Comparing environmental microbiota of different growth stages of fresh water turtle Chinemys reevesii in pond and greenhouse cultivation

Aiguo Zhou
South China Agricultural University
Shaolin Xie
South China Agricultural University
Di Sun
South China Agricultural University
Pan Zhang
South China Agricultural University
Zhengkun Pan
South China Agricultural University
Zhiheng Zuo
South China Agricultural University
Xiang Li
Canadian Food Inspection Agency
Jixing Zou (zoujixing@scau.edu.cn)
South China Agricultural University https://orcid.org/0000-0003-4880-2689

Short report

Keywords: Chinemys reevesii, different growth stages, pond cultivation, greenhouse cultivation, environment microbiota

Posted Date: June 5th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-32288/v1

Abstract

The microbial community structure is an important indicator for evaluating the water quality of the aquaculture environment. In this study, V4 regions of 16S rRNA gene of pond (PC) and greenhouse cultured (GC) *C. reevesii* were sequenced. Results showed that a total of 1,993,090 high quality counts and 105,159 observed OUTs were obtained; and the Chao1 richness estimator of PC was significantly higher than that of GC groups. Beta-diversity showed that the microbiota of two groups were isolated from each other. In addition, the correlation analysis of environmental factors showed that NO2-N, PH, PO4-P, and stocking density played significant roles in the bacterial community composition. The dominant phyla in PC groups were cyanobacteria, proteobacteria, actinobacteria, bacteroidetes, verrucomicrobia, planctomycetes; and in GC groups were proteobacteria, bacteroidetes, firmicutes, cyanobacteria, chloroflexi, actinobacteria. The functional prediction showed that the top5 Picrust prediction gene functions were protein processing in endoplasmic reticulum, retinol metabolism, proteasome, glycan binding proteins, and stilbenoid, diarylheptanoid and gingerol biosynthesis. Meanwhile, the numbers and types of KEGG pathway annotations showed a significant difference between the two cultivation environments. The prediction of bacterial phenotype implied that the GC environment is more likely to deteriorate, and turtles are more susceptible to pathogens than those of PC environment. This is the first report to explore and understand the difference of microbiota characteristics between different cultivation environments in different growth stages of *C. reevesii*, which will provide basic data for water quality adjustment, disease prevention, and healthy breeding of turtle.

1. Introduction

There are about 220 species of turtle in the world, and 15% of them are distributed in China (Peng 2018). The breeding of *Chinemys reevesii* has been a growing trend because of its high-quality variety with edible and medicinal value (Dagoon 2000). The production of turtles is comprised of 50,000 tons per year only in China, and the highest yield is the *C. reevesii* (China Fisheries and Aquaculture, 2019). In recent years, the *C. reevesii* has become a fast-rising wave of breeding for more and more turtle farmers. At the same time, a series of breeding models have been established, and a large number of large-scale breeding farms have been formed, of which pond and greenhouse cultivation have become the mainstream and the largest breeding modes (Huo et al, 2017, 2020; Min et al, 2019). With the expansion of breeding density, the breeding environment has been under tremendous pressure, which has led to the deterioration of farming water quality, disease outbreaks, drug abuse, and has caused huge losses to the turtle industry as well threatened food quality and safety and human health.

As a variable-temperature animal, *C. reevesii* has strict requirements on the environment; hence, a good environment is extremely important for the growth, reproduction, and quality of cultured turtles (Li et al, 2018). However, the living environments of *C. reevesii* are quite different between pond and greenhouse farming (Li et al., 2000; He et al., 2013). The characteristics of the pond culture model are relatively low density, low mortality, and high ecology, but the changes of temperature is large between day and night or seasons, the controlability is poor, and the growth is slow, but a high quality of commodity turtle will be finally obtained. On the contrary, the cultivation mode of the greenhouse environment has an extremely high density, small temperature change, fast growth, and strong controllability, but the mortality rate is relatively high, the pollution is more serious, and the quality of commodity turtle is lower. The cultured environment is closely related to microbial structure, the microbial diversity index is closely related to the water quality of the breeding environment, and the water quality index is closely related to the health of the breeding turtles.

Hence, the Illumina Nova 6000 platform and PE250 sequencing were used to investigate the water microbial communities in different breeding periods between pond and greenhouse cultivation of *C. reevesii*. The ultimate goal is to provide great practical and theoretical significance for the health and ecological farming of *C. reevesii* in order to obtain high-quality and safe raw materials for food and medicine.

2. Materials And Methods

2.1. Sample collection and detection of water quality indexes

11 water samples of ponds (HTPC1-4, JTPC1-4 and ATPC1-3) and 12 water samples of greenhouses (HTGC1-4, JTGC1-4 and ATGC1-4) were collected from the farms in Huadu (N23.29, E113.04) and Zengcheng (N23.28, E113.75), Guangdong, China, the location and display map of different farming environments of *Chinemys reevesii* can be found in Supplementary material S1. 20L of each sample was packed in a plastic sealed barrel, recorded the information, and shipped back to the laboratory. Meanwhile, the Temperature (Temp ℃), PH, transparency (cm) and dissolved oxygen (DO mg/L) of each pond were measured in situ through portable water quality detector (HACH SL 1000, USA), the stocking density (SD ea/m2) were recorded as well. And the total ammonia nitrogen (NH4+-N, mg/L), nitrite nitrogen (NO2-N, mg/L), nitrate nitrogen NO3-N (mg/L) and phosphate (PO43-P, mg/L) were determined according to the standard methods of Jin & Tu (1990) within 24 h. Chlorophyll A(g/L) was determined photometrically after filtration with 0.45 μm membrane using the spectrophotometric method (Sartory and Grobbelaar 1984). The remaining water samples were filtered through 0.22 μm membranes (Sartorius stedimbiotect, Germany) within 24h, and then stored at -80 °C.

2.2. DNA Extraction, amplicon generation, library preparation and sequencing

Total genome DNA of 23 samples was extracted using the MOBIO Powerwater® DNA Isolation Kit (MOBIO Laboratories, Carlsbad, CA, USA), 1% agarose gel electrophoresis was used to determine the integrality of the total DNA, and the NanoDrop One (Thermo Fisher Scientific, MA, USA) was used to measure the concentration and purity of the total DNA.

V4 regions in 16S rRNA genes were amplified using specific primers 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACHVGGGT- WTCTAAT-3') (Caporaso et al., 2011) with 12 bp barcode. Primers were synthesized by Invitrogen (Invitrogen, Carlsbad, CA, USA). PCR reactions system including: 2x Premix Taq (Takara Biotechnology, Dalian Co. Ltd., China) 25 μl, each primer 1 μl (10 mM) and DNA (20 ng/μl) template 3 μl; then PCR reactions were performed by PCR instrument of BioRad S1000 (Bio-Rad Laboratory, CA,USA): 5 min at 94°C for initialization; 30 cycles of 30 s denaturation at 94°C, 30 s annealing at 52°C, and 30 s extension at 72°C; followed by 10 min final elongation at 72°C. The length and concentration of PCR products were determined by using 1% agarose gel electrophoresis. The PCR products were mixed in equidensity ratios according to the Gene Tools Analysis Software (Version 4.03.05.0, SynGene). Then, the mixture PCR products were purified by E.Z.N.A. Gel Extraction Kit (Omega, USA).

Sequencing libraries were generated using NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (NewEngland Biolabs, MA, USA) following manufacturer's recommendations and index codes were added. The library quality was determined by the Qubit@ 2.0 Fluorometer (Thermo Fisher Scientific, MA, USA). At last, the library was sequenced by an Illumina Nova6000 platform and 250 bp paired-end reads were generated according to the manufacturer's instruction.

2.3. 16S rRNA gene sequence analysis

Firstly, the quality of Raw Reads were controlled using Fastp (version 0.14.1, https://github.com/OpenGene/fastp) by sliding window, and the primers were removed by cut adapt software (version 1.9.2; https://github.com/marcelm/cutadapt/) in order to obtain the Paired-end Clean Reads. Then, the Paired-end Clean Reads were merged using USEARCH fastq mergepairs (V10, http://www.drive5.com/usearch/) according to the relationship of the overlap between the Paired-end Reads, and 5 bp maximum mismatch was allowed in an at least 16 bp overlap region in order to obtain Raw Tags. The clean tags were obtained by Raw Tags filtration using Trimmomatic v0.33 (Bolger et al, 2014). Then the chimera sequences were identified and removed using UCHIME v4.2 (Edgar et al. 2011) in order to obtain Effective Tags.

The operational taxonomic units (OTUs) were clustered with 97% similarity by Qiime 2 standardized process (Bolyen et al, 2019) and Open Reference OTU picking algorithm. Each representative sequence was compared with the Green Gene database for generation of Biom data including sample name, strain annotation and abundance information. The taxonomy results of species annotations were divided into 7 levels: kingdom (L1), phylum (L2), class (L3), order (L4), family (L5), genus (L6) and species (L7). And the contaminated OTUs were removed, and the effective Tag sequence number (No. of seqs) and OTU taxonomy comprehensive information table (OUT table) of each sample were obtained.

2.4. Statistical analysis

The readings and OTUs of each sample were counted, the Pan and Core of target classification level in different sample numbers were counted using R program (V5.1.3) (R Development Core Team, 2006). Alpha-diversity analyses including Observed Species index, Chao1 richness estimator, Shannon-Weiner index, Equitability index, Simpson diversity index, Good's Coverage index and Phylogenetic Diversity index; and Beta-diversity measurements, including Principal component analysis (PCA), Analysis of Similarities (Anosim) (Chapman and Underwood 1999), and nonmetric multidimensional scaling (NMDS) (Rabinowitz 1975) were performed to calculate the differences using R program ggplot2 package (Wickham 2016). The correlation analyses of environmental factors, Microbial community structure distribution, LDA Effect Size (LEfSe) and Random forest map of various levels of microbial were analyzed using R package software. Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) (Langille et al, 2013) was used to predict the microbial community function and its differences; BugBase (https://bugbase.cs.umn.edu/) was used to predict the bacterial phenotype; one way analysis of variance (ANOVA) were performed to analyze the differences between groups using SPSS 17.0 (SPSS Inc., USA). $P< 0.05$ was considered significant, while $P< 0.001$ was considered highly significant, all the values were the average of the different groups (mean ± SD).

3. Results And Discussion

3.1 Analyses of bacterial diversity

The bacterial 16S rRNA gene V4 regions of ponds cultured (PC) and greenhouse cultivation (GC) in hatchling, juvenile, and adult periods of *C. reevesii* were sequenced using the Illumina Nova6000 platform. A total of 1,993,090 high quality counts and an average of 86,656.09 counts were obtained from 23 samples belonging to six groups (HTPC1-4, JTPC1-4, ATPC1-3, HTGC1-4, JTGC1-4, and ATGC1-4). All sequences were clustered into 105,159 OUTs in total, and each sample showed different number of OTUs clustering information ranging from 3193 (HTGC4) to 6411 (ATPC2) at 97% similarity. The counts, observed OUTs, Chao1, Equitability, Shannon, Simpson, Good's coverage, Phylogenetic diversity (PD) whole tree statistical estimates of richness and diversity indexes from each sample are presented in Table 1. Comparison and analysis of alpha-diversity for different growth stages of freshwater turtle *C. reevesii* in PC and GC environments showed that the total and each group of the Chao1, Shannon, Simpson in PC groups were higher than those of the GC groups (Fig. 1A-C). These indicated that the bacterial abundance in pond environment was higher than that of greenhouse environment. Meanwhile, beta-diversity was used to calculate the distance between different groups. The principal component analysis (PCoA) showed that the different growth stages of the GC and PC groups could obviously gather together, but the two cultivation environments are isolated from each other, and the cumulative contribution rate of PCoA 1 (30.54%) and 2 (10.52%) with bray curtis, 18.6% and 5.7% of binary jaccard (Fig. 1D, E). Furthermore, the analyses of group differences showed that the Observed R value is 0.7713 ($P=0.0001$), and the distance of all within-groups ranged from 0.5211 to 0.6712, and all between-groups ranged from 0.5781 to 0.7977; the mean maximum and minimum group distances between each other were 0.7868 for HTGC vs. HTPC and 0.6026 for ATPC vs. HTPC, this indicated that there are obvious differences between the groups (Fig. 1F, H). The NMDS analysis showed that the microbiota of six groups was clearly separated into six clusters, especially for PC and GC groups (Fig. 1G). These results indicated that the microbiota communities of different growth stages, especially in different cultivation environments obvious differences, and the similarity of microbiota types in PC groups was higher than those of the GC groups, and the environment of the PC groups was relatively stable, but the environment of the GC groups was greatly affected by the external environment (Pitacco et al., 2019; Lazzari et al., 2020). The raw data of alpha-diversity, the group distance data and PCoA (1, 3) and (2, 3) map can be found in Supplementary material S2.

Table 1 Richness and diversity indexes relative to each sample (OTUs were defined at the 97% similarity level (Threshold is 0.03)).

Sample ID	Counts	Alphadiversity						
		Observed OUTs	Chao1	Equitability	Shannon	Simpson	Goods coverage	PD whole tree
HTPC1	64,997	4,825	11,095.6503	0.65086124	7.964142	0.9648892	0.92931319	398.22413
HTPC2	68,075	5,884	13,892.7109	0.72650334	9.09769732	0.98811813	0.91357092	458.74608
HTPC3	61,492	5,094	10,946.0206	0.66704508	8.21438221	0.97094908	0.92669342	418.07747
HTPC4	91,084	3,779	9,148.1476	0.56251041	6.6847549	0.93072839	0.94304933	338.30432
JTPC1	88,505	4,507	10,507.5396	0.59343605	7.20309818	0.95088012	0.93311305	393.27832
JTPC2	75,770	3,969	9,333.80887	0.5790958	6.92283544	0.9482456	0.94080717	354.25884
JTPC3	74,256	5,068	11,394.1533	0.6663094	8.20040355	0.97460536	0.92570215	417.67826
JTPC4	81,081	3,264	7,848.80331	0.50961591	5.94845363	0.91409333	0.95031862	310.18648
ATPC1	57,370	6,351	14,113.0676	0.74705529	9.4373762	0.99139504	0.90811895	487.5006
ATPC2	67,386	6,411	14,508.5938	0.7503018	9.48856696	0.99136054	0.90644324	512.41029
ATPC3	68,135	6,240	13,547.0426	0.74615905	9.40707362	0.99027745	0.91149398	496.93705
HTGC1	11,5451	4,327	9,133.16981	0.6327939	7.64361337	0.96224922	0.93924947	373.40055
HTGC2	89,853	3,496	7,863.80534	0.56062139	6.59934881	0.92033243	0.94949257	327.23926
HTGC3	88,092	3,601	8,226.01093	0.49108532	5.80177125	0.83141427	0.94680198	329.16926
HTGC4	10,7919	3,193	7,111	0.55434821	6.45299941	0.95083458	0.95374085	281.35905
JTGC1	83,718	3,699	8,325.71587	0.62528623	7.41146738	0.97184881	0.94713241	323.67024
JTGC2	108,954	3,940	8,986.31754	0.62914851	7.51453719	0.96952937	0.94337975	354.61208
JTGC3	99,788	3,933	8,727.00998	0.6344613	7.5763653	0.97553313	0.94333255	344.44308
JTGC4	93,833	5,140	11,155.2791	0.70375189	8.67553846	0.98252414	0.9279679	406.05474
ATGC1	115,298	3,657	9,607.0592	0.45725022	5.41221709	0.75224964	0.94399339	364.64287
ATGC2	99,145	4,728	10,791.9607	0.65504928	7.99619596	0.96793801	0.93186217	421.06455
ATGC3	82,261	4,925	10,803.0905	0.70347102	8.62871079	0.98902736	0.9308945	405.30658
ATGC4	110,627	5,128	11,262.0013	0.68522241	8.44480472	0.98016172	0.92763748	407.46934

Note: HTPC: hatchling turtle of pond cultivation; JTPC: juvenile turtle of pond cultivation; ATPC: adult turtle of pond cultivation; HTGC: hatchling turtle of greenhouse cultivation; JTGC: juvenile turtle of greenhouse cultivation; ATGC: adult turtle of greenhouse cultivation.

3.2. Correlation analysis of environmental factors

Water quality indexes are important for evaluating the breeding environment and also closely related to microbial communities (Zhang et al., 2016; Liao et al., 2020). The stocking density also directly affects water quality (Person-Le Ruyet et al., 2008; Liu et al., 2014). We found that the environmental factors with large differences are concentrated in pH, transparency, DO, NH3-N, PO4-P, chlorophyll A, and stocking density. By comparing two different farming breeding environments, we can see that the stocking density, DO, NH3-N, and PO4-P of the PC environment are significantly less than the GC environment, while the temperature, pH, transparency are significant higher than those of GC environment (Fig. 2A). These implied that the PC environment is more natural and ecological than the GC environment. The relationships between eight environmental factors, 11 samples, and top10

microbiota were determined by redundancy analysis (RDA). The correlation analyses show that the amount of explanation for distribution of species by eight environmental factors were 98.79% (PC) and 86.57% (GC). A total of 64.63% (PC) and 51.74% (GC) of the cumulative variances of the microbiota-environment relationship were represented by the first two axes, and the different environmental factors affected the different growth stages samples and top10 microbiota with different degrees in both PC and GC environments. In the PC environment, the plot demonstrated that NO2-N ($r=0.599$, $P<0.01$) played significant roles in the bacterial community composition (Fig. 2B). In the GC environment, pH ($r=0.501$, $P<0.05$), PO4-P ($r=0.505$, $P<0.05$), and stocking density ($r=0.534$, $P<0.05$) played significant roles in the bacterial community composition (Fig. 2C). These indicated that the environmental factors changeswith the change of the breeding cycle, which affects the growth and living environment of the turtle (Wang et al., 2019; Li et al., 2020). These also explain that the microbial composition and abundance of water are closely related to water quality indicators (Niu et al., 2019; Liu et al., 2019).

3.3. Analysis of microbiota structure

The microbiota of 23 samples (PC 11 and GC 12) was counted. The shared top10 phyla in both PC and GC environments were proteobacteria (28.7%), cyanobacteria (22.4%), bacteroidetes (15.1%), actinobacteria (8.1%), firmicutes (6.9%), chloroflexi (2.9%), verrucomicrobia (2.9%), planctomycetes (2.6%), chlorobi (1.6%), actinobacteria (1.1%), and WWE1 (1.1%); meanwhile, two archaea (euryarchaeota 0.3% and crenarchaeota 0.1%) were also found in two breeding environments (Fig. 3A). Furthermore, both the PC and GC models have 63 different prokaryotic phyla and 51 central intersection strains from the 16S rRNA gene sequences. Surprisingly, all the phyla of HTPC groups belonged to JTPC and ATPC groups (Fig. 3B, C). In addition, the top one phyla in PC groups was cyanobacteria (35.7%), and in GC groups was proteobacteria (35.4%). These may be closely related to the cyanobacteria being the dominant algae, which also explain that the oxygen content of the pond is low, resulting in the large multiplication of aerobic bacteria (Arbib et al., 2017). Conversely, the proteobacteria is the main source of multiple pathogens with a higher pathogenic potential in the GC environment. Furthermore, the genus levels of PC and GC groups showed that the microbiota of different breeding periods and different breeding modes differed, especially for the GC groups (Fig. 3D, E). These indicate that microbiota structures closely related to the breeding environment, feeding method, and management measures (Crippen et al., 2019; Li et al., 2019; Noel et al., 2019). There were 626 (PC) and 649 (GC) OTUs shared among PC and GC groups, representing 26.6% and 24.4% of total reads separately in PC and GC environments, respectively. In addition, the quantity of OTUs were sorted from ATGC (929) >JTGC (878) >HTGC (848) >ATPC (808) >HTPC (777) >JTPC (771) (Fig. 3F). The other taxonomy results of Venn diagrams were provided in the Supplementary material S3.

3.4. Annotation analysis of the microbiota

In order to visualize the annotation results, the KRONA software was used to reveal the average relative abundance of each bacterium in PC (Fig. 4A-C) and GC groups (Fig. 4E-F). The results showed that the most abundant bacterium was p_Cyanobacteria.c_Nostocophycideae.o_Nostocales.f_Nostocaceae.g_*Cylindrosper-mopsis* of HTPC (11%); p_Cyanobacteria.c_Oscillatoriophycideae. o_Chroococcales. f_Microcystaceae. g_*Microcystis* of JTPC (7%); p_Cyanobacteria. c_Synechococcop- hycideae.o_Synechococcales.f_Synechococcaceae.g_*Synechococcus* of ATPC (11%); p_Bacteroidetes.c_Bacteroidia.o_Bacteroidales.f_Porphyromonadaceae.g_*Paludibacter* of HTGC (9%); p_Cyanobacteria.c_Oscillatoriophycideae.o_Chroococcales. f_Microcystaceae.g_*Microcystis* of JTGC (20%); p_Proteobacteria. c_Epsilonproteo- bacteria.o_Campylobacterales.f_Campylobacteraceae.g_*Arcobacter* of ATGC (23%). We can see that the annotated bacteria of different classifications in different growth stages of different cultured environments have significant differences, which are extremely consistent with the actual breeding situation (Castañeda-Monsalve et al., 2019; Wang et al., 2019). Meanwhile, the Archaea in HTPC (0.1%), JTPC (0.06%), and ATPC (0.06%) as well as euryarchaeota in HTGC (0.8%) and ATGC (0.3) and the archaea in JTGC (0.6%) were also found by using KRONA software annotation (Fig. 4A-F). The results showed that the richness of archaea in the GC environment was far greater than in the PC environment. These might be closely related to the water source and the length of aquaculture cycle (Zhong et al., 2016; Lu et al., 2016, 2019; Baker e al., 2020). In addition, a total of 326 biomarkers with statistical differences were obtained in six groups by using LDA Effect Size (LEfSe) analysis, and the LDA values ranged from 3.3745 to 5.5156, and the relative abundances of microbiota in different growth stages of PC and GC environments showed significant differences (Fig. 4G). The mean decrease accuracy (MDA) of random forest model was used to analyze the significant of classification meaning based on the strains with MDA> 3 (Breiman 2001b,

Iverson et al., 2008). The results showed that the acidobacteria, gemmatimonadetes, proteobacteria, OD1, chlorobi, FCPU426, chlamydiae, NKB19, TM6 of the PC groups (Fig. 9H-J) and the parvarchaeota, OP8, Caldithrix, WWE1, WS1, OP11, spirochaetes, verrucomicrobia, planctomycetes, chlorobi, acidobacteria, KSB3, tenericutes, and caldiserica of the GC groups (Fig. 9K-M) have the significant meaning for classification at the phylum level. The LDA values of microbiota communities, digital and taxonomy results of random forest model are provided in the Supplementary material S4.

3.5. Functional prediction of the microbiota

3.5.1. Analyses of Picrust gene function prediction expression

The Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) was used to analyze the microbial species composition information based on the 16S rRNA gene sequences (Langille et al., 2013), the composition of functional gene, and functional differences of different samples in PC and GC environments. Meanwhile, the prediction genes were compared with the COG database, and the results showed that a total of 187 gene functions were predicted, including Cellular Processes (8), Environmental Information Processing (9), Genetic Information Processing (22), Human Diseases (7), Metabolism (113), Organismal Systems (5) and Unclassified (23) (Fig. 5A). As we all know, turtles are fed a large amount of bait during their breeding process, especially the greenhouse cultivation, which greatly accelerates the content of organic matter in the water, and greatly promotes the metabolic function of the water microorganisms (Zhu et al., 2016; Song et al., 2019). In addition, the top10 with the smallest p values were obtained according to the information of Picrust prediction gene function using a non-parametric test, including protein processing in endoplasmic reticulum (p=0.001232), retinol metabolism (p=0.001276), proteasome (p=0.001347), glycan bindng proteins (p=0.001523), stilbenoid, diarylheptanoid and gingerol biosynthesis (p=0.001638), fluorobenzoate degradation (p=0.001691), riboflavin metabolism (p=0.001783), circadian rhythm-plant (p=0.001843), isoflavonoid biosynthesis (p=0.001895), glycosphingolipid biosynthesis lacto, and neolacto series (p=0.001907). Meanwhile, the gene function expressions prediction of the HTGC, ATPC, and JTPC groups showed relatively higher among the six groups based on top10 Picrust prediction gene functions (Fig. 10B-K). These are closely related to the environment of pond farming with a easy to break out of cyanobacteria (Gibbs 1981; Schwarz and Blower 2016). In addition, the both the highest proportion of function annotations in the PC and GC environments were metabolism (51.2% and 57.6%), and the highest proportion of specific functions were concentrated in transporters (5% and 5.3%). These indicated that the prediction functions of the PC and GC environments were similar, but there were differences between the overall levels. The identified bacteria that also related Human Diseases (3.74%) imply potential human pathogens and a certain threat to public health (Nowakiewicz et al., 2015). The gene function prediction results and raw data of each sample are available in the Supplementary material S5.

3.5.2. KEGG pathway annotation and bacterial phenotype prediction

The Picrust gene function prediction results were used to enrich the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation information, and the KEGG pathway annotation heat map with $P<0.05$ was selected according to Kruskal test (Fig. 6A). The original data of heat maps have been standardized and normalized (heat map data=original data-mean/standard deviation). A total of 170 KEGG pathways and 48 significant KEGG pathways were annotated in six groups. There are 15 KEGG pathways showed high relative abundance in the GC groups, especially in the HTGC and JTGC groups, while 31 KEGG pathways showed high relative abundance in the PC groups, especially in the HTPC and JTPC groups (Fig. 6A). It can be seen that the signaling pathways of bacteria in the external pond environment are more abundant than those of greenhouse, which may be closely related to sunlight, rainwater, airflow, and soil geology (De Souza et al., 2001; Brönmark and Hansson 2017). In addition, these results may also be closely related to the difference of growth cycle, stocking density, water quality, feeding conditions, and environments of *C. reevesii* (Hoang et al., 2018; Deng et al., 2019). The gene function prediction and raw data of KEGG pathways for the PC and GC groups were provided in the Supplementary material S6.

The bacterial phenotype of all six groups were predicted by using online tool of BugBase (https://bugbase.cs.umn.edu/) based on the OTUs and mapping files, and the relative abundance of nine phenotype (Gram Positive, Gram Negative, Biofilm Forming, Pathogenic Potential, Mobile Elements, Aerobic, Anaerobic, Facultative anaerobic and Oxidative Stress Tolerant) were

investigated in this study (Reimer et al., 2019). We can see that the highest and lowest relative abundance of Gram Positive (mean) were ATPC and ATGC groups, while the Gram Negative showed the opposite result (Fig. 6B, C). The highest relative abundances of biofilm forming, mobile elements, and aerobic were from the ATPC group (Fig. 6D, F, G). Both the highest relative abundances of Pathogenic Potential and Anaerobic were JTGC groups (Fig. 6E, H). Meanwhile, the highest relative abundances of facultative anaerobic and oxidative stress tolerant were ATGC and JTPC groups (Fig. 6I, J). In addition, the lowest relative abundances of biofilm forming, facultative anaerobic and oxidative stress tolerant were from the HTGC group (Fig. 6D, I, J). The lowest relative abundances of pathogenic potential, mobile elements and anaerobic were from the JTPC group (Fig. 6E, F, H). Finally, the lowest relative abundances of aerobic were from the JTGC group (Fig. 6G). These indicated that the relative abundances of Gram positive and aerobic in pond environment were higher than those of greenhouse environment, as well a relatively high oxidative stress tolerant and which was good for the formation of biofilm. Meanwhile, the relative abundances of anaerobic and facultative anaerobic in the greenhouse were significant higher than those of the pond environment, and an extremely high phenotype of pathogenic potential, especially in JTGC group. These might be related to the enclosed environment, relatively constant temperature, and high breeding density (Zhang et al., 2018; Mobeen et al., 2019). These results also implied that the GC environment is more likely to deteriorate, and turtles are more susceptible to pathogens, while the breeding environment of the pond is more stable and ecological. The statistical test results of BugBase bacterial phenotype prediction can be found in Supplementary material S7.

4. Conclusions

This study confirmed that the microbiota characteristics of different growth stages in pond and greenhouse environments showed differences in phylum and genus levels, and a certain percentage of archaea was also identified in different growth stages of PC and GC environments. The results imply that the water quality of the greenhouse environment is more likely to deteriorate, and the cultured turtles are more susceptible to pathogens, while the pond environment is more stable and ecological. These conclusions may be closely related to the unity and controllability of the greenhouse cultivation environment, and we suggest to carrying out ecological breeding methods and developing better tail water treatment technology in order to improve the culture environment as well as safe aquatic products. Our study will also provide basic data for water quality adjustment, disease prevention and control, and healthy breeding of various turtle species.

Declarations

- Ethics approval and consent to participate (Not applicable)
- Consent for publication (Not applicable)
- Availability of data and material

Please contact author for data requests.

- Competing interests

The authors declare that they have no competing interests.

- Funding

This work was supported by China Scholarship Council (No. 201908440582), China Agriculture Research System (CARS-45-49) and Talent introduction special funds of South China Agricultural University (K16226); and all of them are contribute to the design of the study and collection, analysis, the interpretation of data, and the writing of the manuscript.

- Authors' contributions

AZ and SX carried out the environmental microbiology studies, participated in the sequencing, data analysis and drafted the manuscript. DS, PZ, ZP carried out the samples collection, filter, and DNA extraction. ZZ participated in the statistical analysis.

JZ and XL conceived of the study, and participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

- Acknowledgements

We express our gratitude to Tianbao Jiang, Xizhao Jiang, and Zhenlong Chen of these two farms for their assistance in collecting samples. We also express our appreciation to our anonymous reviewers for providing valuable comments on the manuscript.

References

Ahmed, H., Okba, A., 2017. Determination of Anabaena oryzae Survival by Chlorophyll-a and Carotenoid under Using Treated Sewage Water. Journal of Agricultural Chemistry and Biotechnology, 8(5), 157-161.

Arbib, Z., de Godos Crespo, I., Corona, E. L., Rogalla, F., 2017. Understanding the biological activity of high rate algae ponds through the calculation of oxygen balances. Applied microbiology and biotechnology, 101(12), 5189-5198.

Baker, B. J., De Anda, V., Seitz, K. W., Dombrowski, N., Santoro, A. E., Lloyd, K. G., 2020. Diversity, ecology and evolution of Archaea. Nature Microbiology, 1-14.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114-2120.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G. A., ... Bai, Y., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature biotechnology, 37(8), 852-857.

Brönmark, C., Hansson, L.A., 2017. The biology of lakes and ponds. Oxford University Press.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., Knight, R., 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the national academy of sciences, 108 (Supplement 1), 4516-4522.

Castañeda-Monsalve, V.A., Junca, H., García-Bonilla, E., Montoya-Campuzano, O.I., Moreno-Herrera, C.X., 2019. Characterization of the gastrointestinal bacterial microbiome of farmed juvenile and adult white Cachama (*Piaractus brachypomus*). Aquaculture, 512, 734325.

Chapman, M.G., Underwood, A.J., 1999. Ecological patterns in multivariate assemblages: information and interpretation of negative values in ANOSIM tests. Marine ecology progress series, 180, 257-265.

Crippen, T.L., Sheffield, C.L., Singh, B., Byrd, J.A., Beier, R.C., 2019. How Management Practices Within a Poultry House During Successive Flock Rotations Change the Structure of the Soil Microbiome. Frontiers in microbiology, 10, 2100.

Dagoon, N.J., 2000. Malaysian school engages in bullfrog and turtle farming. SEAFDEC Asian Aquaculture, 22(3), 16-19.

De Souza, M.P., Amini, A., Dojka, M.A., Pickering, I.J., Dawson, S.C., Pace, N.R., Terry, N., 2001. Identification and characterization of bacteria in a selenium-contaminated hypersaline evaporation pond. Appl. Environ. Microbiol., 67(9), 3785-3794.

Deng, Y., Xu, X., Yin, X., Lu, H., Chen, G., Yu, J., Ruan, Y., 2019. Effect of stock density on the microbial community in biofloc water and Pacific white shrimp (*Litopenaeus vannamei*) gut microbiota. Applied microbiology and biotechnology, 103(10), 4241-4252.

Edgar,R.C., 2019. highly accurate OTU sequences from microbial amplicon reads. Nature methods 2019 10(10): 996

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics, 27(16), 2194-2200.

He, R., Xie, J., Huang, S.L., Shi, J.B., Gao, Z.L., Xiong, Q., 2013. Comparison of nutritional composition of Chinese soft-shelled turtles (*Pelodiscus sinensis*) grown in greenhouse and imitative ecological farming conditions. Food Science, 34(13), 234-238.

Gibbs, S.P., 1981. The chloroplast endoplasmic reticulum: structure, function, and evolutionary significance. Int Rev Cytol, 72, 49-99.

Hoang, M.N., Nguyen, P.N., Le, D.V., Nguyen, D.V., Bossier, P., 2018. Effects of stocking density of gray mullet Mugil cephalus on water quality, growth performance, nutrient conversion rate, and microbial community structure in the white shrimp *Litopenaeus vannamei* integrated system. Aquaculture, 496, 123-133.

Huo, J., Dong, A., Yan, J., Dong, A., 2020. Effects of cadmium on the gene transcription of the liver in the freshwater turtle (*Chinemys reevesii*). Environmental Science and Pollution Research, 1-8.

Huo, J., Dong, A., Yan, J., Wang, L., Ma, C., Lee, S., 2017. Cadmium toxicokinetics in the freshwater turtle, *Chinemys reevesii*. Chemosphere, 182, 392-398.

Jin, X.S., Tu, Q.Y., 1990. The investigation standard of lake eutrophication. Environmental Science Press, Beijing, (in Chinese)

Langille, M.G., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R., Beiko, R.G., 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature biotechnology, 31(9), 814.

Li, F., Li, C., Chen, Y., Liu, J., Zhang, C., Irving, B., Fitzsimmons, C., Plastow, G., Plastow, G., 2019. Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. Microbiome, 7(1), 92.

Li, T., Han, X.K., Chen, W., Zhou, H.B., Zhang, Y.P., 2018. Phenotypic Consequences of Embryonic Responses to Developmental Temperatures in Two Latitudinally Separated Populations of Asian Yellow Pond Turtles. Journal of Herpetology, 52(4), 453-457.

Li, X., Yang, J., Qi, D., Huang, Z., Yang, H., 2020. Correlation study of Submerged Macrophytes Growth and Environmental Factors in Lake Qionghai Wetland. E&ES, 440(5), 052005.

Li, Y.S., 2000. Experiment of Industrial Culture of *Chinemys reevesii*. Shandong Fisheries, (4), 6.

Lazzari, N., Martín-López, B., Sanabria-Fernandez, J.A., Becerro, M.A., 2020. Alpha and beta diversity across coastal marine social-ecological systems: Implications for conservation. Ecological Indicators, 109, 105786.

Liao, H., Yen, J. Y., Guan, Y., Ke, D., Liu, C., 2020. Differential responses of stream water and bed sediment microbial communities to watershed degradation. Environment international, 134, 105198.

Liu, D., Jin, J., Liang, S., Zhang, J., 2019. Characteristics of water quality and bacterial communities in three water supply pipelines. RSC advances, 9(7), 4035-4047.

Liu, G., Li, L., Peng, G., Li, J., Yan, W., Tang, J., 2014. Effects of stocking density on growth of *Procambarus clarkii* and aquaculture water quality. Acta Agricultural Jiangxi, 26(4), 86-93.

Lu, S., Liu, X., Liu, C., Wang, X., & Cheng, G. (2019). Review of ammonia-oxidizing bacteria and archaea in freshwater ponds. Reviews in Environmental Science and Bio/Technology, 18(1), 1-10.

Lu, S., Liu, X., Ma, Z., Liu, Q., Wu, Z., Zeng, X., Xu, S., Gu, Z., 2016. Vertical segregation and phylogenetic characterization of ammonia-oxidizing bacteria and archaea in the sediment of a freshwater aquaculture pond. Frontiers in microbiology, 6, 1539.

Min, T., Jian-Ye, Y., Jie, R., Shun-Xiang, L., 2019. Differentiating Trachemys scripta elegans Shell Glue from *Chinemys reevesii* Shell Glue by UPLC-QTOF/MS Coupled with Binary Compare Tool of UNIFI. Digital Chinese Medicine, 2(1), 50-58.

Mobeen, F., Sharma, V., Prakash, T., 2019. Functional signature analysis of extreme Prakriti endophenotypes in gut microbiome of western Indian rural population. Bioinformation, 15(7), 490.

Niu, A., Song, L.Y., Xiong, Y.H., Lu, C. J., Junaid, M., Pei, D.S., 2019. Impact of water quality on the microbial diversity in the surface water along the Three Gorge Reservoir (TGR), China. Ecotoxicology and Environmental Safety, 181, 412-418.

Noel, S. J., Olijhoek, D. W., Mclean, F., Løvendahl, P., Lund, P., Højberg, O., 2019. Rumen and Fecal Microbial Community Structure of Holstein and Jersey Dairy Cows as Affected by Breed, Diet, and Residual Feed Intake. Animals, 9(8), 498.

Nowakiewicz, A., Ziółkowska, G., Zięba, P., Dziedzic, B. M., Gnat, S., Wójcik, M., Dziedzic, R., Kostruba, A,. 2015. Aerobic bacterial microbiota isolated from the cloaca of the European pond turtle (*Emys orbicularis*) in Poland. Journal of wildlife diseases, 51(1), 255-259.

Peng, Y.R., 2018. The development history and training algorithm of artificial neuralnetwork. Public Communication of Science & Technology, 10(21): 129-130.

Person-Le Ruyet, J., Labbé, L., Le Bayon, N., Sévère, A., Le Roux, A., Le Delliou, H., Quéméner, L., 2008. Combined effects of water quality and stocking density on welfare and growth of rainbow trout (*Oncorhynchus mykiss*). Aquatic Living Resources, 21(2), 185-195.

Pitacco, V., Mistri, M., Aleffi, I.F., Lardicci, C., Prato, S., Tagliapietra, D., Munari, C., 2019. Spatial patterns of macrobenthic alpha and beta diversity at different scales in Italian transitional waters (central Mediterranean). Estuarine, Coastal and Shelf Science, 222, 126-138.

Rabinowitz, G.B., 1975. An introduction to nonmetric multidimensional scaling. American Journal of Political Science, 343-390.

Reimer, L.C., Vetcininova, A., Carbasse, J.S., Söhngen, C., Gleim, D., Ebeling, C., Overmann, J., 2019. Bac Dive in 2019: bacterial phenotypic data for High-throughput biodiversity analysis. Nucleic acids research, 47(D1), D631-D636.

Sartory, D.P., Grobbelaar, J.U., 1984. Extraction of chlorophyll a from freshwater phytoplankton for spectrophotometric analysis. Hydrobiologia, 114(3), 177-187.

Schwarz, D.S., Blower, M.D., 2016. The endoplasmic reticulum: structure, function and response to cellular signaling. Cellular and Molecular Life Sciences, 73(1), 79-94.

Song, Y., Wang, Y., Mao, G., Gao, G., Wang, Y., 2019. Impact of planktonic low nucleic acid-content bacteria to bacterial community structure and associated ecological functions in a shallow lake. Science of The Total Environment, 658, 868-878.

Wang, C., Zhou, Y., Lv, D., Ge, Y., Li, H., You, Y., 2019. Change in the intestinal bacterial community structure associated with environmental microorganisms during the growth of Eriocheir sinensis. MicrobiologyOpen, 8(5), e00727.

Wang, S., Gao, Y., Li, Q., Gao, J., Zhai, S., Zhou, Y., Cheng, Y., 2019. Long-term and inter-monthly dynamics of aquatic vegetation and its relation with environmental factors in Taihu Lake, China. Science of The Total Environment, 651, 367-380.

Wickham, H., 2016. Programming with ggplot2. In ggplot2 (pp. 241-253). Springer, Cham.

Xu, L., Wu, F., Guo, Y., Gao, H., 2019. China fishery statistical yearbook.

Zhang, H., Sun, Z., Liu, B., Xuan, Y., Jiang, M., Pan, Y., Zhang, Y., Gong, Y., Lu, X., Yu, D., Kumar, D., Hu, X., Cao, G., Xue, R., Gong, C., 2016. Dynamic changes of microbial communities in *Litopenaeus vannamei* cultures and the effects of environmental factors. Aquaculture, 455, 97-108.

Zhang, Z.S., Huang, Y.F., 1995. Methods for the plankton studyin freshwater. Science Press, Beijing. (in Chinese)

Zhang, W., Jiao, L., Liu, R., Zhang, Y., Ji, Q., Zhang, H., Gao, X., Ma, Y., Shi, H.N., 2018. The effect of exposure to high altitude and low oxygen on intestinal microbial communities in mice. PloS one, 13(9), e0203701.

Zhong, W., Bian, B., Gao, N., Min, J., Shi, W., Lin, X., Shen, W., 2016. Nitrogen fertilization induced changes in ammonia oxidation are attributable mostly to bacteria rather than archaea in greenhouse-based high N input vegetable soil. Soil Biology and Biochemistry, 93, 150-159.

Zhu, J., Dai, W., Qiu, Q., Dong, C., Zhang, J., Xiong, J., 2016. Contrasting ecological processes and functional compositions between intestinal bacterial community in healthy and diseased shrimp. Microbial ecology, 72(4), 975-985.

Figures

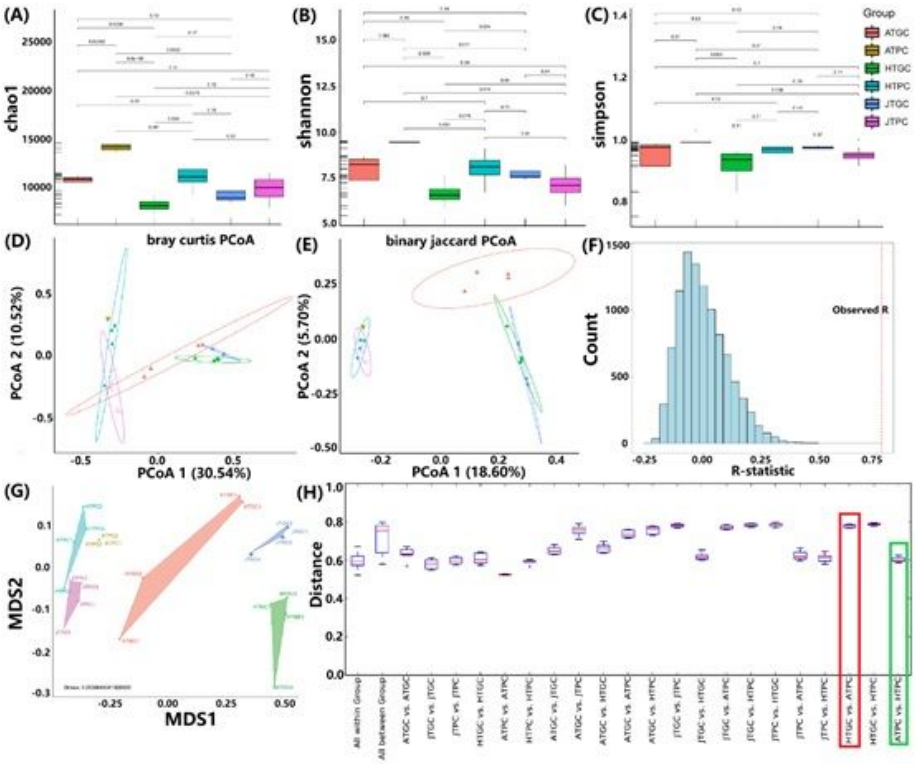

Figure 1

Alpha- and beta-diversity, Anosim group similarity, NDMS, and group distance analyses of different growth stages of freshwater turtle C. reevesii in six groups.

(A)

Samples	HTPC	JTPC	ATPC	HTGC	JTGC	ATGC
Temp (°C)	33.92^{ABd}±0.10	33.85^{Ad}±0.06	34^{Bd}±0.0	31.65^{Bc}±0.44	30.78^{Ab}±0.5	30.08^{Aa}±0.1
PH	9.08^{ABcd}±0.17	9.48^{Bd}±0.47	8.57^{Abc}±0.06	8.12^{Bb}±0.08	8.1^{Bb}±0.16	7.28^{Aa}±0.17
Transparency (cm)	21^{Cc}±1.41	4.75^{Ab}±0.29	14.67^{Bc}±0.58	2.5^{Aa}±0.58	2.5^{Aa}±0.56	18^{Bd}±0.82
DO (mg/L)	4.28^{Bc}±0.17	1.98^{Aa}±0.39	2.5^{Ab}±0.1	7.34^{Bc}±0.06	7.48^{Bc}±0.05	6.43^{Ad}±0.26
NH3-N (mg/L)	0.25^{Aa}±0.12	0.75^{Ba}±0.26	0.46^{ABa}±0.25	23.14^{ABab}±20.5	40.06^{b}±19.9	4.08^{Aa}±0.26
NO2-N (mg/L)	0.03^{Aa}±0.03	0.11^{Aa}±0.09	0.7^{Bb}±0.05	0.08^{Aa}±0.04	0.19^{Ba}±0.04	0.13^{Aa}±0.01
PO4-P (mg/L)	0.04^{Aa}±0.03	0.09^{Aa}±0.09	0.04^{Aa}±0.03	22.29^{Bb}±8.78	22.13^{Bb}±6.18	0.4^{Aa}±0.1
Chlorophyll A (g/L)	81.96^{Aa}±110.08	576.09^{Bb}±687.06	175.86^{Aa}±77.54	120.65^{Ba}±61.35	21.94^{Aa}±15.78	45.3^{ABa}±30.19
Stocking density (ea/m2)	8.5^{Cab}±0.58	6.5^{Bab}±0.58	4.0^{Aa}±0.00	300^{Cd}±4.2	59.7^{Bc}±1.71	9.5^{Ab}±0.58

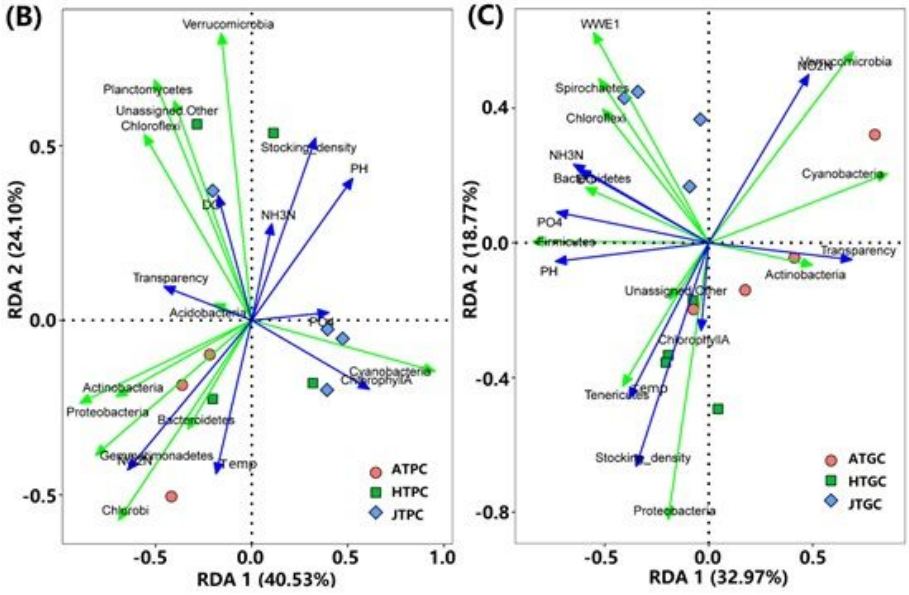

Figure 2

Water quality indexes and RDA analyses of PC and GC environments. A: water quality indexes in different periods of PC and GC environments of C. reevesii.

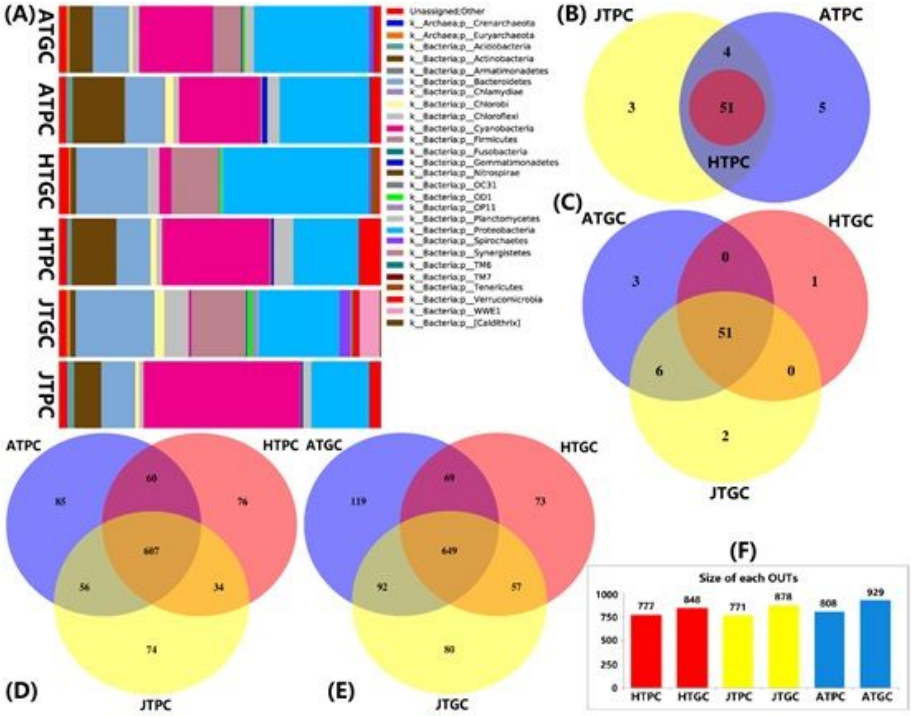

Figure 3

The bacterial community, microbiota structure and OTUs in phylum and genus levels of the PC and GC groups.

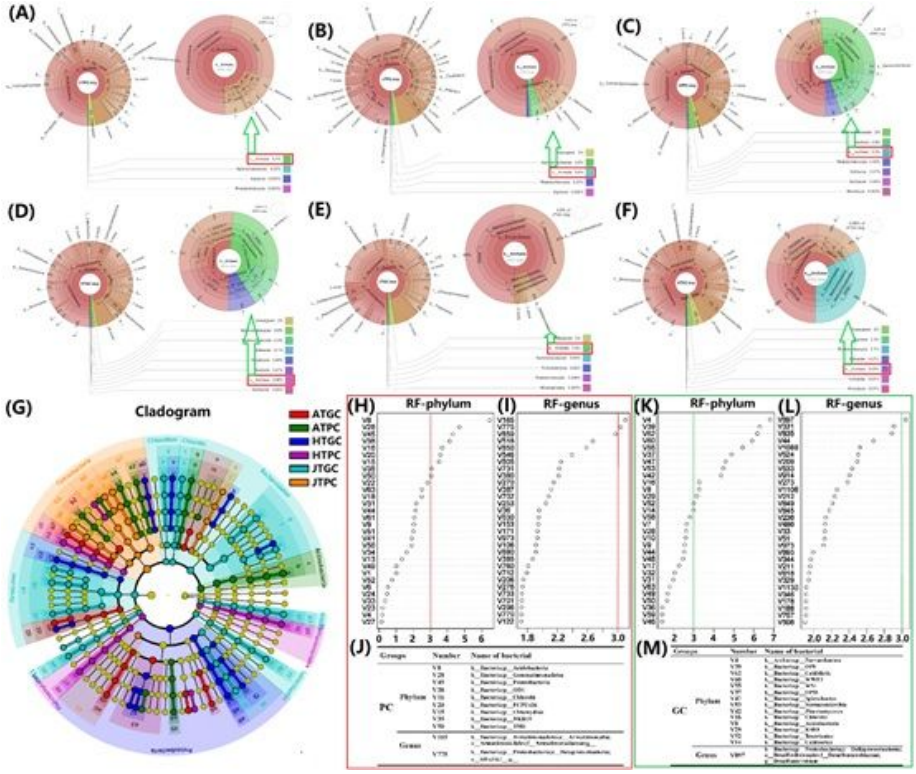

Figure 4

Bacterial and Archaea, LEfSe analysis and random forest model of different classifications annotation of the freshwater turtle C. reevesii in six groups.

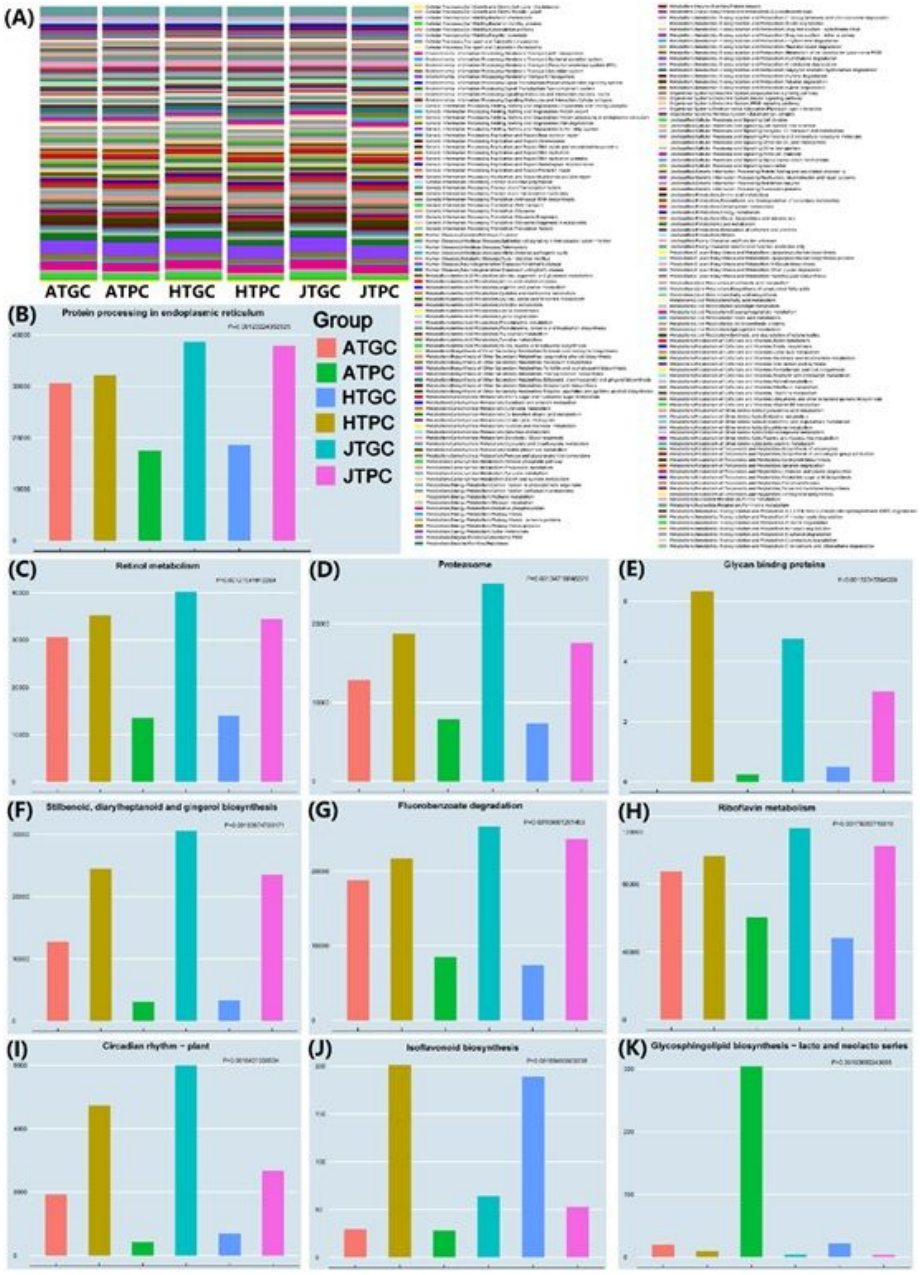

Figure 5

Gene function prediction (A) and top10 in gene function predictive expression P value (B-K) of all six groups based on the COG database.

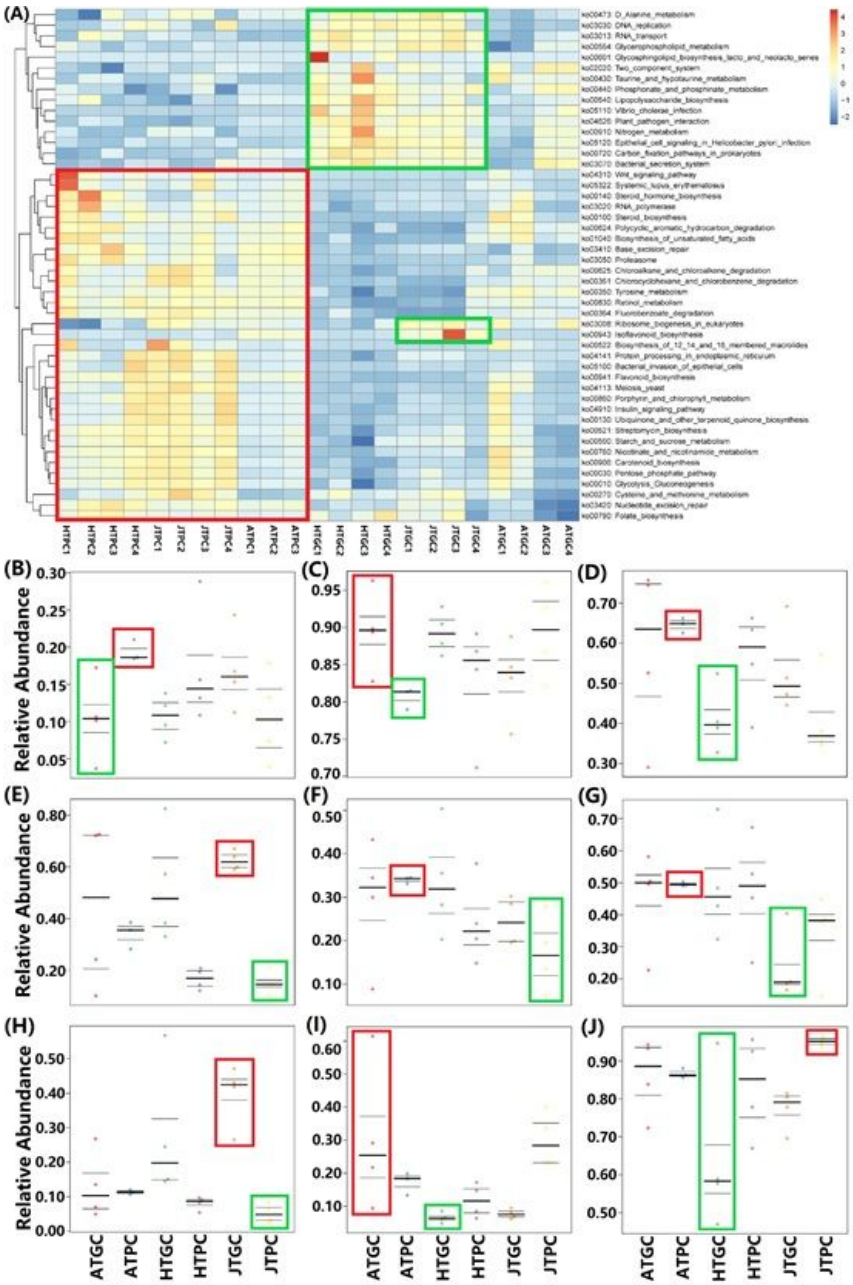

Figure 6

KEGG pathway annotation information and prediction of bacterial phenotype of PC and GC groups.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- S7.pdf
- S2.pdf
- S5.pdf
- S6.pdf
- S3.pdf
- S4.pdf
- S1.tif