Corn starch processing wastewater treated by a full-scale expanded granular sludge bed reactor and comprehensive analysis of microbial community at low and high organic loading rate

H H Yan¹, L Han², Q Yin¹, X Y Guo¹ and Y G Nian¹,³

1 State Key Laboratory of Environmental Criteria and Risk Assessment, Research Center of Pollution Control Technology, Chinese Research Academy of Environmental Sciences, Beijing, China
2 Sichuan Rongxinkai Engineering Design Corporation, Beijing Branch, Beijing, China
E-mail: yanh@craes.org.cn

Abstract. In this study, a full-scale expanded granular sludge bed (EGSB) reactor treating corn starch processing wastewater (CSPW) was monitored. The efficacy and microbial diversity of sludge granules were thoroughly investigated. The results demonstrated that 90.7% of the average chemical oxygen demand (COD) removal efficiency was observed to be in the range of an organic load rate (OLR) of 1.3 to 18.7 g COD/L·d. The Illumina MiSeq sequencing revealed that Proteobacteria, Bacteroidetes, Acidobacteria and Euryarchaeota were the dominant phyla. Acetoclastic and hydrogenotrophic methanogens were shown to play functions in methane production and the maintenance of the granular sludge under high and low OLR operation conditions, respectively. Syntrophorhabdus, Syntrophus and Candidatus Cloacimonas had the correlation with the transformation of the proteinaceous substances into VFAs. The results will guide development of microbial management methods to improve the process stability of EGSB treating CSPW.

1. Introduction
Corn starch accounts for more than 80% of the global starch market, making it an immensely important food and industrial ingredient [1]. As China is the leading and staple corn-planted country in the world, the corn starch processing industry has emerged and developed rapidly over the past few decades there. The starch production process consumes large amounts of fresh water, and most of it is discharged with a high intensity of organic matter and nitrogen compounds. If CSPW is not treated properly and effectively, it will cause serious environmental problems due to its high organic content. Therefore, efficient technologies for these organic wastewater have widely been proposed and promoted.

Due to its advantages of ecologically-benign energy generation, low operational costs and high pollutant removal efficiency, anaerobic digestion has gained enormous popularity and has been widely applied in the treatment of the wastewater with a high concentration of organic matter [2]. Anaerobic digestion consists of a series of sequential biochemical steps, involving a great variety of microbial groups, such as hydrolytic bacteria, acidogenic bacteria, acetogenic bacteria and methanogenic archaea. The microbial community structure is the core for process efficiency and stability, and microbial consortium in anaerobic digestion varies with the change of operational parameters and substrates. In
recent years, many studies have investigated microbial population structures by high-throughput sequencing methods in anaerobic reactor treating potato starch processing wastewater or food waste-recycling wastewater [3]. However, the reference data for profiling of microbial in anaerobic reactor fed by CSPW are limited. Moreover, the transformation of organic matters and related variation of microbial diversity at the different OLRs during anaerobic digestion for CSPW have seldom been reported.

For these reasons, in the present study CSPW was treated in the full-scale EGSB, and the process performance was evaluated, characteristics of the organic matter transformation was investigated by the measure of fluorescence excitation emission matrix (EEM). Major microbial populations and their dynamic shift with different OLRs were revealed via 16S rRNA high throughput pyrosequencing. Thereafter, these findings may provide valuable information regarding the optimization of process parameters and prediction of bioreactor performance in anaerobic reactor treating CSPW.

2. Materials and methods

2.1. Full-scale EGSB and sample collection
The full-scale EGSB of the typical corn starch processing corporation is located in Jilin Province, which has the treatment capacity of 3500 m³/d. The wastewater treatment system consisted of a sedimentation tank, equalization tank, anaerobic digester and aerobic basin in series. The wastewater was mainly generated from two steps, the first during the process of evaporation, and the other from the starch-rinse procedure. The starch-rinsed wastewater was discharged discontinuously, which led to a fluctuation of OLR in the influent.

The anaerobic digester was operated at an HRT of 20~26 h, under a constant temperature of 38~40℃. The wastewater samples were collected every day for eight months, in order to evaluate the performance of the digester. All of the collected samples were transported to the laboratory within 24 h, and stored at 4℃ as soon as possible.

2.2. Physicochemical analysis
The COD and volatile fatty acids (VFAs) were measured following the Standard Methods. The wastewater sampled at the 132th day (at an OLR of 7.0 g COD/L·d) was then further analyzed by EEM. The water samples were filtered through a filter paper with a mean pore size of 0.45 μm for EEM analysis, then measured by a Hitachi F-7000 spectrofluorometer. The EEM spectra were recorded with scanning emission and excitation wavelengths ranging from 200 to 500 nm. The excitation and emission bandwidths were set at 5 nm. The fluorometer was set at a speed of 1200 nm/min and a PMT voltage of 700 V. The samples were diluted to a final TOC concentration of 1 mg/L with Milli-Q water to avoid the inner-filter effect.

2.3. High-throughput sequencing analysis
Archaeal 16S rRNA genes and bacterial 16S rRNA genes of the collected sludge on the 132th day (at an OLR of 7.0 g COD/L·d) and 170th day (at an OLR of 2.0 g COD/L·d) from the digester were sequenced. The total genome DNA from samples was extracted using the CTAB method according to the manufacturer’s instructions. The V5 and V6 regions of the archaeal 16S rRNA gene were amplified in a polymerase chain reaction (PCR) with universal primers ARC519F (CAGCCGCCGCGGTAA) and ARC915R (GTGCTCCCCCGCCAATT CCT). Primers 340F (CCTACGGGNBGCASCAG) and 805R (GACTACNVGGGTATC TTG) were used to amplify the V3-4 regions of the bacterial 16S rRNA gene. The 30 μL PCR reaction mixture consisted of 15 μL of Phusion® High-Fidelity PCR Master Mix (New England Biolabs), 5 μL 0.2 μM of forward and reverse primers, and 10 μL template DNA. The PCR amplification was performed under the following conditions: initial denaturation at 98℃ for 1 min, followed by 30 cycles of denaturation at 98℃ for 10 s; annealing at 50℃ for 30 s and extension at 72℃ for 30 s; and a final extension at 72℃ for 5 min. The PCR products were detected in 2% agarose gel, and purified with a GeneJETTM Gel Extraction Kit (Thermo Scientific). The amplicons of every sample were analyzed quantitatively by Qubit 2.0.
Qualified DNA was sequenced finally on the Ion S5(TM) XL platform at the Beijing SinoGenoMax Co., Ltd. (Beijing, China).

3. Results and discussion

3.1. Operation performance

The long-term results of the COD removal efficiency in the full-scale EGSB system are shown in figure 1. During the experiment, the EGSB feed COD fluctuated between 3014 to 12462 mg/L, with an average content of 6184 mg/L. Meanwhile, the concentration of COD in the effluent was constantly controlled at under 1300 mg/L, with an average content of 686 mg/L. The reactor showed a stable and efficient performance, with an average COD removal efficiency of 90.7%. The EGSB was able to handle a wide variation of COD concentration in the feed. This results showed that EGSB is an effective measure for CSPW treatment.

![Figure 1](image1.png)

Figure 1. COD removal rate of EGSB treating CSPW during the operation.

![Figure 2](image2.png)

Figure 2. Relationship between OLR and VFA of EGSB treating CSPW during the operation.

VFAs are major intermediates in anaerobic bioconversion of organic matter to methane and carbon dioxide. As figure 2 shows, the mean concentration of VFAs in the effluent was 281 mg/L. It was noticeable that VFAs in the effluent gradually decreased with the OLR. When OLR was during the range of 6.0 - 7.0 g COD/L·d, VFAs in the effluent got the minimal concentration of 237 mg/L, then VFAs showed the rising trend with OLR increased.

Figure 3 shows the change of COD removal at different OLRs. It was illustrated that the OLR varied considerably from 1.3 to 18.7 g COD/L·d, with the maximum COD average removal rate of 93.0% being achieved at an OLR in the range of 6.0-7.0 g COD/L·d. This result was higher than the lab-scale experiment conducted by Antwi, who had reported an OLR range of 2.7-13.27 g COD/L·d in the upflow anaerobic sludge blanket reactor treating potato starch processing wastewater [4]. In general, the granules in EGSB reactors have the higher transfer rate of substrates, which could result in higher microbial activity and more efficient removal of organic matters.

The EGSB reactor had a wide variation of OLR in the feed, which was due to the discontinuous discharge of the starch-rinsed wastewater. A large amount of protein-like substances were found under the high OLR. Meanwhile, inflow with the low OLR mainly consisted by the relatively small molecular organic matter, which was generated in the process of evaporation. As we known, starch wastewater is rich in biodegradable components such as starch and protein, which are rate limiting matters in COD decomposition. However, it was noticeable that the COD removal in the study was relatively efficient at the high OLR. This could be ascribed to the dynamics of microbial community. Rincón et al. reported that bacterial community structures will be firstly affected by the OLR fluctuation [5]. Therefore, the study on microbial community under different OLRs is of high importance in performance regulation and optimization.
3.2. Overall analysis of high-throughput sequencing

In order to reveal the bacterial and archaea community structure and diversity in the EGSB treating CSPW under different OLRs, Illumina Miseq pyrosequencing was employed to analyze the 16S rRNA gene of the sludge samples obtained at the OLR of 2.0 g COD/L·d and 7.0 g COD/L·d, respectively. As shown in table 1, 63,585 to 80,200 effective sequence reads were obtained at the high and low OLR. The Good’s coverage was above 0.98, which demonstrated that the gene library accurately reflected the profile of microorganisms in the anaerobic digester. The Chao estimator was used to evaluate the richness of the microbial community. In general, the Chao index of bacterial and archaea apparently increased as the OLR elevated, thereby indicating the richness of microbial community increased when the reactor operated at the high OLR. The Shannon and Simpson values, which represent bacterial diversity, were higher at the low OLR than at the high OLR. This demonstrated that the microbes under the condition of low OLR had more abundant microbial species.

| Microbes | Mode   | Sequence | Coverage | Chao   | Shannon | Simpson |
|----------|--------|----------|----------|--------|---------|---------|
| Bacterial| low OLR| 63,585   | 0.983    | 3389   | 5.46    | 0.034   |
|          | high OLR| 80,200   | 0.987    | 6521   | 4.67    | 0.027   |
| Archaea  | low OLR| 72,890   | 0.991    | 4620   | 3.59    | 0.238   |
|          | high OLR| 68,976   | 0.986    | 5653   | 3.07    | 0.196   |

3.3. Bacterial communities under different OLRs

Figure 4 shows the bacterial community at the phylum level in granules collected from the anaerobic digester (relative abundance greater than 1% was considered). The bacterial samples of the different OLRs showed a high predominance of *Proteobacteria* with an average relative abundance of 25.60%, followed by *Bacteroidetes* (average 21.12%). The phyla of *Proteobacteria* and *Bacteroidetes* were typically reported as the most dominant phyla in anaerobic sludge and methanogenic granular sludge [6]. *Proteobacteria* was believed to have performed heterotrophic functionalities, and to have degraded a broad range of organic pollutants. *Bacteroidetes* exerted the fermentative effects by breaking down macromolecules, such as protein and starch, during the anaerobic process. In addition, the phyla *Chloroflexi*, *Firmicutes* and *Spirochaetes* were also identified in this study. The phylum *Chloroflexi* was attributed to the degradation of carbohydrates and other cellular components [7]. The presence of the phylum *Firmicutes* performed a major role in the protein and amino acid decomposition. Antwi et al. revealed that the synergistic coexistence of *Firmicutes* and *Chloroflexi* mainly contributed to the hydrolysis and acidogenesis process to produce methane in the anaerobic reactor treating potato starch wastewater [4]. *Spirochete* is an obligate fermenter that catabolizes proteins and carbohydrates, thus resulting in acetate, ethanol, and molecular hydrogen production [8]. Notably, the relative abundance phyla *Acidobacteria* decreased from 16.22% to 0.87% when the anaerobic digester operated at a low OLR, whereas *Synergistetes* were more abundant if the digester
operated at low OLR, with the average relative abundance of 18.2%. The phyla Acidobacteria and Synergistetes were reported in previous work, where these microbes were observed as the prominent phyla in the stable state of the mesophilic digester treating food waste [9].

The genera distribution of the bacterial communities is presented in figure 5. It was demonstrated that the genera with relative abundance greater than 1% only accounted for 47.5% and 50.8% of the total bacterial in the high and low OLR samples, which indicated that there is no absolute host for the bacteria in the anaerobic sludge. This phenomenon was mainly a result of the varied organic composition and different OLRs of the influent. The fermentation and hydrolyzation of the substrates were carried out by the synergistic actions of the complex groups of microorganisms, though many bacteria could not be identified in a certain genus unit.

The major genera affiliated with the phylum Proteobacteria were Syntrophorhabdus and Syntrophobacter, with respective average relative abundances of 4.60% and 4.54%. Previous studies have reported that the genus Syntrophorhabdus is a mesophilic bacterium capable of degrading aromatic substrates in syntrophic cooperation with a partner methanogen [10]. Syntrophobacter was considered as a syntrophic metabolizer, with the potential abilities to generate hydrogen and provide the substrate for hydrogenotrophic methanobacters. It was also reported that Syntrophobacter are a stable and resilient functional group of bacteria in anaerobic digestion systems, which grow on acetate, propionate, butyrate and sulfate [11]. Unidentified Acidobacteria (average 9.47%) and Thermoanaerobaculum (average 1.56%) belonged to the phylum of Acidobacteria. Thermoanaerobaculum, have been identified to have the capabilities of fermentative growth on pyruvate, and reducing Fe (III) and Mn (IV) [12]. The dominant genera belonging to the phylum Bacteroidetes included unidentified Lentimicrobiaceae (average 4.98%) and unidentified Rikenellaceae (average 4.42%). It was observed that the genus Candidatus Cloacimonas, classified as part of the phylum Cloacimonetes, showed a rising trend to the mount of 2.45% when the digester
operated at the high OLR. It was reported that these genera are typically found in anaerobic environments, and derive most of their energy from the fermentation of amino acids.

3.4. Archaea communities under different OLRs
As the inherent phylogenetic low diversity of archaea, the common detected methanogens in anaerobic digesters are far more less than bacterials. The archaea in the studied digester were mainly concentrated in the phyla of *Euryarchaeota* (average 84.50%), *Diapherotrite* (average 5.01%) and *Crenarchaeota* (average 4.88%) (figure 6). Noticeably, *Euryarchaeota* was the most dominant phylum in the samples of the different OLRs, and is believed to perform acetoclastic-hydrogenotrophic functionalities in metabolizing VFAs produced from proteinaceous degradation and byproducts of acetogenesis.

![Figure 6](image1.png) 
*Figure 6.* Distribution of archaeal population in anaerobic sludge obtained from the EGSB treating CSPW at the phylum level: high OLR (a) and low OLR (b).

In order to grasp further details of community structures, the population should be studied at the finer taxonomic levels. The relative abundance of the sample was calculated at the genus level. For archaea at the genus level (as shown in figure 7), *Methanosaeta*, unidentified *Thermococci*, *Methanobacterium* and *Methanolinea* dominated in sludge samples at the high and low OLRs, all of which were affiliated with the phylum *Euryarchaeota*. The results exhibited no notable change in terms of methanogens compositions, but significantly discrimination was observed in the relative abundance at different OLRs. *Methanosaeta* and unidentified *Thermococci* were predominant in the high OLR sample, with the relative abundances of 35.42% and 23.92%. This was greatly supported by the morphological observation of microbes through scanning electron microscope analysis, where *Methanosaeta* and cocci-shaped archaea were found in the core zone of the granule consortia. However, for the low OLR sludge, the relative abundance was increased for *Methanolinea* and *Methanobacterium*, with the relative abundances of 35.16% and 15.26%.

![Figure 7](image2.png) 
*Figure 7.* Distribution of archaeal population in anaerobic sludge obtained from the EGSB treating CSPW at the genus level: high OLR (a) and low OLR (b).
3.5. Functional population and related role in organic matter removal

There was a distinction of microbial community among the sludge samples at high and low OLRs, which implied that the variation of the influent had a great influence on the microbial community structure in the EGSB reactor. The relative abundances of *Syntrophorhabdus*, *Syntrophus* and *Candidatus Cloacimonas* decreased significantly when the OLR was reduced from 7.0 to 2.0 g COD/L·d. These functional bacteria are interpreted as a microbial response to transforming the proteinaceous substances into VFAs. Both the well-known acetoclastic methanogen *Methanosaeta* and hydrogenotrophic methanogens, such as *Methanobacterium* and *Methanolinea*, were observed in the samples. However, the predominance of archaea bacteria changed among the different OLRs. The relative abundance of *Methanosaeta* increased if the EGSB was operated at the high OLR. When the EGSB was operated with the low OLR, the growths of *Methanobacterium* and *Methanolinea* were improved. The variation of dominant methanogens may be the reason for the relatively high content of VFAs in the EGSB effluent when digester was operated at the low OLR (as shown in figure 2). These results indicate that the acetoclastic methanogens and hydrogenotrophic methanogens respectively perform functions in the methane production and maintenance of the granular sludge at high and low OLR operation conditions, which is due to the large range of variation in influent composition and OLR. This is in line with a previous research, showing that methanosaeta-like acetoclastic methanogens are generally accepted to be competitive with hydrogenotrophic methanolinea-like methanogenic community under the different OLRs [13].

![Figure 8](image_url)

**Figure 8.** EEM of influent and effluent in EGSB reactor: influent (a), effluent (b).

The EEM spectra provide transformation of organic matters during the EGSB treatment at the high OLR (as shown in figure 8). The spectra showed that four peaks related to potein-like matters were detected in the influent, of which the fluorescence is associated with the tyrosine or tyrosine-like substances (Peaks A and C) and tryptophan protein-like substances (Peaks B and D). These matters were in agreement with the characteristics of CSPW [14]. After the treatment of EGSB reactor, only fluorescence peaks B and D were monitored in the effluent. It was also noted that Peaks B and D showed a red shift (5-10 nm) in terms of emission wavelength in the effluent. These results suggested that the monocyclic fluorescence compounds could be completely degraded into fatty acids and/or inorganic matter through a microbiological mediated process. Whereas, the configuration of the complicated matters had changed or transformed by the cleavage of C-N in the cyclic ring and hydrolyzation, making them more likely to generate carbonyl-containing substituent, hydroxyl, alkoxy, amino groups and carboxyl constituents. It was reported that the reduction of the amino group into amine is quite common in anaerobic environments [15]. In the present study, it was found that the amine mean concentration in the effluent was 495 mg/L, which was much higher than that of the inflow (68 mg/L).
According to the analysis, we propose that different metabolic pathways for the mineralization of organic matters at the high OLR could be active simultaneously in these microcosms (figure 9).

4. Conclusions
In this study, the full-scale EGSB reactor was employed to treat CSPW, the efficacy and microbial diversity were studied. When the OLR ranged from 1.3 to 18.7 g COD/L·d, the average COD removal could reach 90.7%. Proteobacteria, Bacteroidetes, Acidobacteria and Euryarchaeota were the dominant phyla. Syntrophorhabdus, Syntrophus and Candidatus Cloacimonas played a function in the degradation and transformation of the proteinaceous substances. The predominance of archaea bacteria in the granule sludge shifted from Methanoseta and unidentified Thermococci to Methanolinea and Methanobacterium at the high and low ORLs, respectively.

5. References
[1] Desam G P, Li J S, Chen G B, Campanella O and Narsimhan G 2018 Carbohyd. Polym. 199 331
[2] Lee J Y, Kim E J, Han G S, Tongco J V and Shin S G 2018 Bioresour. Technol. 259 388
[3] Kim E J, Lee J Y, Han G S and Hwang S H 2018 Bioresour. Technol. 259 442
[4] Antwi P, Li J Z, Boadi P O, Meng J, Quashie F K, Wang X, Ren N Q and Buelna G 2017 Bioresour. Technol. 239 105
[5] Rinćon B, Borja R, González J M, Portillo M C and Sáiz-Jiménez C 2008 Biochem. Eng. J. 40 253
[6] Guo J, Peng Y, Ni B J, Han X, Fan L and Yuan Z 2015 Microb. Cell Fact 14 33
[7] Song G Q, Xi H B, Zhou Y, Fu L Y, Xing X and Wu C Y 2017 Bioresour. Technol. 243 502
[8] Dong X, Greening C, Brüls T, Conrad R, Guo K, Blaskowski S, Kaschani F, Kaiser M, Laban N A and Meckenstock R U 2018 ISME J. 12 2039
[9] Li L, He Q, Ma Y, Wang X and Peng X 2015 Bioresour. Technol. 189 113
[10] Nobu M K, Narihiro T, Tamaki H, Qiu Y L, Sekiguchi Y, Woyke T, Goodwin L, Davenport K W, Kamagata Y and Liu W T 2014 Genome Announc. 2 1064
[11] Cetecioglu Z, Dolfing J, Taylor J, Purdy K J and Eyice Ö 2019 Water Res. 155 444
[12] Losey N A, Stevenson B S, Busse H J, Singinghe Damsté J S, Rijpstra W I, Rudd S and Lawson P A 2013 Int. J. Syst. Evol. Micr. 63 4149
[13] Yang B, Xu H, Wang J F, Song X S, Wang Y H, Li F, Tian Q, Ma C Y, Wang D Y, Bai J H and Sand W 2017 Bioresour. Technol. 244 750
[14] Yan H H, Nian Y G, Zhou Y X, Yin Q, Zhong Z Z and Guo XY 2017 Microchem. J. 130 295
[15] García-Mancha N, Monsalvo V M, Puyol D, Rodríguez J J and Mohedano A F 2017 J. Hazard. Mater. 339 320

Acknowledgements
The work is financially supported by the China special S&T project on treatment and control of water pollution (2012ZX07202-009), the Fundamental Research Funds for Central Public Welfare Scientific Research Institutes of China (Grant No. 2019YSKY-014).