

下水処理場の活性汚泥内に存在する *Patescibacteria* 門のメタゲノム解析

広島大学院先進理工系科学研究科 藤井直樹、大橋晶良、尾崎則篤、金田一智規

1. 研究背景

現在多くの下水処理場では活性汚泥法をはじめとする生物処理法が用いられており、活性汚泥内の細菌を中心とした微生物の働きによって下水中の有機物や栄養塩類の除去が行われている。そういった微生物のなかで単離や培養ができるものは全体の 15%未満であり¹⁾、そのなかでも地球上の細菌を門レベルで分類すると約 1/4 が *Patescibacteria* に属する²⁾。*Patescibacteria* は培養された例がほぼ無い門レベルで未分類の細菌群であり、海洋水、淡水、バイオリアクター、海底熱水などのさまざまな環境に存在している。また、ゲノムのサイズが小さく、生命維持に必要な生合成能力が制限されている。東広島浄化センターの活性汚泥内にも *Patescibacteria* が数種類存在することが既往の研究³⁾⁴⁾により分かっており、これらの細菌を初めとする微生物の生理機能の研究は、微生物の増殖メカニズムの解明につながり、活性汚泥法の効率化につながると考える。

2. 目的

東広島浄化センターで採取した活性汚泥内に存在する細菌群 *Patescibacteria* 門に属する細菌の内、存在割合が高い上位3種である *Gracilibacteria*、*Saccharimonadia*、*Parcubacteria* のゲノムをメタゲノム解析によって再構築する。

3. 方法

3-1. サンプル収集、DNA 抽出

2019年2月6日、2020年4月13日に東広島浄化センターで活性汚泥を採取し、冷凍保存したものをサンプル1、サンプル2とした。また、4月13日に採取した活性汚泥から100 mLを滅菌したバイアル瓶に入れて気相部を窒素置換し、20°Cで3日間静置した後に冷凍保存したものをサンプル3とした。

サンプル1、2、3をFast DNA SPIN Kit for Soilを用いてDNAを抽出した。抽出したDNAは、ライブラリーキットTruSeq DNA PCR Free (350)を使用し、Illumina社のHiSeq XおよびMiSeq、Pacific Biosciences社のPacBio Sequel IIを用いて解析した。

3-2. 16S rRNA 遺伝子を対象としたアンプリコン解析による *Patescibacteria* の存在ポテンシャル推定

MiSeqによってDNA断片から解析された塩基配列データを用いてQiime2-2020.2⁵⁾を使用し、*Patescibacteria* の存在ポテンシャル推定を行った。また、EMIRGE⁶⁾を使用してHiSeq Xによって解析された塩基配列のRaw Dataから16S rRNA 遺伝子を再構築、データのポテンシャルを確かめた。さらに、Pacbio CCSのRaw Dataを直接アノテーションし、16S rRNAだけを抽出してデータのポテンシャルを確かめた。

3-3. メタゲノム解析

HiSeq XおよびPacBio Sequel IIによってDNA断片から解析された塩基配列データを用いてメタゲノム解析を行った。メタゲノム解析はTrimmomatic-0.39⁷⁾を用いてアダプター配列やクオリティの低い部分配列を除去するトリミングを行った後、SPAdes-3.13.1⁸⁾を用いて断片的な塩基配列をつなぎ合わせて元の配列を再構築するアセンブルを行った。その後、Maxbin2.2.7⁹⁾を用いて同一微生物由来の配列断片をグループ化するビンニングを行い、ゲノムを再構築した。その際に1000 bp未満のコンティグはビンニングに含めなかった。最後に、CheckM-v1.1.2¹⁰⁾を用いたクオリティチェックと、Prokka1.13¹¹⁾で再構築されたゲノム配列上がどの遺伝子であるかを調べるアノテーションを行った。

キーワード 活性汚泥、*Patescibacteria*、メタゲノム解析、*Gracilibacteria*、*Saccharimonadia*、*Parcubacteria*

連絡先 〒739-8527 広島県東広島市鏡山1-4-1 広島大学院先進理工系科学研究科 環境保全工学研究室

TEL 082-424-5718

4. 結果

4-1. 16S rRNA 遺伝子を対象とした *Patescibacteria* の存在ポテンシャル推定

Qiime2 による解析の結果からサンプル 1~3 について 3 種類の細菌それぞれの存在割合の推定値が得られた。細菌の存在割合はサンプルごとに異なっていたが、サンプル 1、2 から、活性汚泥内での 3 種類の細菌の存在割合は大きく変化していないことが分かった(Table 1)。また Hiseq X、Pacbio の Raw Data には 3 種類の細菌の 16S rRNA 遺伝子が含まれていた(Table 2、Table 3)。16S rRNA 遺伝子は再構築した細菌のゲノムがどの細菌であるかを推定する指標であり、16S rRNA 遺伝子を含むゲノムを再構築することは後段の解析において非常に重要である。したがって、本解析において用意したサンプルには後段のメタゲノム解析で 16S rRNA 遺伝子を含むゲノムを再構築するポテンシャルがあることが確かめられた。

Table 1 Analysis with Qiime2 (%)

Phylum	1	2	3
<i>Gracilibacteria</i>	0.22	1	0.33
<i>Parcubacteria</i>	2.33	3.34	1.32
<i>Saccharimonadia</i>	10.6	9.46	2.09

Table 2 Number of 16S rRNA of Hiseq X (species)

Species	1	2	3
<i>Gracilibacteria</i>	1	1	0
<i>Parcubacteria</i>	1	0	0
<i>Saccharimonadia</i>	9	2	2
Total	154	112	81

Table 3 Number of 16S rRNA of Pacbio CCS (species)

Species	1	2	3
<i>Gracilibacteria</i>	4	2	0
<i>Parcubacteria</i>	13	0	0
<i>Saccharimonadia</i>	15	7	0
Total	245	22	3

4-2. メタゲノム解析

メタゲノム解析によって再構築することができたゲノムについて、Completeness、Contamination、Contig 数を Table 4 に示す。細菌のゲノムは Completeness > 85%、Contamination < 5% のものを高品質とし、Contig 数は少ないほど塩基配列の欠損が少ないことを表す。メタゲノム解析の結果、*Gracilibacteria* のゲノムを 1 つ、*Saccharimonadia* のゲノムを 5 つ、*Parcubacteria* のゲノムを 5 つ再構築することができた。特に、*Gracilibacteria* は高品質のゲノムを再構築することが出来た。一方で *Saccharimonadia* と *Parcubacteria* は複数種のゲノムを再構築することが出来たものの、その多くが Contamination を多く含んでおり品質が低かった。bin.007 は *Proteobacteria* の 16S rRNA を同じビン内に含んでいた。このゲノムは Bin size も他の細菌と比べて 2 倍以上大きくなっており、明らかに他の細菌の遺伝子が混入している可能性が高い。このことから *Saccharimonadia* と *Parcubacteria* に対してビンニングが適切に行えていないことが考えられる。

Table 4 Result of metagenomic analysis

BIN.ID	Taxonomy	Completeness (%)	Contamination (%)	No.contigs	Bin size (Mbp)
bin.150	<i>Gracilibacteria</i>	100	0	6	1.42
bin.004	<i>Saccharimonadia</i>	93.02	4.65	6	0.94
bin.007	<i>Saccharimonadia</i>	95.35	9.3	21	3.58
bin.060	<i>Saccharimonadia</i>	88.37	0	8	0.75
bin.079	<i>Saccharimonadia</i>	97.67	18.6	29	1.42
bin.116	<i>Saccharimonadia</i>	95.35	16.28	20	1.82
bin.032	<i>Parcubacteria</i>	62.79	0	8	0.76
bin.069	<i>Parcubacteria</i>	86.05	6.98	16	1.29
bin.082	<i>Parcubacteria</i>	100	27.91	42	1.75
bin.149	<i>Parcubacteria</i>	20.93	9.3	43	1.38
bin.154	<i>Parcubacteria</i>	90.7	0	25	1.6

5. 結論

メタゲノム解析によって対象とする3種の細菌のゲノムを複数種再構築することができた。特に *Gracilibacteria* については高品質のゲノムを再構築することが出来た。しかし、*Saccharimonadia* と *Parcubacteria* は再構築することができたゲノムのほとんどが Contamination を多く含んでおり品質が低い。そのため今後はビンニング方法の再検討、再解析を行い結果の改善を試みる。そして再解析の結果をもとに機能推定、検証を行う。

参考文献

- 1) Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., et al.: Insights into the phylogeny and coding potential of microbial dark matter, *nature*, 499(7459), pp.431-437, 2013
- 2) Parks, D.H., Rinke, C., Chuvochina, M., Chaumeil, P., Woodcroft, B.J. et al.: Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life, *nature microbiology*, 2, pp.1533-1542, 2017
- 3) 西村恭平, 金田一智規, 大橋晶良, 尾崎則篤: 活性汚泥内に存在する Candidate divisio 細菌群の FISH 法を用いた可視化, 第 68 回土木学会中国支部研究発表会, 7, pp.6, 2016
- 4) 竹中 亮太, 金田一 智規, 大橋 晶良, 尾崎 則篤, 青井 議輝: 活性汚泥内に存在する Candidatus Saccharibacteria(TM7)に対するプライマーの特異性の検証, 第 69 回土木学会中国支部研究発表会, 7, pp.6, 2017
- 5) Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., et al.: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2, *Nature Biotechnology*, 37, pp.852-857, 2019
- 6) Miller, C.S., Baker, B.J., Thomas, B.C., Singer, S.W. and Banfield, J.F.: EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data, *Genome Biology*, 12, pp.1-14, 2011
- 7) Bolger, A.M., Lohse, M. and Usadel, B.: Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics*, 30(15), pp. 2114-2120, 2014
- 8) Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., et al.: SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing, *JOURNAL OF COMPUTATIONAL BIOLOGY*, 19, pp.455-477, 2012
- 9) Wu, Y.W., Simmons, B.A. and Singer, S.W.: MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets, *Bioinformatics*, 32(4), pp.605-607, 2016
- 10) Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P. and Tyson G.W.: CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes, *Genome Research*, 25(7), pp.1043-1055, 2015
- 11) Seemann, T.: Prokka: rapid prokaryotic genome annotation, *Bioinformatics*, 30(14), pp.2068-2069, 2014