Functional conservation of microbial communities determines composition predictability in anaerobic digestion

Qiang Lin[1,#], Lingjuan Li[2,#], Jo De Vrieze[3], Chaonan Li[1], Xiaoyu Fang[1], Xiangzhen Li[1]*

[1] Key Laboratory of Environmental and Applied Microbiology, CAS; Environmental Microbiology Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China

[2] Department of Biology, University of Antwerp, 2610 Wilrijk, Belgium

[3] Center for Microbial Ecology and Technology (CMET), Ghent University, Coupure Links 653, Ghent 9000, Belgium

* Corresponding author: Xiangzhen Li, E-mail: lixz2015@163.com; telephone numbers: +86-028-82890506; fax numbers: +86028-82890288

[#] These authors contributed equally.

Abstract

A major challenge in managing and engineering microbial communities is determining whether and how microbial community responses to environmental alterations can be predicted and explained, especially in microorganism-driven systems. We addressed this challenge by monitoring microbial community responses to the periodic addition of the same feedstock throughout anaerobic digestion, a typical microorganism-driven system where microorganisms degrade and transform the feedstock. The immediate and delayed response consortia were assemblages of microorganisms whose abundances significantly increased on the first or third day after feedstock addition. The immediate response consortia were more predictable than the delayed response consortia and showed a reproducible and predictable order-level composition across multiple feedstock additions. These results stood in both present (16S rRNA gene) and potentially active (16S rRNA) microbial communities and in different feedstocks with different biodegradability and were validated by simulation modeling. Despite substantial species variability, the immediate response consortia aligned well with the reproducible CH_4 production, which was attributed to the conservation of expressed functions by the response consortia throughout anaerobic digestion, based on metatranscriptomic data analyses. The high species variability might be attributed to intraspecific competition and contribute to biodiversity maintenance and functional redundancy. Our results demonstrate reproducible and predictable microbial community responses and their importance in stabilizing system functions.

Keyword: Microbial community response; Predictability; Diversity; Functional redundancy; Anaerobic digestion

1. Introduction

Microorganisms play crucial roles in a wide range of engineered and natural processes, from waste treatment and human health to ecosystem restoration and services [1-3]. Developing the ability to rationally manipulate the microbial community toward beneficial states for ecosystem function, especially in natural and complex settings, is one of the central challenges in microbial ecology [4, 5] and depends on revealing and elucidating the reproducibility and predictability of microbial community dynamics and assemblages. Prior efforts have experimentally validated the reproducibility and predictability of microbial community assembly in laboratory synthetic media [4]. Despite the different environmental origins of microbial communities, all the stabilized communities under the same nutrient resource converged into a similar community structure at the family taxonomic level, therefore indicating that the stable status of microbial communities is predictable and governed by nutrient availability [4]. A previous experiment showed that the stabilized microbial community was not reproducible across replicated sediment–water microcosms and thus indicated unpredictability of the microbial community assemblage [6]. Such predictability, which refers to the ability to predict, depends on the reproducibility among duplicates, such as duplicated cultural conditions [4] or environmental conditions [6]. These prior studies mainly focused on the stable status of microbial communities [4, 6]. Under changes in environmental conditions, the shift process of the microbial community usually includes microbial responses and subsequent stabilization (Fig. 1A). For example, the microbial community responds instantly to environmental changes in hours or days and then stabilizes over a longer time (e.g., weeks, months or years). In this context, it raises a new question (Fig. 1A): Is it reproducible and predictable for microbial community response patterns? If it is predictable, two scenarios are hypothesized: (I) The predictability would gradually decrease as the taxonomic resolution increases (Fig. 1B). This is because the microbial community at a lower taxonomic resolution (e.g., phylum level) shows a simpler structure and a lower taxon diversity than that at a higher taxonomic resolution (e.g., species level). For example, during decades of ecological succession, bacterial community structure at the phylum level remains relatively stable and thus could be predictable, but bacterial community structure at the operational taxonomic unit (OTU, based on 97% similarity of sequences) level changes substantially in an unpredictable manner [7]. (II) There is an inflection point at a specific taxonomic level after which the predictability would substantially decrease (Fig. 1B). This is because stoichiometric and thermodynamic constraints could shape specific microbial community functions that are usually conserved at specific taxonomic levels [8]. For example, the degradation of a specific carbon source shapes the functional conservation of the microbial community at the family level, so the predictability of community structure occurs

at the family level despite substantial variability at the species level [4]. Addressing the above questions and hypotheses (Fig. 1) would advance the mechanistic understanding of community assembly and boost the ability to rationally manipulate microbial communities in the expected directions, yet the above questions and hypotheses remain underexplored, probably due to the high complexity of microbial processes.

Anaerobic digestion (AD) has been widely applied to treat livestock manure and other wastes for renewable energy recovery in the form of CH_4. The feedstock (e.g., manure and/or other wastes) is repeatedly and periodically added to the digesters to maintain (semi)continuous AD. Feedstock addition could increase nutrient availability and change other environmental characteristics in AD, causing shifts in AD performance. AD could be conducted in replicated digestors under manipulable conditions, and there is high microbial diversity, nutrient cycling and energy flow in AD, which allows us to explore the link between microorganisms and system functioning. The feedstock addition could be well controlled to be the same and periodical across the AD process. Due to these advantages, AD has been used as a model system to understand microbial community assembly under changes in feedstock or sludge properties [9, 10], reveal microbial community assemblage differences among replicated reactors [11], and predict digestor performance based on operational parameters [12]. These advantages of the AD system provide a unique opportunity to address the above questions and hypotheses (Fig. 1) by exploring whether and how the responses of the microbial community as well as AD performance to the same and periodical feedstock addition across the AD process are reproducible and predictable and are linked to each other. As continuous feedstock addition is important for AD function maintenance, answering these questions facilitates an understanding of the microorganism mechanisms underlying AD performance, yet such knowledge remains lacking.

To test the above questions and hypotheses (Fig. 1) in this study, AD was conducted with periodic and repeatable addition of swine manure, a widely used livestock manure in AD. Specifically, the responses of performance (mainly represented by CH_4 production) and microbial community to repeatable and periodical feedstock addition across AD were explored and whether microbial response patterns differed between present (revealed by 16S rRNA gene) and potentially active (revealed by 16S rRNA) microorganisms, between immediate (the first day after feedstock addition) response and delayed (the third day after feedstock addition) responses, and between feedstocks with different biodegradability was investigated. The higher reproducibility of the responses of AD performance and microbial community to repeatable and periodical feedstock additions indicated the higher response predictability. If the microbial response was predictable, the microbial community

transcriptional mechanisms underlying this predictability were explored. In addition to predictability, how the variability of the response microorganisms underlies AD functioning was elucidated.

2. Materials and methods

2.1. Setup of the anaerobic digesters

The AD experiments were initiated in nine digesters (volume of each digester was 2 L) and contained 1050 mL swine manure (total solid (TS) content in the total mass was adjusted to 8% by using water based on previous reports [13, 14]), 450 mL seed sludge (TS content = 8%) and 18 g dry rice straw (length of approximately 10 mm). Fresh swine manure was collected from a pig farm near Chengdu and had a TS content of 25.0 ± 0.4% and a volatile solid (VS) content in TS of 77.6 ± 0.5%. Seed sludge was collected from long-term maintained anaerobic digesters of swine manure at 37 ± 1 °C in the laboratory. Straw was collected from a farm near Chengdu, had a TS content of 97.4 ± 0.6% and a VS content in TS of 86.6 ± 0.4% and was added only once at the beginning of AD to balance the nutrients. At the start of the experiments, the digesters were purged with nitrogen gas for 5 minutes to generate anaerobic conditions, and the pH in the slurry was adjusted to 7 with NaOH and HCl solutions. The digesters were operated at a hydraulic retention time (HRT) of 15 days, an organic loading rate (ORL) of 4.1 g VS of swine manure L^{-1} day^{-1}, and at 37 ± 1 °C in temperature-controlled incubators. Each digester was fed 300 mL swine manure (VS = 18.6 g) every three days for 30 days (i.e., initial stage) to reach relatively stable and comparable performances (CH_4 = 0.9 ± 0.2 L L^{-1} d^{-1}). Each digester had two holes on the upper and lower flask walls for feeding and digestate removal (Supplementary Fig. S1) and was manually shaken twice per day. After the initial stage, the nine digesters were randomly grouped into three treatments (triplicate digesters in each treatment): the control treatment, apple waste treatment (AWT) and fructose addition treatment (FAT). The control treatment was fed swine manure for another 10 feeding cycles (FCs) of 30 days, with the same operational conditions as before. The digestate samples in each digester under the control treatment were collected at 13 time points throughout AD, specifically, 2 hours before the first FC and on the first and third days within the first, second, fifth, sixth, seventh and tenth FCs (Fig. 2A).

For the AWT and FAT, the feedstock biodegradability was increased by partially replacing swine manure with apple waste (AWT; rotten apples collected from markets in Chengdu, with TS content of 17.3 ± 2.2% and VS content of 94.0 ± 1.6%) or fructose (FAT; $C_6H_{12}O_6$, purity ≥ 99%, CAS number: 57–48-7, Sangon Biotech, China) in the easily biodegradable (EB) stage. In this stage, feedstock was added every three days for a total of 5

FCs, keeping the 18.6 g VS and 300 mL volume constant across different FCs and different treatments. In each FC, AWT and FAT were fed 300 mL of swine manure (VS=11.4 g) with 44.3 g of apple waste (VS = 7.2 g) or 7.2 g of fructose powder (VS = 7.2 g). In the swine manure (SM) stage, the feedstock addition in AWT and FAT was changed back to before and contained only 300 mL of swine manure (VS = 18.6 g) for another 5 FCs. All operational parameters remained the same across different stages and treatments. The VS ratio of apple waste versus manure in AWT was determined based on a preliminary experiment [15]. The digestate samples in each of the six digesters separately in AWT and FAT were collected at 13 time points, specifically, 2 hours before the first FC, and on the first and third day within the first, second, fifth (the fifth and earlier FCs were within EB stage), sixth (the sixth and later FCs were within SM stage) seventh and tenth FCs during AD (Supplementary Fig. S2).

In the control treatment, the periodical addition of the same feedstock (i.e., swine manure, a common feedstock widely applied in AD [16, 17]) under the same operational conditions and multiple samplings across the AD process allowed us to explore the reproducibility and predictability of the responses of the microbial community as well as AD performance to feedstock addition. In the AWT and FAT, apple waste (a major fruit waste widely applied in AD [17, 18]) and fructose, which showed a higher biodegradability (indicated by a higher methane production [19]) than swine manure (Supplementary Table S1-S2), were used to explore the influence of feedstock biodegradability on the reproducibility and predictability of microbial responses.

2.2. Measurements and chemical analysis

The characteristics of swine manure and apple waste were measured (Supplementary method S1) and are shown in Supplementary Table S1. The digestate samples were centrifuged at 13,400 ×*g* for 5 minutes and were immediately used for DNA and RNA extraction using an Ezup Column Soil DNA Purification Kit (Cat. No. B518263; Sangon Biotech, China) and an RNAprep pure Cell/Bacteria Kit (Cat. No. DP430; TIANGEN, China) including DNase, respectively. A reverse transcription kit (Cat. No. PR6901; Thermo, USA) was used to synthesize complimentary DNA (cDNA) from the RNA. The supernatant was filtered through a 0.22 μm filter (Cat. NO. SLGP033RS; Millipore, USA) and then used to quantify the volatile fatty acids (acetic acid, propionic acid and butyric acid) that are important metabolic intermediates contributing to methane production in AD by high-performance liquid chromatography (HPLC, Agilent 1260), as described previously [15], the pH by a pH monitor (Cat. No. G001630; Sangon Biotech, China) and the NH_4^+-N by Nessler's reagent colorimetric method

[20]. The biogas production and CH_4 and H_2 contents within the biogas were measured daily by using a water displacement method [21] and by a gas chromatography system (SP-2000, Beifen, China) as described previously [15]. The ideal gas law [22] was used to standardize the volumes of CH_4 and H_2 at standard temperature (273.15 K) and pressure (101.325 kPa) (Supplementary Table S2). The details of these measurements are shown in Supplementary method S1.

2.3. Sequencing and data analysis

To characterize the microbial community, the 16S rRNA gene was amplified from the DNA and cDNA with the primers 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3') (Supplementary method S2) [23]. In total, 234 amplification samples were sequenced on the MiSeq System (Illumina) with the Reagent Kit v3 2 × 250 bp. To achieve clean sequences, low-quality sequences and short reads (length < 200 bp) were removed by using QIIME Pipeline Version 1.9.0 [24], and UCHIME was used to remove chimeric sequences [25]. The clean sequences were clustered at 97% similarity with Cd-hit [26] into operational taxonomic units (OTUs). The taxonomic assignment of OTU representative sequences was conducted based on the RDP database (version 18) [27] using blastn with an e-value $\leq 1\times10^{-7}$. Absolute singletons and nonprokaryotic reads were removed, and the final reads were normalized to 8,713 per sample using the *otutab_rare* function in the Usearch Pipeline [28].

To reveal the gene expression of microbial communities during AD, total RNA, excluding ribosomal RNA, with the RiboMinusTM kit (Lot. No. 1539791; Invitrogen, USA), was applied for metatranscriptomic sequencing by a HiSeq 2000 System (Illumina). To explore the stability of microbial community transcriptional profiles under the same feedstock during AD, nine samples from triplicate digesters in the control treatment (continuously fed swine manure throughout the AD) on the first day within the first, fifth and sixth FCs were subjected to metatranscriptomic sequencing. Metatranscriptomic samples with an average of 33,069,276 ± 169,270 sequences per sample were uploaded to MG-RAST (version 4.0.3) [29] for quality control and functional and taxonomic annotation based on the KEGG Orthologs database and RefSeq database with default parameters. A total of 3,170,885 ± 71,070 functional categories per sample were identified. The sequences in specific taxonomic populations were extracted and then functionally assigned using the KEGG Orthologs database with default parameters in MG-RAST. The sequences in specific KEGG pathways were extracted and then taxonomically assigned using the RefSeq database with default parameters in MG-RAST.

198

199 **2.4. Statistical analysis**

200 To reveal the microbial response to feedstock addition, microbial communities before and after feedstock
201 addition within an FC were compared by DESeq2 [30]. The digesters were fed every three days, and the
202 sampling on the third day within an FC occurred 2 h before the next feedstock addition for the next FC. We
203 compared microbial communities specifically between 2 h before and on the first day within the first FC, the
204 second FC, the sixth FC and the seventh FC to reveal microbial community responses of four times throughout
205 the AD process (Fig. 2A and Supplementary Fig. S2). This comparison could reflect the immediate responses of
206 microorganisms to feedstock addition. We also revealed the delayed responses of microorganisms on the third
207 day after feedstock addition within an FC by comparing microbial communities between 2 h before and on the
208 third day within the first FC, the second FC, the sixth FC and the seventh FC (Fig. 2A and Supplementary Fig.
209 S2). As increases in microbial abundances usually facilitate microbial contributions to ecosystem functioning,
210 only the microorganisms whose relative abundances significantly ($P < 0.05$) increased after feedstock addition
211 were revealed by the above comparisons in DESeq2. The *p* values were adjusted by the Benjamini-Hochberg
212 correction [31]. The composition and structure of the response consortia (i.e., the assemblages of response
213 microorganisms) were compared across four different responses throughout the AD, and the consortium
214 heterogeneity based on Bray–Curtis dissimilarity was calculated by the function *betadisper* in the R vegan
215 package. Lower heterogeneity implied higher reproducibility and predictability.

216 To generalize the features of the observed communities 2 h before feedstock addition despite different
217 feeding cycles, the statistical model "SparseDOSSA 2" [32] was used to generate 100 simulated communities by
218 parameterizing the observed microbial community profiles before feedstock addition. During simulation, the
219 function fit_SparseDOSSA2, which fits the SparseDOSSA2 model to microbial count, was used. Another 100
220 simulated communities to generalize the features of the observed communities on the first day after feedstock
221 addition were generated by using the same procedure. DESeq2 was used to compare the two sets of the 100
222 simulated communities to further validate the microbial community changes after feedstock addition.

223 To reveal the metabolic functions of OTUs, FAPROTAX, which is a database that assigns prokaryotic
224 microorganisms to established metabolic functions based on the current literature on cultured strains [33], was
225 used with default parameters. Principal coordinate analysis (PCoA) based on Bray–Curtis dissimilarity with 999

permutations was conducted in the R vegan package [34]. The Wilcoxon test was conducted by the R vegan package. The results were visualized using the R package ggplot2 [35].

3. Results and Discussions

3.1. Predictability of response consortia shows an inflection point at the order taxonomic level

In AD, the repeatable and periodical feedstock addition in the control treatment resulted in reproducible and predictable CH_4 daily production, namely, CH_4 production reached a similar peak (1.07 ± 0.04 L L^{-1}) on the first day and decreased to a similar extent (0.84 ± 0.06 L L^{-1}) on the third day after each feedstock addition during the AD process (Fig. 2B). The cumulative CH_4 production from the first day to the third day after feedstock addition reached a similar level across different feeding cycles (Supplementary Fig. S3). Therefore, CH_4 production was reproducible and predictable under constant feedstock addition in AD.

CH_4 is a major end-product in the AD food chain (hydrolysis, acidogenesis, acetogenesis and methanogenesis) that is driven jointly by different microorganisms. Under the dynamics of microbial community composition and structure during AD (Supplementary Fig. S4), it is important to clarify whether the microbial community response to feedstock addition is also reproducible and predictable, underlying the reproducible AD performance. To address this question, responses of the microbial community to feedstock addition at four times within different feeding cycles throughout the AD process were explored (Fig. 2), and the microorganisms that showed an immediate response (IR, the first day after feedstock addition) were revealed (Fig. 2C). Microorganisms with distinct life strategies (r- and K-strategies, or copiotrophs and oligotrophs) usually show different response patterns to changes in nutrient availability [36, 37]. To thoroughly reveal microbial response patterns, microorganisms that showed a delayed response (DR, the third day after feedstock addition) were also revealed (Fig. 2C). To reveal the composition and structure of the response consortia, the response OTUs (Fig. 2C) were grouped into different taxonomic levels (Fig. 3A). The IR OTUs converged into similar phylum-level consortia dominated by Firmicutes (average 25 OTUs), Bacteroidetes (15) and Proteobacteria (12), similar class-level consortia dominated by *Clostridia* (20), *Bacteroidia* (14) and *Gammaproteobacteria* (9), and similar order-level consortia dominated by *Clostridiales* (20) and *Bacteroidales* (14) across the different responses (Fig. 3A). However, the response consortia at the higher taxonomic levels (i.e., family or genus levels) did not show a similar composition and thus were less conserved across different responses (Fig. 3A). This finding was further supported by the fact that the heterogeneity of the response consortia across different responses remained

relatively stable and lower at the phylum (mean = 0.22), class (0.24) and order (0.24) levels but steeply increased at the family (0.35) and genus (0.37) levels (Wilcoxon test $P < 0.05$ for both the 16S rRNA gene and 16S rRNA datasets) (Fig. 3B). Higher heterogeneity indicates lower reproducibility and predictability. Therefore, the IR consortia were reproducible and predictable at the order level, after which the reproducibility and predictability dramatically dropped. This finding occurred in both datasets of the 16S rRNA gene and 16S rRNA. The two datasets of the 16S rRNA gene and 16S rRNA could reveal present or potentially active microorganisms, respectively [38, 39]. Therefore, not only present but also potentially active microorganisms showed reproducibility and predictability in response to feedstock additions. This might be because feedstock additions caused selection on present microorganisms and triggered microbial activities. Unlike the IR consortia, the DR consortia across different responses were less reproducible and predictable. For example, the DR consortia specifically based on the 16S rRNA gene did not show a similar composition across different responses (Fig. 3A). The heterogeneity of the DR consortia across different responses was higher (Wilcoxon test $P = 0.002$) than that of the IR consortia (Fig. 3B). Interestingly, the heterogeneity of both IR and DR consortia dramatically dropped after the order level (Wilcoxon test $P < 0.05$ for all) (Fig. 3B). This finding answered our hypotheses (Fig. 1) and revealed that the reproducibility and predictability did not gradually decrease with increasing taxonomic resolution but showed an inflection point at the order level.

The predictable microbial community assemblage has been found to be highly dependent on the single limiting resource [4], implying an effect of feedstock specificity on predictability. Swine manure, as a major livestock manure treated through AD, represents a common feedstock used in the AD process [16, 17]. Swine manure is not a mono-component but consists of complex and diverse components (Supplementary Table S1) [40], so it might not cause a high feedstock specificity to influence our findings. This study did not explore all the diverse feedstocks for AD, so the possible impact of feedstock specificity on our findings may not be entirely ruled out. Feedstock biodegradability is a crucial determinant of AD performance. To explore the effect of feedstock biodegradability on our findings, some swine manure was replaced with more easily degradable apple waste or fructose (Supplementary Fig. S2). The addition of more easily degradable feedstocks revealed a similarly changing trend of the response consortia reproducibility and predictability along taxonomic resolution, and the reproducibility and predictability showed an inflection point at the order level, especially for the IR consortia (Supplementary Fig. S5). Therefore, the reproducibility and predictability of the microbial community response to feedstock addition was barely influenced by feedstock biodegradability. This is probably because the feedstock composition was not completely changed, despite the change in biodegradability.

The differences in the reproducibility and predictability between the IR and DR consortia are probably attributed to their life strategies. The IR and DR consortia could separately approximate the r-strategists and K-strategists that usually show different response patterns to nutrient availability [36]. The IR consortia obviously differed from the DR consortia (see PCo1 in Supplementary Fig. S6), with less than 32% of common OTUs, further supporting that they could approximate distinct strategists. The r-strategists are more sensitive than K-strategists to the shift of nutrient availability, and thus would respond more intensively, which is supported by the higher OTU number and change-folds in the immediate responses than in the delayed responses (Fig. 2C). The higher response intensity might contribute to reproducibility and predictability. The nutrient availability and other environmental factors are immediately changed after feedstock addition, and subsequently, such changes gradually vanish, which probably causes an immediate increase and subsequent decrease in environmental selection. Therefore, the IR consortia were subjected to a higher level of environmental selection than the DR consortia, which reduced the randomness and thus increased the predictability.

3.2. Function of response consortia is conserved at the order taxonomic level

The IR consortia always corresponded to a peak of CH_4 production after feedstock addition (Fig. 2). A question arises: how did the reproducible and predictable IR consortia underlie the reproducible and predictable CH_4 production? To reveal generic characteristics of IR consortia, two sets of 100 simulated communities were generated separately to generalize the features of the observed communities before and on the first day after feedstock addition, and then the two sets of simulated communities were compared. The comparison resulted in a similar result with that in Fig. 3A, that the IR consortia at the phylum level were dominated by Firmicutes, Bacteroidetes and Proteobacteria, the IR consortia at the class level were dominated by *Clostridia*, *Bacteroidia* and *Gammaproteobacteria*, and the IR consortia at the order level were dominated by *Clostridiales* and *Bacteroidales* (Fig. 4). This result further confirmed the reproducibility and predictability of the IR consortia at the order level, as well as the dominance of the orders *Clostridiales* and *Bacteroidales*. The reproducibility and predictability of the response consortia at the order level implied functional conservation at the order level, underlying predictable and reproducible CH_4 production. Simple carbon resource utilization and organic phosphorus acquisition are supposed to be conserved at a finer taxonomic resolution between the genus and species levels [8]. In our study, methane production from swine manure was through complex processes involving diverse microorganisms, so the functions of the response consortia were conserved at a relatively

lower taxonomic resolution, namely, the order level. In the metatranscriptomic data, the expressed genes were taxonomically assigned to achieve microbial communities that expressed these genes. Across different feeding cycles during AD, the variations in the microbial community that expressed genes substantially increased from the order level to the family level, with a significantly higher increasing extent (Wilcoxon test $P < 0.001$) than those between other adjacent taxonomic levels (Fig. 5A). These results supported the inflection point of predictability and functional conservation at the order level.

The functional conservation at the order level is also probably linked to dominance of the orders *Clostridiales* and *Bacteroidales*. The relative contributions of these two taxa to the top 10 metabolic pathways expressed during AD were revealed (Fig. 5B). Compared to the other orders, these two taxa contributed majorly to the expression of most of these pathways, such as amino acid metabolism, carbohydrate metabolism, nucleotide metabolism, metabolism of cofactors and vitamins, translation and membrane transport. This result revealed the critical functional roles of these two taxa in AD performance. As these metabolic pathways are critical to CH_4 production in AD, the consistent dominance of these two taxa in the IR consortia underlies the reproducible peak of CH_4 production after feedstock addition and contributes to the conservation of these critical functions at the order level. The top 5 orders in the IR consortia (Fig. 3A and 4) were retrieved. The variations in gene expression across different feeding cycles during AD were significantly lower (Wilcoxon test $P < 0.001$ for all) within *Clostridiales* and *Bacteroidales* than within *Synergistales*, *Lactobacillales* and *Pseudomonadales* (Fig. 5C). This result indicated a high stability of the expressed functions within these two taxa during AD, which provided functional conservation and thus underlay the reproducible CH_4 production. Collectively, the consistent dominance of *Clostridiales* and *Bacteroidales* in the response consortia and their transcriptional profiles indicated their crucial roles in the alignment between the reproducible and predictable IR consortia and CH_4 production. *Clostridiales* and *Bacteroidales* have been previously found in AD [41, 42], but their roles in AD remain elusive. This study provided novel experimental evidence to reveal their roles in underlying the predictable microbial community response and reproducible CH_4 production. In particular, the two taxa provided functional conservation, resulting in a predictability inflection point at the order level. Only a few methanogens (0-3 OTUs) were detected in each response consortium during AD (Fig. 3A), probably because the added feedstock could barely be utilized directly by methanogens. The expression of total methanogenic genes after feedstock addition was abundant (6%-9%) in total transcripts and did not significantly change (Wilcoxon test $P > 0.05$ for all) across the AD process, with the dominance of acetoclastic and hydrogenotrophic methanogenesis compared to methylotrophic methanogenesis (Supplementary Fig. S7). Therefore, despite only a

few response methanogens, the abundant and relatively stable expression of methanogenic genes also contributed to the reproducible peak of CH_4 production after feedstock addition.

3.3. High OTU variability of the response consortia promotes biodiversity maintenance and functional redundancy

The reproducibility and predictability of the response consortia substantially decreased at a finer taxonomic resolution because of the high OTU variability. For the delayed response, there were no common OTUs across all responses; for the immediate response, less than 4.4% of the total response OTUs were common across all responses (Fig. 6A). This result was in line with a prior study in which laboratory incubations of diverse-sourced microbial communities under a single limiting resource resulted in a highly variable community assemblage at the species level, despite a predictable community assemblage at the family level [4]. Under functional conservation, high species variability is also observed in natural microbial community assemblages [43]. However, no specific reasons for such high species variability have been found yet. Feedstock addition caused a substantial increase in nutrient availability. High nutrient availability could generate strong environmental selection [44], so the high species variability in the response consortia is unlikely to be attributed to random effects. The identical feedstock (i.e., swine manure) was added in the same way across different feeding cycles in the control treatment during AD, and a similar and reproducible CH_4 production was detected after feedstock addition, so there was unlikely to be a batch effect. The storage effect tells that no species would be the best consistently in a fluctuating environment [45], which is one of the important theories explaining diverse species coexistence. One of the mechanisms underlying the storage effect is intraspecific competition that occurs among individuals within a population, is usually triggered as the population expands, and could in turn constrain species population expansion to maintain the coexistence of diverse species. In this study, the response species were identified if their abundances significantly increased after feedstock addition. The increased population (i.e., abundance) of a species would intensify intraspecific competition, especially in resource-rich environments [46]. Therefore, in the resource-rich AD system, especially after feedstock addition, increased intraspecific competition is likely to reduce the fitness of the species, which could result in the species not responding to the following feedstock addition or their response extent becoming too weak to be detected. This might explain why few species consistently responded to feedstock addition in all feeding cycles during AD, namely, the high species variability.

373 Despite the high species variability in the response consortia, 92%-96% of all the response OTUs were
374 present (revealed by DNA data) and active (revealed by RNA data) at no less than 10 time points of the 13 time
375 points throughout the AD (Fig. 6B), reflecting a high extent of coexistence of diverse species. This result
376 indicated that the high species variability facilitated the maintenance of species diversity in AD. In a laboratory
377 incubation of synthetic microbial communities, a similar result has also been found that the
378 fluctuation/variability of microbial communities reinforces community diversity [5]. Microbial diversity has
379 been supposed to drive multifunctionality in ecosystems [47]. In AD, CH_4 production from swine manure
380 necessitates multifunctionality (i.e., hydrolysis, acidogenesis, acetogenesis and methanogenesis), so microbial
381 diversity is the guarantee of AD functioning. Therefore, the high species variability in the response consortia
382 facilitated AD functioning by maintaining biodiversity.

383 Species variability within functional groups has been found to facilitate functional redundancy in
384 natural microbial communities [43]. The response consortia played crucial roles in the degradation and
385 transformation of the added feedstock. However, it is unknown whether the high species variability of the
386 response consortia facilitates functional redundancy. To address this question, we revealed the metabolic
387 functions of each response OTU using a previously established approach [43]. The AD process primarily
388 depends on the microbial functions of chemoheterotrophy and fermentation, and we found that most of response
389 OTUs could perform the two functions (Fig. 6C). This result indicated a high functional redundancy of the two
390 functions. In addition to the two functions, multiple OTUs were also involved in aromatic compound
391 degradation, cellulolysis and the metabolism of other compounds (Fig. 6C), further implying high functional
392 redundancy. Therefore, the high species variability in the response consortia provided high functional
393 redundancy. Under high functional redundancy, AD functioning will be maintained and stable, even if some
394 species are lost due to disturbances. Feedstock addition in AD usually causes environmental disturbances by
395 shifting nutrient availability and other environmental characteristics. This study found a high functional
396 redundancy in the response consortia, which provided another mechanism (in addition to functional conservation
397 at the order level) underlying the stable and reproducible CH_4 production after feedstock addition.

398 Environmental alterations or fluctuations are ubiquitous, and microbial community responses to such
399 alterations or fluctuations are usually linked to shifts in microorganism-driven functions and provide clues to
400 understand the subsequent and stabilized assemblage of communities. Based on the constant feedstock in
401 microorganism-driven AD, this study reveals the reproducibility and predictability of the response
402 microorganisms and AD performance to feedstock addition. The reproducibility and predictability did not

change gradually along the taxonomic resolution but showed an inflection point at the order level. Importantly, we provide a mechanistic understanding that despite substantial species variability, the expressed functions of the response consortia are conserved, underlying the stable and reproducible AD performances. These findings provide novel evidence for the predictability and reproducibility of microbial responses, which contribute to the understanding of the mechanisms underlying microbial community assembly and facilitate performance prediction in microorganism-driven systems. These findings could be utilized as a stepping stone toward developing a quantitative theory to engineer microbial communities and system functions.

Availability of data and material

The original amplicon sequencing data are deposited at the European Nucleotide Archive by accession number PRJEB59216. The original metatranscriptomic sequencing data are deposited in the project “mgp80358” at MG-RAST.

Funding

This study was supported by the Open Found of Key Laboratory of Environmental and Applied Microbiology CAS (11050004GB), and China Biodiversity Observation Networks (Sino BON).

Authors' contributions

QL conceived this study, conducted the experiment, analyzed data, and wrote and revised the manuscript. LJL analyzed data and revised the manuscript. JDV, CNL and XZL revised the manuscript. XYF conducted the experiment.

Declaration of Competing Interest

The authors declare that they have no known conflict of interest.

Acknowledgements

Thank all people who work against COVID-19 pandemic and appreciate the support from our families for our works.

References

1. Lin Q, De Vrieze J, Li C, Li J, Li J, Yao M, et al. Temperature regulates deterministic processes and the succession of microbial interactions in anaerobic digestion process. Water Res. 2017;123:134-143.
2. Coban O, De Deyn GB, van der Ploeg M. Soil microbiota as game-changers in restoration of degraded lands. Science. 2022;375:abe0725.
3. Coyte KZ, Schluter J, Foster KR. The ecology of the microbiome: networks, competition, and stability. Science. 2015;350:663-666.
4. Goldford JE, Lu N, Bajić D, Estrela S, Tikhonov M, Sanchez-Gorostiaga A, et al. Emergent simplicity in microbial community assembly. Science. 2018;361:469-474.
5. Hu J, Amor DR, Barbier M, Bunin G, Gore J. Emergent phases of ecological diversity and dynamics mapped in microcosms. Science. 2022;378:85-89.
6. Pagaling E, Strathdee F, Spears BM, Cates ME, Allen RJ, Free A. Community history affects the predictability of microbial ecosystem development. ISME J. 2014;8:19-30.
7. Lin Q, Dini-Andreote F, Meador TB, Angel R, Meszárošová L, Heděnec P, et al. Microbial phylogenetic relatedness links to distinct successional patterns of bacterial and fungal communities. Environ Microbiol. 2022;24:3985–4000.
8. Martiny JB, Jones SE, Lennon JT, Martiny AC. Microbiomes in light of traits: A phylogenetic perspective. Science. 2015;350:aac9323.
9. Klang J, Szewzyk U, Bock D, Theuerl S. Effect of a profound feedstock change on the structure and performance of biogas microbiomes. Microorganisms. 2020;8:169.
10. Langer SG, Gabris C, Einfalt D, Wemheuer B, Kazda M, Bengelsdorf FR. Different response of bacteria, archaea and fungi to process parameters in nine full-scale anaerobic digesters. Microb Biotechnol. 2019;12:1210-1225.
11. Steinberg LM, Martino AJ, House CH. Convergent Microbial Community Formation in Replicate Anaerobic Reactors Inoculated from Different Sources and Treating Ersatz Crew Waste. Life. 2021;11:1374.
12. Long F, Wang L, Cai W, Lesnik K, Liu H. Predicting the performance of anaerobic digestion using machine learning algorithms and genomic data. Water Res. 2021;199:117182.
13. Lin Q, He G, Rui J, Fang X, Tao Y, Li J, et al. Microorganism-regulated mechanisms of temperature effects on the performance of anaerobic digestion. Microb Cell Fact. 2016;15:1-18.
14. Wang Z, Jiang Y, Wang S, Zhang Y, Hu Y, Hu Z-h, et al. Impact of total solids content on anaerobic co-digestion of pig manure and food waste: Insights into shifting of the methanogenic pathway. Waste Manage. 2020;114:96-106.
15. Lin Q, De Vrieze J, Fang X, Li L, Li X. Labile carbon feedstocks trigger a priming effect in anaerobic digestion: An insight into microbial mechanisms. Bioresour Technol. 2022;344.
16. Cao Q, Zhang W, Lian T, Wang S, Yin F, Zhou T, et al. Revealing mechanism of micro-aeration for enhancing volatile fatty acids production from swine manure. Bioresour Technol. 2022;365:128140.
17. Karki R, Chuenchart W, Surendra K, Shrestha S, Raskin L, Sung S, et al. Anaerobic co-digestion: Current status and perspectives. Bioresour Technol. 2021;330:125001.
18. Lian T, Zhang W, Cao Q, Wang S, Yin F, Chen Y, et al. Optimization of lactate production from co-fermentation of swine manure with apple waste and dynamics of microbial communities. Bioresour Technol. 2021;336:125307.
19. Li Y, Zhang R, Liu G, Chen C, He Y, Liu X. Comparison of methane production potential, biodegradability, and kinetics of different organic substrates. Bioresour Technol. 2013;149:565-569.
20. Hart SC, Stark JM, Davidson EA, Firestone MK. Nitrogen mineralization, immobilization, and nitrification. Methods of Soil Analysis: Part 2—Microbiological and Biochemical Properties. 1994:985-1018.
21. Lin Q, De Vrieze J, Li J, Li X. Temperature affects microbial abundance, activity and interactions in anaerobic digestion. Bioresour Technol. 2016;209:228-236.
22. Bludman SA, Vanriper KA. Equation of state of an ideal ferni gas. Astrophys J. 1977;212:859-872.
23. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012;6:1621-1624.
24. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7:335-336.
25. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics. 2011;27:2194-2200.
26. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 2006;22:1658-1659.

27. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 2014;42:D633-D642.
28. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460-2461.
29. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, et al. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics. 2008;9:386.
30. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:1-21.
31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B. 1995;57:289-300.
32. Ma S, Ren B, Mallick H, Moon YS, Schwager E, Maharjan S, et al. A statistical model for describing and simulating microbial community profiles. PLoS Comput Biol. 2021;17:e1008913.
33. Louca S, Parfrey LW, Doebeli M. Decoupling function and taxonomy in the global ocean microbiome. Science. 2016;353:1272-1277.
34. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara R, et al. Package 'vegan'. Community ecology package, version. 2013;2:1-295.
35. Wickham H. ggplot2. Wiley Interdiscip Rev Comput Stat. 2011;3:180-185.
36. Klappenbach JA, Dunbar JM, Schmidt TM. rRNA operon copy number reflects ecological strategies of bacteria. Appl Environ Microb. 2000;66:1328-1333.
37. Fierer N, Bradford MA, Jackson RB. Toward an ecological classification of soil bacteria. Ecology. 2007;88:1354-1364.
38. De Vrieze J, Pinto AJ, Sloan WT, Ijaz UZ. The active microbial community more accurately reflects the anaerobic digestion process: 16S rRNA (gene) sequencing as a predictive tool. Microbiome. 2018;6:63.
39. Kamke J, Taylor MW, Schmitt S. Activity profiles for marine sponge-associated bacteria obtained by 16S rRNA vs 16S rRNA gene comparisons. ISME J. 2010;4:498-508.
40. Yang Y, Yang F, Huang W, Huang W, Li F, Lei Z, et al. Enhanced anaerobic digestion of ammonia-rich swine manure by zero-valent iron: With special focus on the enhancement effect on hydrogenotrophic methanogenesis activity. Bioresour Technol. 2018;270:172-179.
41. Poirier S, Bize A, Bureau C, Bouchez T, Chapleur O. Community shifts within anaerobic digestion microbiota facing phenol inhibition: towards early warning microbial indicators? Water Res. 2016;100:296-305.
42. Tsapekos P, Khoshnevisan B, Zhu X, Treu L, Alfaro N, Kougias PG, et al. Lab-and pilot-scale anaerobic digestion of municipal bio-waste and potential of digestate for biogas upgrading sustained by microbial analysis. Renewe Energ. 2022;201:344-353.
43. Louca S, Jacques S, Pires AP, Leal JS, Srivastava DS, Parfrey LW, et al. High taxonomic variability despite stable functional structure across microbial communities. Nat Ecol Evol. 2016;1:1-12.
44. Lin Q, Li L, Adams JM, Heděnec P, Tu B, Li C, et al. Nutrient resource availability mediates niche differentiation and temporal co-occurrence of soil bacterial communities. App Soil Ecol. 2021:103965.
45. Chesson PL, Warner RR. Environmental variability promotes coexistence in lottery competitive systems. Am Nat. 1981;117:923-943.
46. LaManna JA, Walton ML, Turner BL, Myers JA. Negative density dependence is stronger in resource-rich environments and diversifies communities when stronger for common but not rare species. Ecol Lett. 2016;19:657-667.
47. Delgado-Baquerizo M, Maestre FT, Reich PB, Jeffries TC, Gaitan JJ, Encinar D, et al. Microbial diversity drives multifunctionality in terrestrial ecosystems. Nat Commun. 2016;7:10541.

562 **Figure legend**

563 Fig. 1 The shift process of microbial community under environmental change (A). The shift process of microbial
564 community usually includes microbial responses and subsequent stabilization. At the beginning when microbial
565 community face to environmental change, they show responses (changes in community composition and
566 structure) in hours or days, which usually shifts microorganism-driven system performance/functions (e.g.,
567 nutrient degradation and transformation). Whereafter, it might take weeks, months or years for microbial
568 community to be stabilized. If microbial community response pattern is predictable, two scenarios are
569 hypothesized (B): (I) the predictability decreases as the increase of taxonomic resolution (the blue line), because
570 community structure at a higher taxonomic resolution has a higher taxa diversity to be predicted; (II) the
571 predictability shows an inflection point at specific taxonomic level, underlying function conservation (the orange
572 line), because of stoichiometric and thermodynamic constraints.

573 Fig. 2 Experimental design to explore microbial community response (A) to feedstock addition (i.e., swine
574 manure), by comparing microbial communities before and after feedstock addition. In total, four responses (i.e.,
575 Res1, Res2, Res3 and Res4) across multiple feeding cycles (FC) throughout anaerobic digestion are studied. The
576 response of AD performance to feedstock addition is represented by CH_4 production in triplicate digesters (B).
577 The immediate response (IR) and delayed response (DR) OTUs in DNA (16S rRNA gene) and RNA (cDNA
578 from 16S rRNA) datasets are shown by points that are sized by OTUs average relative abundances (C).

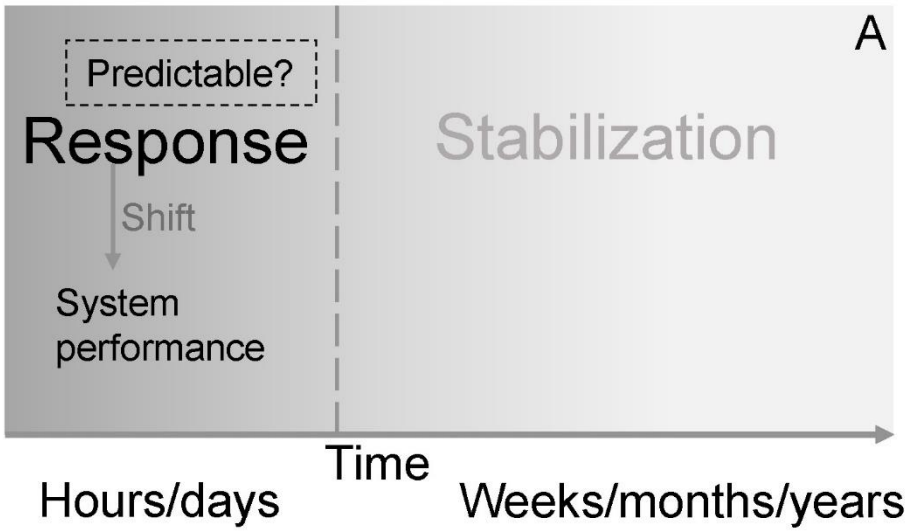
579 Fig. 3 The composition and structure of response consortia (the assemblages of response microorganisms) at
580 different taxonomic levels (A). The Y-axis shows the number of response OTUs affiliated to specific taxonomic
581 taxa. At each taxonomic level, only the top 10 taxa are shown. The heterogeneity (B) of the response consortia
582 across Res1, Res2, Res3 and Res4 is shown at different taxonomic levels. DNA:16S rRNA gene; RNA: cDNA
583 from 16S rRNA.

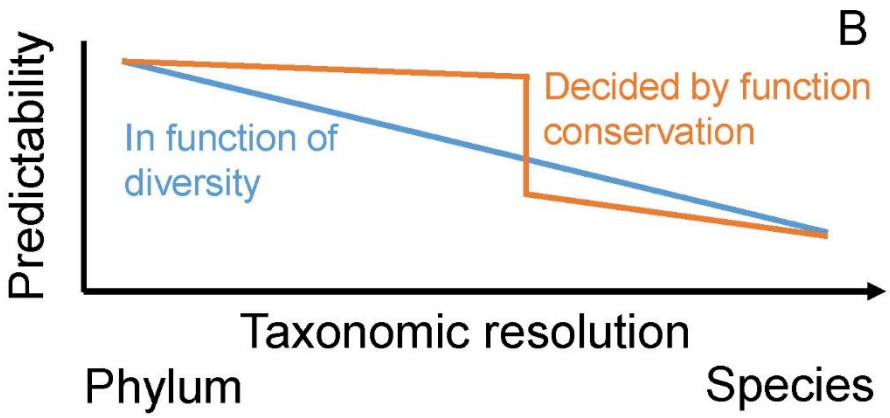
584 Fig. 4 The generic characteristics of immediate response consortia (the assemblages of immediate response
585 microorganisms) at different taxonomic levels, based on the comparison between 100 simulated communities
586 generalizing community features before feedstock addition, and 100 simulated communities generalizing
587 community features on the first day after feedstock addition. The Y-axis shows the number of response OTUs
588 affiliated to specific taxonomic taxa. At each taxonomic level, only the top 10 taxa are shown. DNA:16S rRNA
589 gene; RNA: cDNA from 16S rRNA.

590 Fig. 5 Variations of gene expressions throughout anaerobic digestion at different taxonomic levels (A). The sub-
591 figure on the top-left shows the variation ratio between two neighboring taxonomic levels. The gene expression
592 abundances of the top 10 metabolic pathways (KEGG level 2) and the relative abundances of *Clostridiales*,
593 *Bacteroidales* and other orders within these expressed pathways are shown (B). A point represents an order.
594 Within the top 5 response orders (shown in the Fig. 3A), the variations of their gene expressions throughout
595 anaerobic digestion are shown (C). Mean ± SD are shown.

596 Fig. 6 The percentage (A) of common OTUs that are present in all Res1, Res2, Res3 and Res4, among all the
597 response OTUs. The frequency (B) of each response OTU present in 13 time points (shown in Fig. 2A) during
598 AD. Functional profiles (C) of the response OTUs. Some OTUs are involved in more than one functions. Only
599 top ten functions are shown.

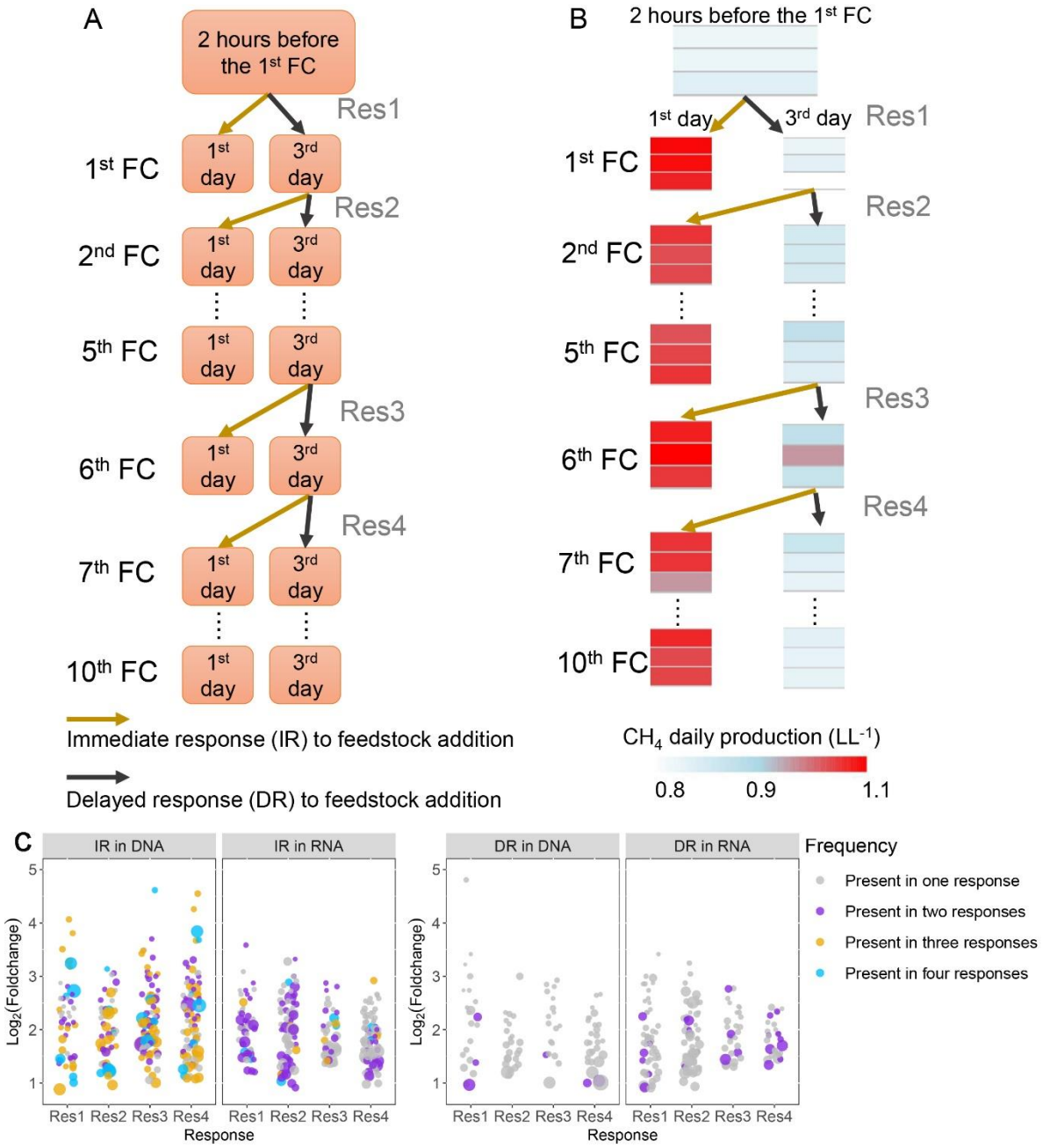
600





601

602 Fig. 1


A
2 hours before the 1st FC
Res1
1st FC
1st day
3rd day
Res2
2nd FC
5th FC
Res3
6th FC
Res4
7th FC
10th FC
B
2 hours before the 1st FC
1st day
3rd day
Immediate response (IR) to feedstock addition
Delayed response (DR) to feedstock addition
CH4 daily production (LL-1)
0.8
0.9
1.1
C
IR in DNA
IR in RNA
DR in DNA
DR in RNA
Log2(Foldchange)
Res1 Res2 Res3 Res4
Response
Frequency
Present in one response
Present in two responses
Present in three responses
Present in four responses

603

604 Fig. 2

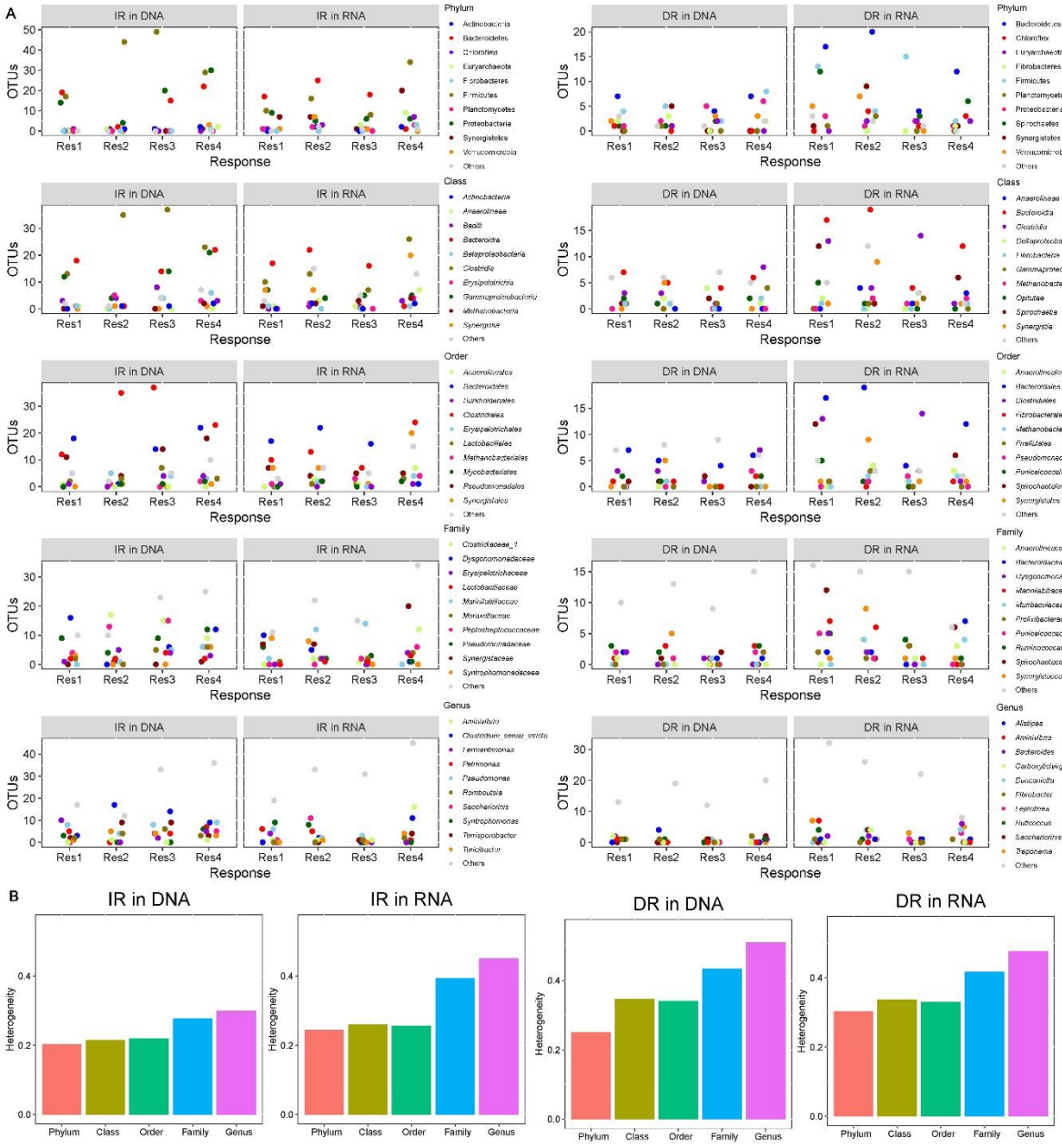
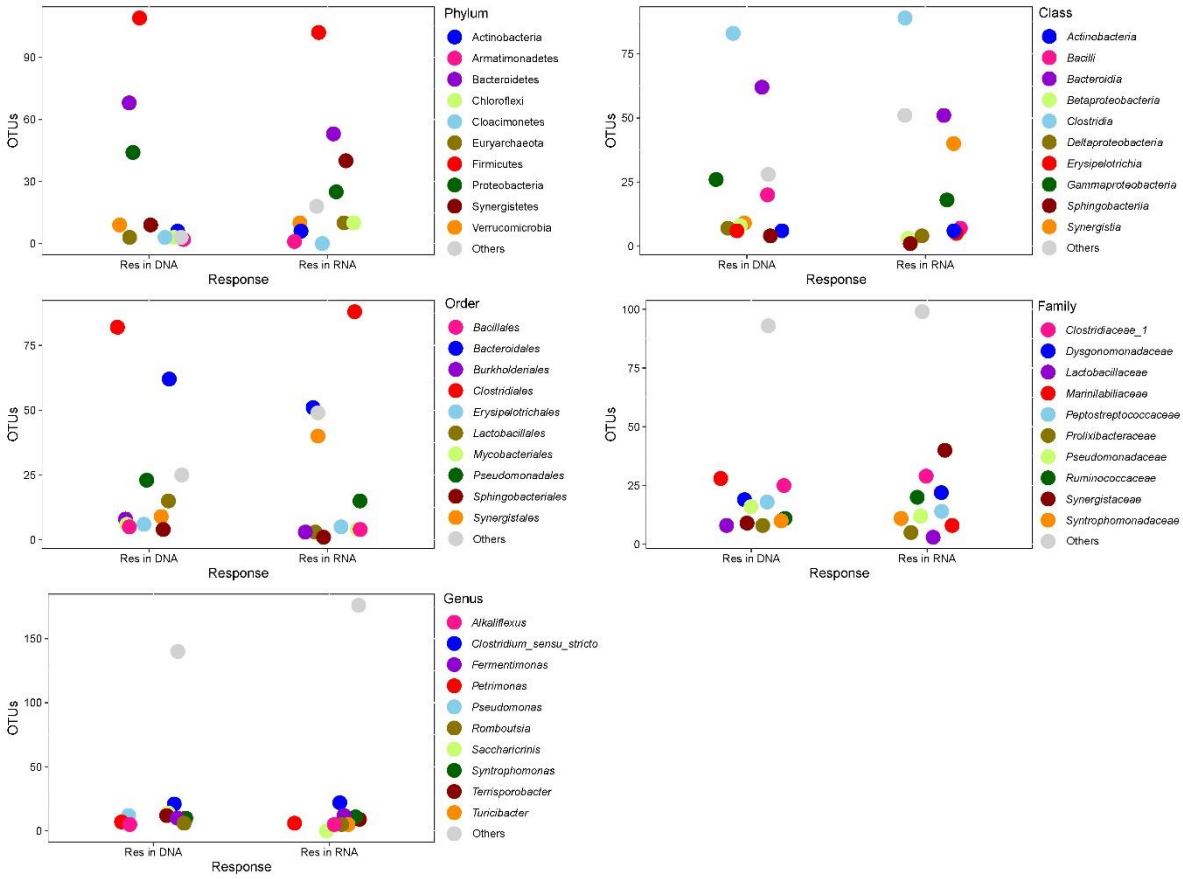
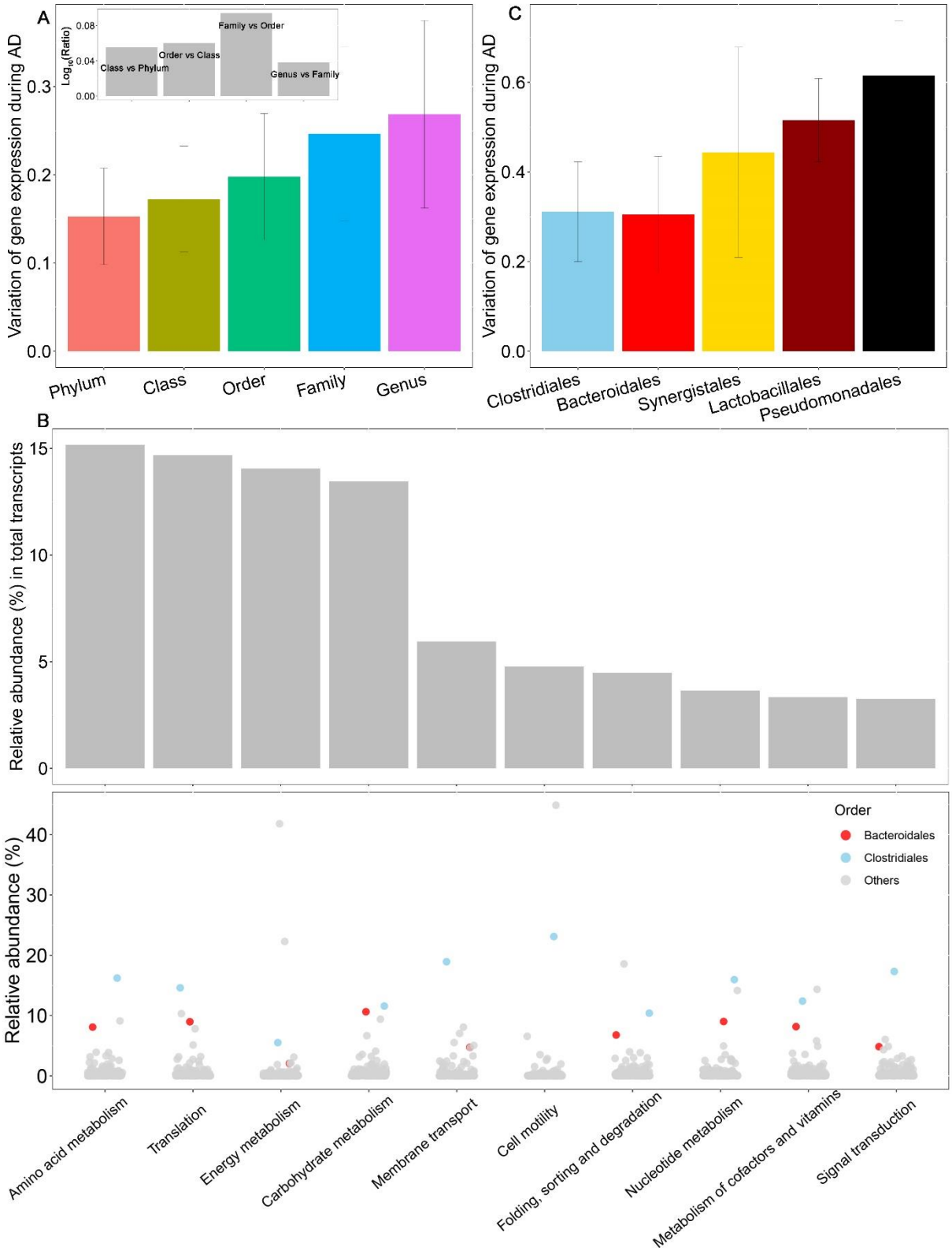

A
IR in DNA
IR in RNA
DR in DNA
DR in RNA
OTUs
Response
Res1
Res2
Res3
Res4
Phylum
Class
Order
Family
Genus
B
Heterogeneity
Phylum
Class
Order
Family
Genus

Fig. 3


Phylum
Actinobacteria
Armatimonadetes
Bacteroidetes
Chloroflexi
Cloacimonetes
Euryarchaeota
Firmicutes
Proteobacteria
Synergistetes
Verrucomicrobia
Others
Class
Actinobacteria
Bacilli
Bacteroidia
Betaproteobacteria
Clostridia
Deltaproteobacteria
Erysipelotrichia
Gammaproteobacteria
Sphingobacteriia
Synergistia
Others
Order
Bacillales
Bacteroidales
Burkholderiales
Clostridiales
Erysipelotrichales
Lactobacillales
Mycobacteriales
Pseudomonadales
Sphingobacteriales
Synergistales
Others
Family
Clostridiaceae_1
Dysgonomonadaceae
Lactobacillaceae
Marinilabiliaceae
Peptostreptococcaceae
Prolixibacteraceae
Pseudomonadaceae
Ruminococcaceae
Synergistaceae
Syntrophomonadaceae
Others
Genus
Alkaliflexus
Clostridium_sensu_stricto
Fermentimonas
Petrimonas
Pseudomonas
Romboutsia
Saccharicrinis
Syntrophomonas
Terrisporobacter
Turicibacter
Others
OTUs
Response
Res in DNA
Res in RNA

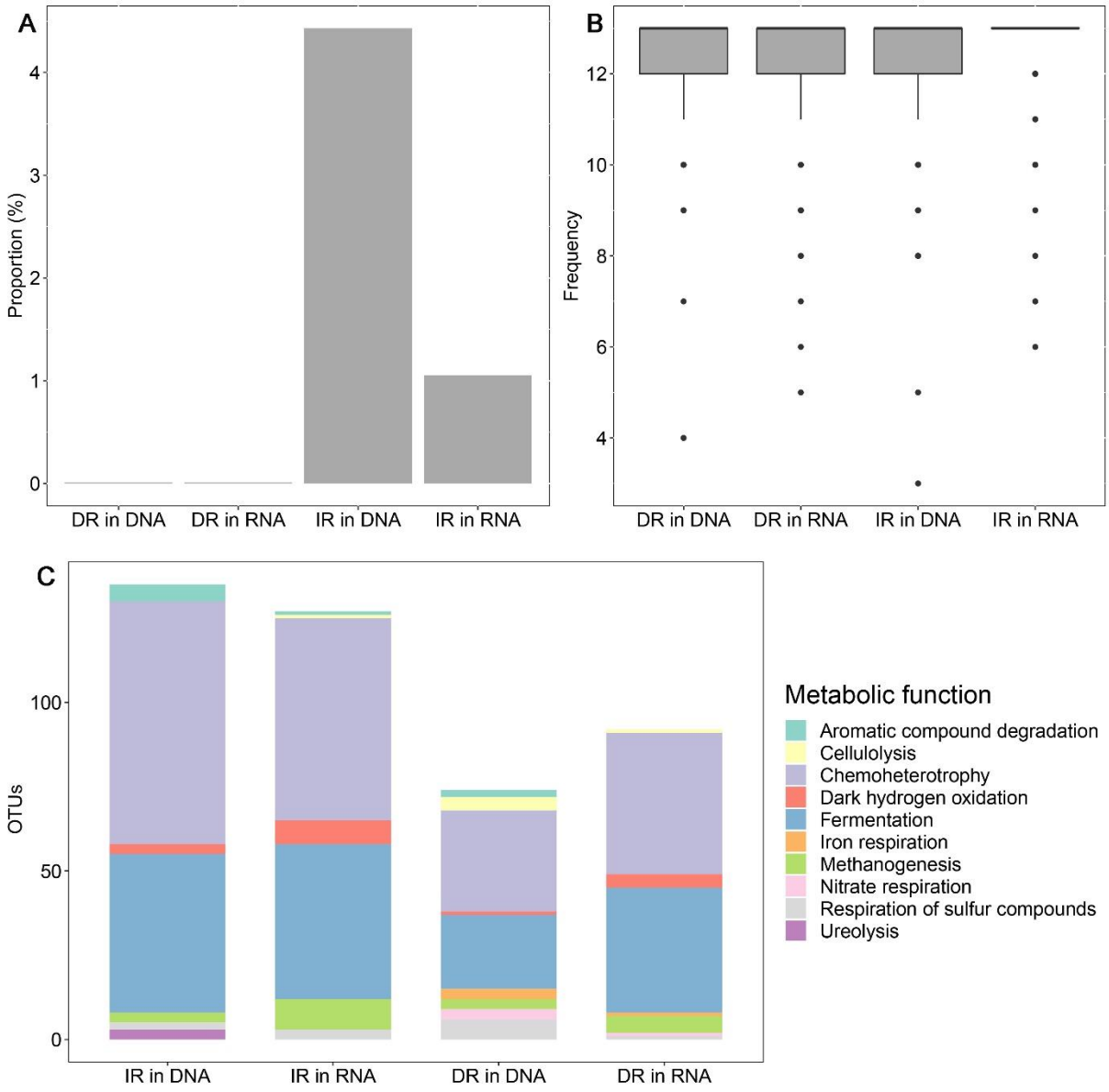
607

608 Fig. 4


A
Variation of gene expression during AD
0.0
0.1
0.2
0.3
Log10(Ratio)
0.00
0.04
0.08
Class vs Phylum
Order vs Class
Family vs Order
Genus vs Family
Phylum
Class
Order
Family
Genus
C
Variation of gene expression during AD
0.0
0.2
0.4
0.6
Clostridiales
Bacteroidales
Synergistales
Lactobacillales
Pseudomonadales
B
Relative abundance (%) in total transcripts
0
5
10
15
Relative abundance (%)
0
10
20
30
40
Order
Bacteroidales
Clostridiales
Others
Amino acid metabolism
Translation
Energy metabolism
Carbohydrate metabolism
Membrane transport
Cell motility
Folding, sorting and degradation
Nucleotide metabolism
Metabolism of cofactors and vitamins
Signal transduction

609

610 Fig. 5


A
Proportion (%)
0
1
2
3
4
DR in DNA
DR in RNA
IR in DNA
IR in RNA
B
Frequency
4
6
8
10
12
DR in DNA
DR in RNA
IR in DNA
IR in RNA
C
OTUs
0
50
100
IR in DNA
IR in RNA
DR in DNA
DR in RNA
Metabolic function
Aromatic compound degradation
Cellulolysis
Chemoheterotrophy
Dark hydrogen oxidation
Fermentation
Iron respiration
Methanogenesis
Nitrate respiration
Respiration of sulfur compounds
Ureolysis

611

612 Fig. 6