An Approach to Identify ‘Microbial Dark Matter’ Parcubacteria in Environmental Samples and Enrichments from Woods Hole

Lei Zhou

2017 Microbial Diversity Course student

School of Chemistry and Molecular Engineering, East China University of Science and Technology, 130 MeilongRoad, Shanghai 200237, P.R. China

Engineering Research Center for Microbial Enhanced Energy Recovery, Shanghai 200237, P.R. China

Email: leizhou@ecust.edu.cn

Introduction:

Microorganisms are the most diverse and abundant life forms on Earth[1,2]. As the development of genome-resolved metagenomics, Hug *et al* [3] presented a new view of the tree of life with sequenced genomes. However, large numbers of major lineages have no isolated representatives. The phyla composed exclusively of uncultured representatives are referred to as Candidate Phyla Radiation (CPR). Because these organisms likely account for a large proportion of the biomass and diversity on earth, and their metabolisms are not well known yet, so microbiologists define them as ‘microbial dark matter’[1]. Investigating the phylogenetics and metabolic capacities of these members will open a new door for the ecology and evolutionary history.

In the CPR, Parcubacteria is a superphylum, which has 497 genomes (generally<1Mb) in NCBI (by 8/21/2017). To date, there is no isolation of Parcubacteria, which are mostly identified in anoxic environments. Metabolic predictions show that the members have anaerobically fermentative metabolisms, and some likely impact hydrogen and sulfur cycles. Because they are members of the CPR, they are non-respiring due to lacking genes for electron transport and the tricarboxylic acid cycle (TCA)[4]. But a new candidate Parcubacteria genome showed that it had unusual respiratory capacities and nitrogen metabolism[5]. They lack genes for biosynthesis of amino acids, nucleotides, vitamins, and lipids. Metagenomic analysis suggested that they obligately ferment sugars to fatty acids and some likely degrade complex carbon source[4-6].

So how are Parcubacteria distributed in the environment? How to detect them? Can they be cultured? The goal of the miniproject was to identify and quantify Parcubacteria in environmental samples and enrichments. The strategy for the study is shown in Figure 1. Environmental samples were collected from Trunk River and Cedar Swamp. Different methods were used for metagenomic identification. According to the properties of samples,

the media were designed and applied to enrich Parcubacteria. Specific FISH-probe and qPCR primers were used for identification and quantification of Parcubacteria.

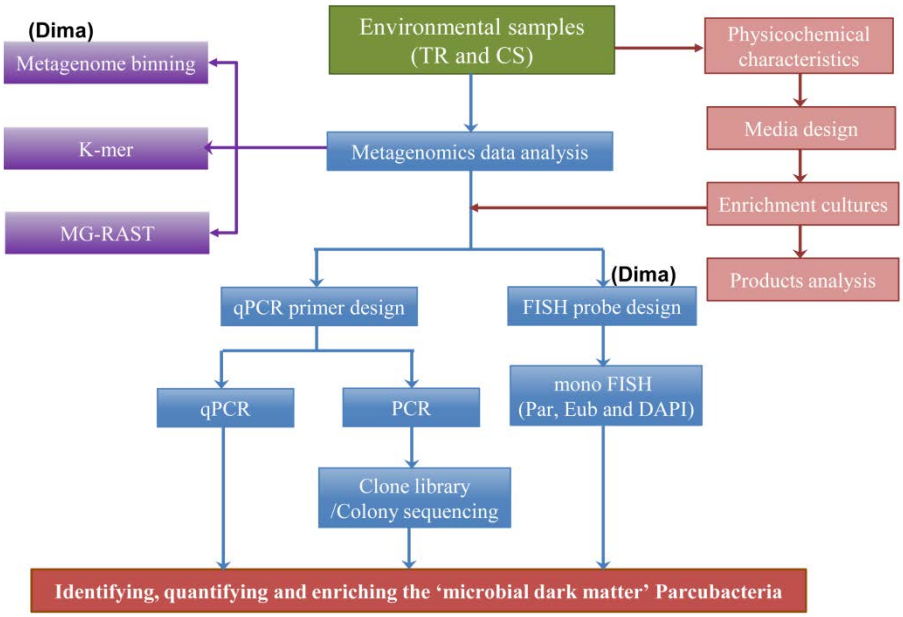

Figure 1 Research strategy for the miniproject

Materials and Methods

Sample collection

Samples were collected from Trunk River's 'Lemonade' (TR, 41.534207N, -70.641519W) by inserting a syringe below the water and Cedar Swamp (CS, 41.527726N, -70.654074W), Woods Hole (Figure 2). There was a great deal of sediment in both samples which showed green and brown. Additionally, sea grass was also collected from Trunk River for later use in enrichment cultures.

Figure 2 Sampling sites and samples appearance.

Characterization of samples

Multiparameter water quality meter was applied for *in-situ* monitoring the temperature, pH, salinity, oxidation reduction potential (ORP), conductivity, dissolved oxygen (DO) of samples. After taken back to the lab, samples were analyzed by Ion Chromatography (IC,

Dionex ICS-2000). Samples were filtered using an 0.2 μm filter and diluted with MilliQ water. Anions were detected by using the Dionex IonPac AS22 column using isocratic with an eluent of 1.4 mM $NaHCO_3$ and 4.5 mM Na_2CO_3 for 15 minutes, while cations were detected by using the Dionex IonPac CS19 column using isocratic with an eluent of 7 mM Methanesulfonic Acid for 15 minutes. Data were shown in Table 1.

Table 1 Physicochemical characteristics of samples

Sample	Temp. (ºC)	pH	Salinity (ppt)	ORP (mV)	DO (ppm)	Ions (mg/L)									
						F^-	Cl^-	Br-	NO_3^-	PO_4^{3-}	SO_4^{2-}	Na^+	K^+	Mg^{2+}	Ca^{2+}
TR	25.3	6.0	27.5	-279.0	19.5	--	57.8	0.2	0.1	--	4.4	26.3	0.6	3.2	1.0
CS	19.6	6.8	0.5	-307.3	2.6	0.03	13.4	0.3	0.04	1.5	0.4	8.5	0.9	1.2	2.3

Enrichment cultures

To enrich Parcubacterium, medium was specifically designed adapted for the conditions of each environmental sample with different types of sugar (Table 2, according to 2017 Microbial Diversity manual). After preparation, 9 mL medium was transferred into an anaerobic tube under anaerobic condition and then sealed with a butyl rubber stopper and aluminum crimp seal. The headspace was vacuumed and filled with N_2 gas. 1 mL environmental sample was inoculated into the tube and then 5 dilution tubes were made for each sugar (Table 3).

Table 2 Medium components of enrichment cultures

Component	TR DW (mL/L)	CS FW (mL/L)
1× SW	700	--
100× FW	--	700
1M MOPS	20	20
1M NH_4Cl	10	10
100× K phosphate	10	10
1000× Vitamin solution	1	1
1000× Trace elements	1	1
10% Yeast stock	1	1
10% Tryptone stock	1	1
2.5% Na_2S	3	3
Habitat juice	53	53
Sugar:		
18% Glucose solution	100	100
Cellobiose solution(saturated)	100	100

Table 3 Enrichment cultures

	ID	Inoculums	Amended	Dilution Number
1	TG	TR	Glucose	5
	TC	TR	Cellobiose	5

	T	TR	--	5
2	CG	CS	Glucose	5
	CC	CS	Cellobiose	5
	C	CS	--	5

Taxonomic classification of metagenome

K-mer analysis of TR metagenome was conducted with a command-line tool sourmash (https://sourmash.readthedocs.io/en/latest/). K-mer sizes for comparison were chosen at 21, 31, or 51, respectively. TR assemblies were also submitted to MG-RAST (http://metagenomics.anl.gov/) for identification.

Bioproduct analysis

Headspace samples (250 µL) were collected using a 250 µL gas-tight syringe and were injected into a gas chromatograph (GC, SHIMADZU GC-2014) coupled with a flame ion detector (FID) and thermal conductivity detector (TCD). The oven temperature was isothermal at 130 ºC. The injector and detector temperatures were kept at 200 ºC. High Performance Liquid Chromatography (HPLC, Shimadzu LC-2010c) equipped with BioRad Aminex HPX-87H column, a UV/vis Detector and Refractive Index Detector was used for liquid product detection. 1 mL sample was centrifuged, and 900 µL supernatant was taken followed by adding 100 µl of 5M H_2SO_4. Then the sample was centrifuged and filtered using an o.2 µm filter for HPLC analysis.

monoFISH

2 mL sample was fixed with 37% paraformaldehyde solution (final conc. 4%) for 1 h at room temperature. Fixed sample was washed three times with 1×PBS, centrifuged at 16000×g for 5 min and then stored in 1×PBS at -20 ºC. 200 µL fixed sample was filtered on 0.2 µm and 0.025 µm filter separately. Each filter was cut into pieces. 9 µL of hybridization buffer was mixed with 1 µL of Parcu537 probe and 1 µL of EUB338-I-III probe on a slide (Table 4). The filter piece was put on the mix and the slide was placed into a tube incubating at 46 ºC for 2 h. Filter was transferred to washing buffer incubating at 48 ºC for 10 min, and washed with milliQ water and 96% EtOH. After air-dry, the filter was placed on glass slide with a mix of Citifluor, Vectashield and DAPI. The microscope images were recorded on a ZEISS AXIO Imager. A2.

Table 4 FISH probes used in this study

Probe ID	Target Organism	Sequence (5’-3’)
Parcu537	Parcubacteira	GGATAACGCTTGAGGTCTCT
EUB338-I-III	Bacteria	GCWGCCWCCCGTAGGWGT

DNA extraction, PCR amplification and clone library

Sample was centrifuged at 14000×g for 5 min and then the supernatant was discarded. 10 µL of sample was added into 20 µL of ALP reagent (alkaline PEG200) followed by boiling cells with a PCR program. PCR reaction mixtures contained 25 µL GoTaq® Green Master Mix (Promega), 2.5 µL primers (Table 5), 4 µL DNA and 16 µL Nuclease-free water. PCR reaction was performed as follows: 95 ºC for 2 min; 35 cycles of 95 ºC for 30 s, 55 ºC for 30 s, followed by 72 ºC for 1.5 min; and a final 72 ºC for 10 min. Amplification was checked by running 5 µL of PCR product on a 1% agarose gel with stain. The PCR product can be directly cloned with Pgem T-easy vector system (Promega) and plated on to Luria Broth (LB) plates with 100 µg/mL ampicillin, IPTG (final conc. 0.5 mM) and 80 µg/mL X-Gal. The plates were incubated at 37 ºC overnight. White colonies were picked and sent for sequencing.

Table 5 PCR and qPCR primers used in this study

Primers	Target Group	Sequence (5'-3')	References
8F	Universal	AGAGTTTGATCCTGGCTCAG	[7]
1391R	Universal	GACGGGCGGTGWGTRCA	[7]
515F	Universal	GTGCCAGCMGCCGCGGTAA	[7]
907R	Universal	CCGTCAATTCMTTTRAGTTT	[8]
OD1_32F	Parcubacteria	CGGATTTTCCCGTGCAGG	[6]
OD1_1427R	Parcubacteria	TACCCGTGCCTTGTTACGAC	[6]
OD1_987F	Parcubacteria	GGTGCTGCATGGTTGTCGTC	[6]
OD1_1186R	Parcubacteria	GCTGCCCTCTGTAACTGCCA	[6]

qPCR

DNA quantities were measured by Quantus Fluorometer (Promega) according to manufacturer's instruction. Triple replicates for 5 DNA samples were applied to conduct qPCR. qPCR reaction mixture contained: 25 µL GoTaq® qPCR Master Mix (Promega), 2.5 µL each primer, 10 µL template and 10 µL Nuclease-free water. Two-step method was applied with the setting as: 15 s at 95 °C, 60 s at 60 °C for 40 cycles.

Results and Discussion

Taxonomic classification of metagenome

Last year Dima found Parcubacteria in TR metagenome by using binning. In this study, another method sourmash was tried to use for searching Parcubacteria in the same data. Results are shown in Table 6~8. There are 6, 11 and 16 matches found in data with different k-mer, respectively. But Parcubacteria was not found in it, and that maybe depended on the method used. MG-RAST annotations also showed no finding.

Table 6 Sourmash results with k=51

	overlap	p_query	p_match	Reference
1	1.2 Mbp	2.3%	47.3%	FNAQ01000039.1 Desulfuromonas thiophi...
2	440.0 kbp	0.8%	9.4%	JXYD01000001.1 Pseudoalteromonas sp. ...
3	80.0 kbp	0.2%	3.4%	CP016432.1 Prosthecochloris sp. CIB 2...
4	360.0 kbp	0.7%	1.3%	CP013138.1 Pseudoalteromonas sp. Bsw2...
5	60.0 kbp	0.1%	2.9%	CP001099.1 Chlorobaculum parvum NCIB …
6	70.0 kbp	0.1%	0.9%	AUTP01000001.1 Pseudoalteromonas sp. ..

Table 7 Sourmash results with k=31

	overlap	p_query	p_match	Reference
1	1.6 Mbp	3.0%	57.7%	FNAQ01000039.1 Desulfuromonas thiophi...
2	0.7 Mbp	1.2%	13.9%	CP013138.1 Pseudoalteromonas sp. Bsw2...
3	190.0 kbp	0.4%	8.8%	AE006470.1 Chlorobium tepidum TLS, co...
4	180.0 kbp	0.3%	7.4%	CP016432.1 Prosthecochloris sp. CIB 2...
5	180.0 kbp	0.3%	6.2%	CP001099.1 Chlorobaculum parvum NCIB ...
6	190.0 kbp	0.4%	3.3%	CP017305.1 Chlorobaculum limnaeum str...
7	0.7 Mbp	1.2%	1.3%	JXYD01000001.1 Pseudoalteromonas sp. ...
8	60.0 kbp	0.1%	2.1%	FREM01000944.1 Enterococcus faecium i...
9	180.0 kbp	0.3%	1.1%	BADV01000121.1 Pseudoalteromonas sp. ...
10	50.0 kbp	0.1%	1.3%	CP001896.1 Allochromatium vinosum DSM...
11	150.0 kbp	0.3%	2.0%	LUZT01000020.1 Chlorobiales bacterium...

Table 8 Sourmash results with k=21

	overlap	p_query	p_match	Reference
1	1.4 Mbp	2.6%	59.1%	FNAQ01000039.1 Desulfuromonas thiophi...
2	0.7 Mbp	1.2%	13.9%	JXYD01000001.1 Pseudoalteromonas sp. ...
3	310.0 kbp	0.6%	13.0%	CP001099.1 Chlorobaculum parvum NCIB …
4	300.0 kbp	0.5%	7.9%	CP017305.1 Chlorobaculum limnaeum str...
5	230.0 kbp	0.4%	8.9%	CP016432.1 Prosthecochloris sp. CIB 2...
6	300.0 kbp	0.5%	6.0%	LUZT01000020.1 Chlorobiales bacterium...
7	80.0 kbp	0.1%	1.5%	AFWV01000062.1 Thiocapsa marina 5811 ...
8	80.0 kbp	0.1%	1.7%	LMSL01000001.1 Frateuria sp. Soil773 ...
9	80.0 kbp	0.1%	2.9%	CP001108.1 Prosthecochloris aestuarii...
10	60.0 kbp	0.1%	1.7%	LSYU01000001.1 Marichromatium gracile...
11	50.0 kbp	0.1%	0.7%	KB902362.1 Lamprocystis purpurea DSM ...
12	240.0 kbp	0.4%	1.1%	BADV01000121.1 Pseudoalteromonas sp. ...
13	50.0 kbp	0.1%	0.6%	BDCU01000001.1 Nocardia fusca NBRC 14...
14	50.0 kbp	0.1%	0.5%	LRMT01000001.1 Cupriavidus sp. UYMMa0...
15	60.0 kbp	0.1%	0.6%	JABX01000001.1 Thiohalocapsa sp. ML1 ...
16	60.0 kbp	0.1%	0.4%	ASSI01000001.1 Burkholderia sp. AU4i ...

Bioproducts in enrichments

After 10 days of cultivation, products were detected in all the enrichments (Table 9).

Glucose was found in the enrichments with cellobiose. Various fatty acids and different amount of H_2 were detected in all the enrichments with sugar, but no methane. That means lots of fermentative organisms grew in enrichments.

Table 9 Metabolites detected in environmental samples and enrichments

(mM)	TR	TG	TC	T	CS	CG	CC	C
Glucose	--	--	61.694	--	--	--	27.108	--
Maltose	--	27.268	--	--	--	32.327	--	--
Formate	--	7.089	8.533	--	--	7.805	1.553	0.309
Acetate	--	14.165	10.438	--	--	--	5.090	--
Propionate	--	1.469	24.807	--	--	1.017	0.439	--
Lactate	--	2.096	0.570	--	--	12.267	5.583	--
Citrate	--	0.246	0.317	--	--	0.327	0.380	--
H_2	--	✓	✓✓	✓	--	✓✓	✓✓✓	✓

Note: ✓ detected --- not detected

monoFISH

The FISH-probe specific for Parcubacteria designed last year was remade with a colored dye (Cy3), which made signals more obvious to see than before. Depending on the different microbial abundance, there are a lot of signals in TR, while few in CS (Figure 3). All the enrichments of TR had signals (Figure 4) and only CG had signals (Figure 5). However, it's worth noting that some phototrophs also have fluorescence in red filter but their sizes are larger than Parcubacteria. The results showed some indication of searching Parcubacteria.

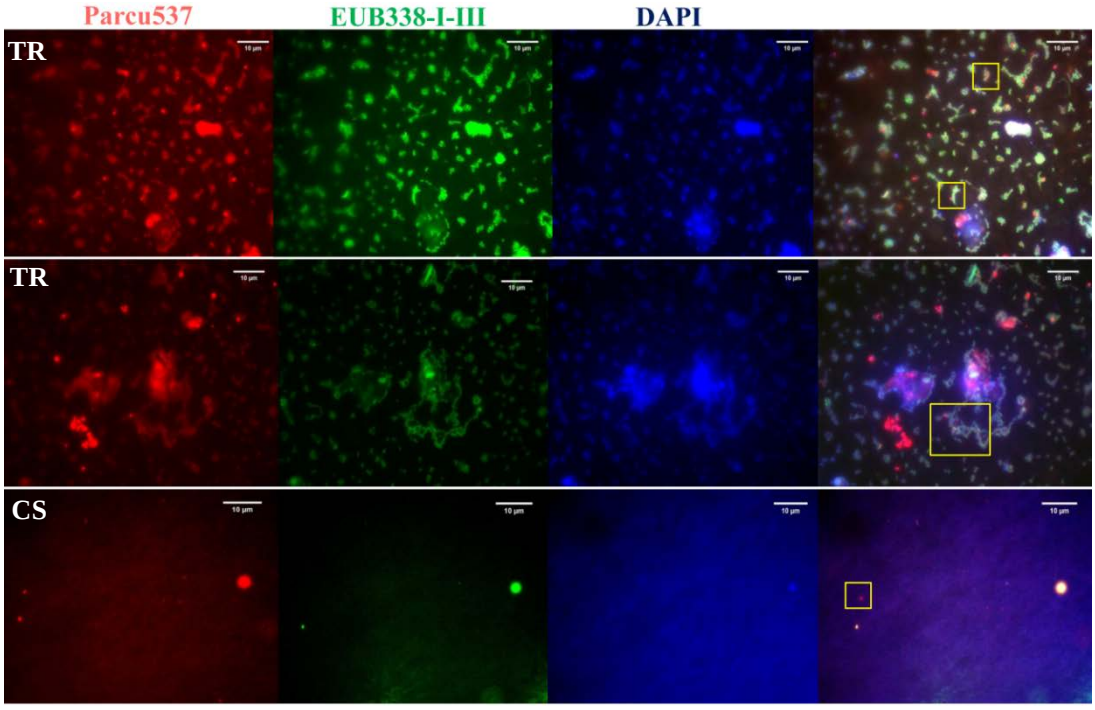

Figure 3 monoFISH images of evironmental samples

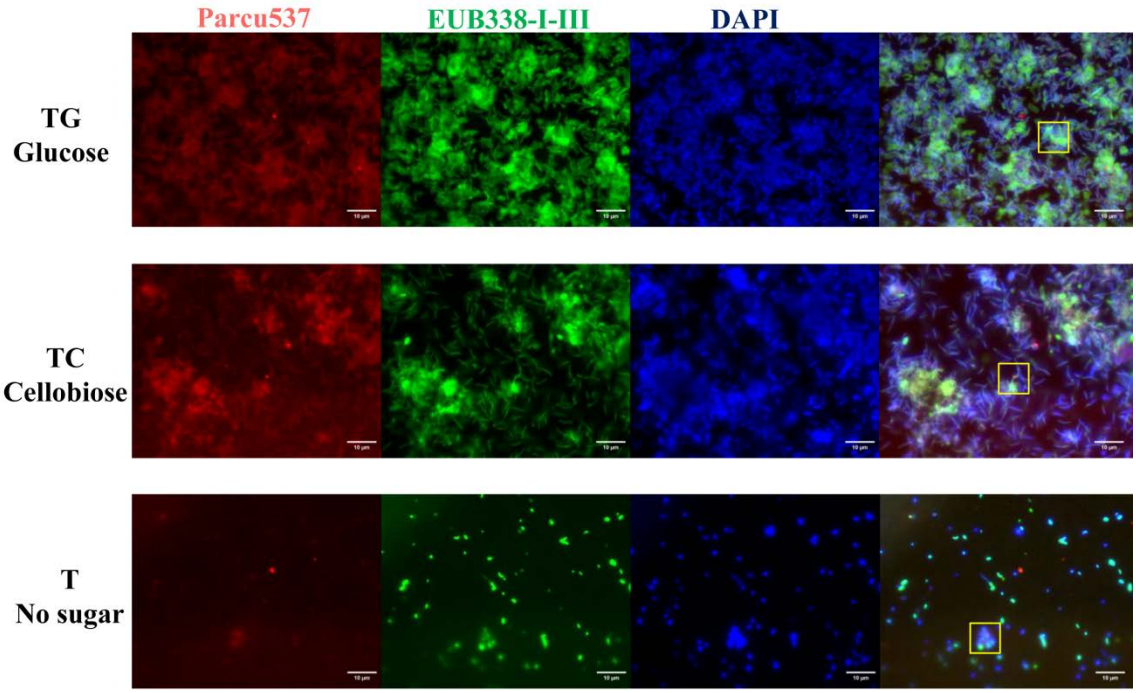

Figure 4 monoFISH images of TR enrichments

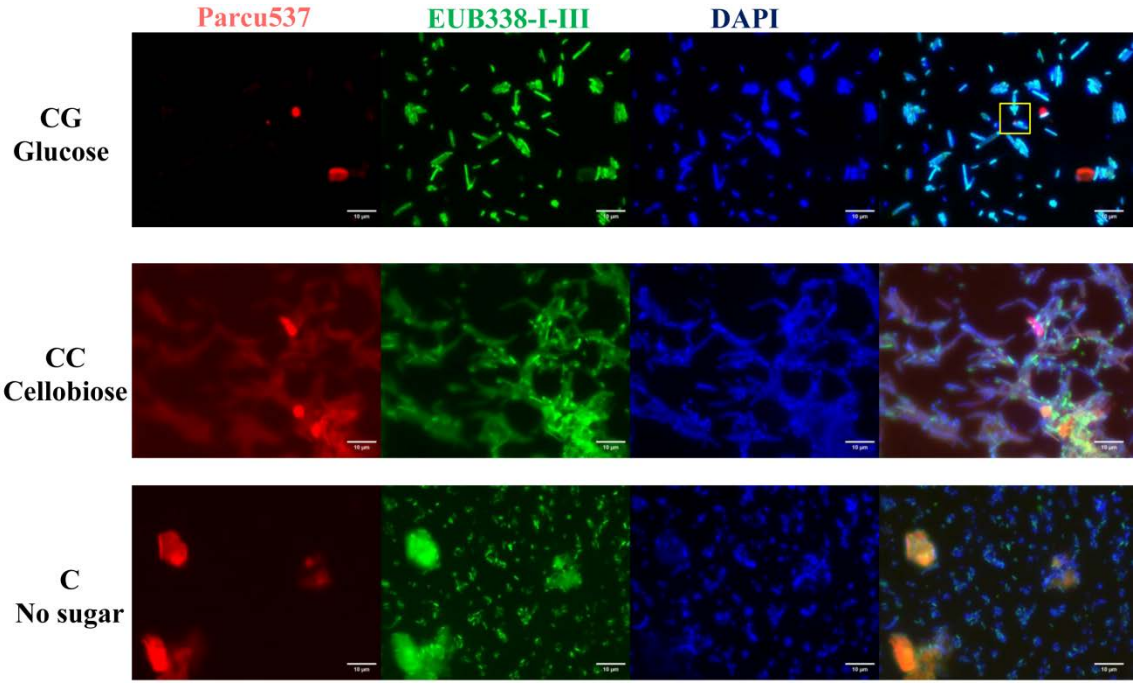

Figure 5 monoFISH images of CS enrichments

PCR for Parcubacteria

qPCR primers specific for Parcubacteria was used to detect their existence. As is shown in Figure 6A, all the samples had bands with primer1 (Bacteria). There are multiple bands (close to each other) in the gel with primer 2 (Parcubacteria), while no band appeared with primer 3 (Parcubacteria). So, enrichment TC and T amplified with primer 2 were sent for

sequencing. After exchanged the forward primers for Parcubacteria, we used primer 4 and 5 to amplify. Enrichment T had band for both primers (Figure 6B) and was sent for sequencing.

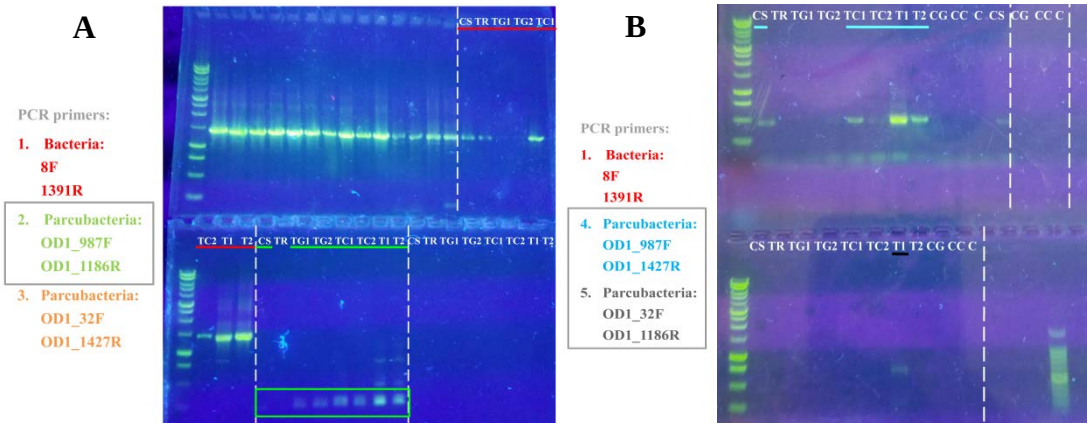

Figure 6 Agarose gel electrophoresis with different primers

qPCR

Primer 4 and universal bacterial primers were used for qPCR. Because of limited time, no standard curve was made. So, the calculation way is the Ct value of the Parcubacteria gene divided by the Ct value of the 16S bacterial control gene (Figure 7).

Parcubacteria: OD1_987F
OD1_1427R

Bacteria: 515F
907R

1~5: 10 times dilution series

Parcubacteria			16S Bacteria			Par/16S	
	CT	Ct Mean		CT	Ct Mean	CT	Ct Mean
A1	34.940	28.063	A6	Undetermined	20.750	--	1.352417
A2	34.519	32.904	A7	26.108	26.142	1.322151	1.258668
A3	33.190	27.743	A8	24.873	20.527	1.334375	1.351544
A4	32.289	25.802	A9	31.417	23.880	1.027758	1.080461
A5	34.650	27.776	A10	20.844	26.535	1.662357	1.046763
B1	24.837	28.063	B6	21.375	20.750	1.161922	1.352417
B2	33.159	32.904	B7	25.230	26.142	1.314235	1.258668
B3	25.510	27.743	B8	18.967	20.527	1.344969	1.351544
B4	22.536	25.802	B9	20.547	23.880	1.09679	1.080461
B5	23.136	27.776	B10	39.402	26.535	0.587179	1.046763
C1	24.413	28.063	C6	20.125	20.750	1.213057	1.352417
C2	31.036	32.904	C7	27.088	26.142	1.145725	1.258668
C3	24.530	27.743	C8	17.742	20.527	1.382644	1.351544
C4	22.580	25.802	C9	19.677	23.880	1.147556	1.080461
C5	25.543	27.776	C10	19.360	26.535	1.319332	1.046763

Figure 7 qPCR results

Summary

The miniproject might provide an approach and some indications of identifying the 'microbial dark matter' Parcubacteria by using specific FISH-probe and PCR primers. To date, there is no isolation of Parcubacteria. It is a challenging work. Metagenomic analysis is the key to know more about the energy and material metabolism, which is significant for

designing targeted primers and probes, enriching and isolating Parcubacteria.

Acknowledgements

The author would like to thank everyone in the Microbial Diversity Course 2017. Dianne Newman and Jared Leadbetter, the course directors, gave me a lot of good suggestions on the miniproject design; Kurt Hanselmann helped me sampling and designing media; Scott Dawson provided the new ideas and designed primers for me; Gabriela Kovacikova was very concerned about my life and health; Viola Krukenberg gave me the strongest supports on FISH and the whole miniproject; Matthew Tien, Kyle Costa and Adalid Pelayo helped me with PCR and qPCR; Lisa Cohen assisted with the informatics; Scott Saunders, Alicja Dabrowska and Bingran Cheng assisted with the chemical analysis. I also would like to thank my supervisor Dr. Bo-zhong Mu in East China University of Science and Technology, Dr. Ji-dong Gu in The University of Hong Kong, and Dr. Joseph M. Suflita in University of Oklahoma, who gave me important support on my academic research and life. Finally, I want to thank all the funding for the course.

References

[1] Solden L, Lloyd K, Wrighton K. The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Curr Opin Microbiol*, 2016, 31: 217-226.

[2] Rinke C, Schwientek P, Sczyrba A, *et al*. Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, 2013, 499: 431-437.

[3] Hug LA, Baker BJ, Anantharaman K, *et al*. A new view of the tree of life. *Nature Microbiology,* 2016, 1, doi:10.1038/nmicrobiol.2016.48.

[4] Nelson WC, Stegen JC. The reduced genomes of Parcubacteria (OD1) contain signatures of a symbiotic lifestyle. *Front Microbiol*, 2015, 6, doi: 10.3389/fmicb.2015.00713.

[5] Castelle CJ, Brown CT, Thomas BC, *et al*. Unusual respiratory capacity and nitrogen metabolism in a Parcubacterium (OD1) of the Candidate Phyla Radiation. *Scientific Report*, 2017, 7, doi:10.1038/srep40101.

[6] Luo F, Devine CE, Edwards EA. Cultivating microbial dark matter in benzene-degrading methanogenic consortia. *Environ Microbiol*, 2016, 18: 2923-2936.

[7] Turner S, Pryer KM, Miao VPW, *et al*. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J Eukaryot Microbiol*, 1999, 46: 327-338.

[8] Lane DJ. 16S/23S rRNA sequencing. In Nucleic Acid Techniques in Bacterial Systematic. Edited by Stackebrandt and Goodfellow M. John Wiley and Sons, New York, 1991: 115-175.