand F, Narbad A. Meta-analysis of the Parkinson’s disease gut microbiome suggests alterations linked to intestinal inflammation. npg Parkinson’s Dis. 2021;7(1):27.
15. Bajaj JS, Ahluwalia V, Steinberg JL, Hoogboord S, Boling PA, Godschalk M, et al. Elderly patients have an altered gut-brain axis regardless of the presence of cirrhosis. Sci Rep. 2016;6(1):38481.
16. Anderson JR, Carroll I, Azcarate-Peril MA, Rochette AD, Heinberg LJ, Peat C, et al. A preliminary examination of gut microbiota, sleep, and cognitive flexibility in healthy older adults. Sleep Med. 2017;38:104–7.
17. Manderino L, Carroll I, Azcarate-Peril MA, Rochette A, Heinberg LJ, Peat C, et al. Preliminary evidence for an association between the composition of the gut microbiome and cognitive function in neurologically healthy older adults. J Int Neuropsychol Soc JINS. 2017;23(8):700–5.
18. Carlson AL, Xia K, Azcarate-Peril MA, Goldman BD, Ahn M, Stynar MA, et al. Infant gut microbiome associated with cognitive development. Biol Psychiatr. 2018;83(2):148–59.
19. Verdi S, Jackson MA, Beaumont M, Bower RC, Bell JT, Spector TD, Steves CJ. An investigation into physical frailty as a link between the gut microbiome and cognitive health. Front Aging Neurosci. 2018;10:398.
20. Almo-Buturago ME, Sabater-Maeude M, Moreno-Navarrete JM, Caballano-Infantes E, Armoniaga-Rodriguez M, Coll C, et al. Glutamate interactions with obesity, insulin resistance, cognition and gut microbiota composition. Acta Diabetol. 2019;56(5):569–79.
21. Saji N, Murotani K, Hisada T, Tsuduki T, Sugimoto T, Kimura A, et al. The gut microbiome and cognitive health. Front Aging Neurosci. 2018;10(2):141–50.
22. van Soest APM, Hermes GDA, Berendsen AAM, van de Rest Q, Zoetendal EG, Fuentes S, et al. Associations between pro- and anti-inflammatory gastro-intestinal microbiota, diet, and cognitive functioning in Dutch healthy older adults: The NU-AGE Study. Nutrients. 2020;12(11):3471.
23. Tamahana SK, Tun HM, Konya T, Chari RS, Field CJ, Guttmann DS, et al. Bacteroides-dominant gut microbiome of late infancy is associated with enhanced neurodevelopment. Gut Microbes. 2021;13(1):1–17.
24. Streit F, Prandovszky E, Send T, Zillich L, Frank J, Sabuncyan S, et al. Microbiome profiles are associated with cognitive functioning in 45-month-old children. Brain Behav Immun. 2021;98:151–60.
25. Hicks JF, Flamez BN. Young adulthood: physical and cognitive development across the lifespan: applications for counselors. Wiley; 2016. p. 389–414.
26. Harada CN, Natelson Love MC, Triebel K. Normal cognitive aging. Clin Geriatr Med. 2013;29(4):737–52.

Oluwagbemigun et al. Gut Pathogens (2022) 14:15
49. Schröder C, Matthies A, Engst W, Blaut M, Braune A. Identification and
47. Frost G, Sleeth ML, Sahuri-Arisoylu M, et al. The short-chain fatty acid
44. Landete JM, De Las Rivas B, Marcobal A, Muñoz R. Updated molecular
41. Strandwitz P, Kim KH, Terekhova D, Liu JK, Sharma A, Levering J, et al.
36. Nagpal R, Neth BJ, Wang S, Craft S, Yadav H. Modified Mediterranean-
34. Steenbergen L, Sellaro R, van Hemert S, Bosch JA, Colzato LS. A rand-
32. Zheng P, Zeng B, Liu M, Chen J, Pan J, Han Y, et al. The gut microbiome
31. Tremlett H, Fadrosh DW, Farooq AA, Zhu F, Hart J, Roalstad S, et al. Gut
microbiota in early pediatric multiple sclerosis: a case-control study. Eur J
Neurol. 2016;23(8):1308–21.
30. Bloemendaal M, Szopinska-Tokov J, Belzer C, Boverhoff D, Papalini S,
29. Liu P, Wu L, Peng G, Han Y, Tang R, Ge J, et al. Altered microbiomes distin-
28. Liu Y, Wu L, Peng G, Han Y, Tang R, Ge J, et al. Altered microbiomes distin-
27. Elias L, Saucier D. Neuropsychology: clinical and experimental founda-
tions. Boston: Pearson Education, Inc; 2006.
26. Stavski RS, Almeida DW, Lachman ME, Tun PA, Rosnick CB. Fluid cognitive
ability is associated with greater exposure and smaller reactions to daily
stressors. Psychol Aging. 2010;25(2):330–42.
25. Liu P, Wu L, Peng G, Han Y, Tang R, Ge J, et al. Altered microbiomes distin-
24. Schroder C, Matthies A, Engst W, Blaut M, Braune A. Identification and
expression of genes involved in the conversion of daidzein and genistein
by the equal-forming bacterium Stäckisia isoflavoniconvertens. Appl Envi-
ron Microbiol. 2013;79(11):3494–502.
23. Freedman SN, Shahi SK, Mangalam AK. The “Gut feeling”: breaking down
the role of gut microbiome in multiple sclerosis. Neurotherapeutics.
2018;15(1):109–25.
22. Parada Venegas D, De La Fuente MK, Landskron G, González MJ, Quera R,
Dijkstra G, et al. Short chain fatty acids (SCFAs)-mediated gut epithelial
and immune regulation and its relevance for inflammatory bowel dis-
eases. Front Immunol. 2019;10:277.
21. Oluwagbemigun K, Foonter J, Watkins C, Fohuy F, Stanton C, Bergmann
MM, et al. Dietary patterns are associated with serum metabolite patterns
and their association is influenced by gut bacteria among older German
adults. J Nutr. 2020;150(1):149–58.
20. Leong C, Haszard JJ, Heath A-LM, Tannock GW, Lawley B, Cameron SL,
et al. Using compositional principal component analysis to describe
children’s gut microbiota in relation to diet and body composition. Am J
Clin Nutr. 2018;11(1):88–99.
19. Kühn S, Düzel S, Colzato L, Norman K, Gallinat J, Brandmaier AM,
et al. Food for thought: association between dietary tyrosine and
cognitive performance in younger and older adults. Psychol Res.
2019;83(6):1097–106.
18. Weber D, Skirbekk V, Freund J, Herlitz A. The changing face of cognitive
gender differences in Europe. Proc Natl Acad Sci. 2014;111(32):11673–8.
17. Schlegel K, Wittmer JS, Ramaswamy TH. Intelligence and sensory sensitivity
as predictors of emotion recognition ability. J Intell. 2017;5(4):35.
16. Ahrenfeldt LJ, Scheel-Hincke LL, Kjaergaard S, Møller S, Christensen K,
Lindahl-Jacobsen R. Gender differences in cognitive function and grip
strength: a cross-national comparison of four European regions. Eur J Pub
Health. 2019;29(4):667–74.
15. Flensborg-Madsen T, Falgreen Eriksen H-L, Mortensen EL. Early life
predictors of intelligence in young adulthood and middle age. PLoS ONE.
2021(10):e0228141.
14. Oluwagbemigun K, O’Donovan AN, Berding K, Lyons K, Alexy U, Schmid
M, et al. Long-term dietary intake from infancy to late adolescence is
associated with gut microbiota composition in young adulthood. Am J
Clin Nutr. 2021;113(3):567–576.
13. van Aken L, Kessels RPC, Wingbärmhüle E, van der Veld WM, Egger JIM.
Fluid intelligence and executive functioning more alike than different?
Acta Neuropsychiatrica. 2016;28(1):31–7.
12. Kroke A, Manz F, Kersting M, Reimer T, Sichert-Hellert W, Alexy U, et al. The
DONALD Study: History, current status and future perspectives. Eur J Nutr.
2004;43(1):45–54.
11. Weiss RH. Grundintelligenztest skala 2, CFT 20-R [Basic intelligence test,
revised]. Gottingen: Hogrefe; 2006.
10. Hron K, Filzmoser P, Thompson K. Linear regression with compositional
explanatory variables. J Appl Stat. 2012;39(5):1115–28.
9. Quinn TP, Erb I. Amalgams: data-driven amalgamation for the dimension-
ality reduction of compositional data. NAR Genomics and Bioinformatics.
2020(2). https://academic.oup.com/nargab/article/2/4/faq076/59173
00.
8. Austin PC, Steyerberg EW. The number of subjects per variable required
in linear regression analyses. J Clin Epidemiol. 2015;68(6):627–36. https://
doi.org/10.1016/j.clinepi.2014.12.014.
7. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed.
Boston: Pearson Education, Inc; 2006.
6. O’Toole PW, Jeffery IB. Gut microbiota and aging. Science.
2019;366(6466):647–56.
5. van Aken L, Kessels RPC, Wingbärmhüle E, van der Veld WM, Egger JIM.
Fluid intelligence and executive functioning more alike than different?
Acta Neuropsychiatrica. 2016;28(1):31–7.
4. Kroke A, Manz F, Kersting M, Reimer T, Sichert-Hellert W, Alexy U, et al. The
DONALD Study: History, current status and future perspectives. Eur J Nutr.
2004;43(1):45–54.
3. Weiss RH. Grundintelligenztest skala 2, CFT 20-R [Basic intelligence test,
revised]. Gottingen: Hogrefe; 2006.
2. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed.
Boston: Pearson Education, Inc; 2006.
72. Hawkins MAW, Keirns NG, Helms Z. Carbohydrates and cognitive function. Curr Opin Clin Nutr Metab Care. 2018;21(4):302–7.
73. Lee E, Lee I-E. Impact of drinking alcohol on gut microbiota: recent perspectives on ethanol and alcoholic beverage. Curr Opin Food Sci. 2021;37:91–7.
74. Cheng B, Chu X, Yang X, Wen Y, Jia Y, Liang C, et al. Dietary habit is associated with depression and intelligence: an observational and genome-wide environmental interaction analysis in the UK Biobank cohort. Nutrients. 2021;13(4):1150.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.