"*Candidatus* Accumulibacter" Gene Expressions in Response to Dynamic EBPR Conditions

by Shaomei He and Katherine D. McMahon

Supplementary Materials

Supplementary Material Included:

Supplementary Text 1. RT-qPCR primer design for Accumulibacter clades IA and IIA
Supplementary Text 2. PCR condition optimization and standard curve generation
Supplementary Text 3. Eligibility of absolute quantification of mRNA
Supplementary Text 4. Changes in 16S rRNA abundance across a cycle
Supplementary Text 5. Discussion on observed difference in phaA and phaC regulation

Table S1
Figure S1
Figure S2
Supplementary Reference

RT-qPCR primer design for Accumulibacter clades IA and IIA

Twelve primer sets were designed to target genes of interest from both clades IA and IIA Accumulibacter. The metagenome sequencing recovered a nearly complete genome of clade IIA (binned as "*Candidatus* Accumulibacter phosphatis" on IMG/M [http://img.jgi.doe.gov/cgi-bin/m/main.cgi]), but only retrieved partial sequences from clade IA (Garcia Martin *et al.*, 2006). It was estimated that clades IA and IIA genomes had an average nucleotide sequence divergence of 15% (Kunin *et al.*, 2008). Since these two clades are the closest relatives in the sludge

microbial community, non-clade IIA genes that share the highest DNA sequence identify to clade IIA in the sludge metagenome are likely from clade IA. Table S1 lists the non-clade IIA genes from the sludge metagenome that have the closest match to clade IIA, as well as non-Accumulibacter genes in Genbank that share the highest DNA sequence identity to clade IIA, as a comparison. For every gene studied, the percent identity between clade IIA gene and its homolog from the sludge metagenome is higher than that from the closest non-Accumulibacter match in Genbank. Therefore, these closest homologs of IIA genes in the sludge metagenome are very likely from clade IA.

PCR optimization and standard curve generation

To generate an appropriate positive control for PCR optimization, we selected and pooled a total of 14 DNA samples with varied relative abundances of clades IA and IIA, collected from the reactor during a 5-month period. The samples were mixed to generate a community DNA pool that had a balanced distribution of clades IA and IIA (verified by ppk1-targeted qPCR). A 30-cycle PCR with an annealing temperature gradient range of 58-65^{o}C was performed on the pooled DNA sample to determine the optimal annealing temperatures for the newly designed primer sets. The reaction mixtures contained 1X PCR buffer II, 3.0 mM $MgCl_2$, 200 μM of each dNTP, 400 nM of each forward and reverse primer and 0.05 U/μl of AmpliTaq Gold® DNA polymerase (Applied Biosystems, Foster City, CA). Since the Accumulibacter genome has a high GC content (64%), betaine was added to 0.5 M final concentration as a denaturant to facilitate the reaction. PCR products were visualized by 2% agarose gel electrophoresis. A single sharp band at the expected amplicon size was obtained for each individual primer pair, indicating

that the amplification was specific. The optimal annealing temperature was determined as the highest temperature that still resulted in a high intensity band (Table 1).

Positive controls for individual primer sets were generated by PCR amplification with corresponding primer sets, using the pooled sludge DNA sample described above. PCR products were purified (PureLink^{TM} PCR purification kit ,Invitrogen) and the mass concentration of PCR products was determined (PicoGreen dsDNA quantification kit, Invitrogen). Copy number was calculated based on mass concentration and the average molecular weight of each individual PCR amplicon. Six-point calibration curves for qPCR were produced by tenfold serial dilution of positive controls in duplicate, ranging from 10^3 to 10^8 target copies per reaction. Quantification was performed using the iCycler^{TM} iQ optical system software Version 3.0a (Biorad). PCR amplification efficiencies were estimated from the standard curve slope by the formula $10^{-1/slope}$-1.

Eligibility of absolute quantification of mRNA

Accurate comparison of gene expression between samples using RT-qPCR often requires normalizing unwanted variations introduced during procedures such as RNA extraction and reverse transcription. Usually a housekeeping gene that has constant expression levels among different samples is used as a reference gene or an internal control to quantify the relative change of target genes (referred to as “relative quantification”). However, choosing an appropriate housekeeping gene that does not vary in all samples under different conditions can be very difficult, even in well-studied bacteria, such as *Pseudomonas* and *Staphylococcus* (Savli *et al.*,

2003; Vandecasteele *et al.*, 2001). As we know much less about Accumulibacter, it is more challenging to select such an internal standard.

Quantifications normalized to sample volumes or cell numbers without comparing to a reference gene (referred to as "absolute quantification") have also been applied in some cases, especially when evaluating the stabilities of housekeeping genes (Vandecasteele *et al.*, 2001). In our study, due to the difficulties in identifying a good internal standard, we employed absolute quantification. We first evaluated the variation in RNA extraction and RT-qPCR, since these two steps are expected to be the major sources of variation across samples. Replicated extractions of 8 different samples indicated a good reproducibility of RNA yield, with an average coefficient of variation (CV) of 7%. A subset of these samples with replicated extractions was used in downstream RT-qPCR, with primer sets targeting cbb3-COX, ppk1-IA, ppk1-IIA and 16S rRNA genes. For these primer sets, an average CV of 21% was obtained from the overall analyses. By partitioning the variance from the ANOVA (Analysis of Variance), we found that the RNA extraction and reverse transcription steps contributed to 41% of the total observed variance on average, while the remaining 59% was contributed by qPCR step alone. The results indicated that the variance from RNA extraction and reverse transcription is acceptable, particularly when comparing fold changes. Therefore, in this study, we used quantification normalized to the starting sludge volume.

Trina McMahon 6/14/10 9:54 PM
Deleted: and

Trina McMahon 6/14/10 9:54 PM
Deleted: . The

Trina McMahon 6/14/10 9:54 PM
Deleted: arising from the RT-qPCR step

Trina McMahon 6/14/10 9:54 PM
Deleted: 9% and the average CV resulted

Trina McMahon 6/14/10 9:54 PM
Deleted: analysis (starting from biomass collection

Trina McMahon 6/14/10 9:54 PM
Deleted: transcript quantification)

Trina McMahon 6/14/10 9:54 PM
Deleted: 13%

Trina McMahon 6/14/10 9:54 PM
Deleted: variation

Changes in 16S rRNA abundance across a cycle

As anticipated, the 16S rRNA level changed during an EBPR cycle, since bacteria regulate their ribosome content by the growth rate, stringent response to amino-acid starvation, and other

mechanisms (Nomura *et al.*, 1984). Accumulibacter 16S rRNA levels doubled during the anaerobic phase, and except for a short increase lasting for about 35 min in the beginning of aerobic phase, levels decreased during the remaining aerobic phase. Its anaerobic increase is indeed surprising, because EBPR metabolic models suggest PAO growth only occurs in the aerobic phase, due to the high energy cost, which is more likely to be provided under aerobic conditions. It is not clear whether the increase is simply associated with cells synthesizing a large amount of enzymes necessary for rapid anaerobic acetate-uptake and transformation, or with some other mechanisms. A preliminary study showed that the transcription of relA (the synthase of (p)ppGpp, a global stringent response regulator), was negatively correlated to 16S rRNA during the majority of EBPR cycle (correlation coefficient = -0.83, n = 14) (data not shown). This might indicate some stringent control mechanism during EBPR. Such stringent control could promote cyclical turnover and re-synthesis of ribosomes (including rRNA), leading to the observed but otherwise unanticipated changes in 16S rRNA levels. Further study is needed to elucidate the mechanism.

Observed difference in phaA and phaC regulation

In Accumulibacter, genes phaC (included in our study) and phaE form an operon, encoding the two subunits of the heterodimeric Class III PHA synthase, respectively. Gene phaA (included in our study), encoding acetyl-CoA acetyltransferase, is located in an adjacent operon downstream, in the same transcriptional direction as phaCE. In *Cyanothece sp.* PCC 7425, to which this phaC has the highest similarity, these genes are in a single operon. While in Accumulibacter, two short genes (both encoding MaoC domain protein dehydratases) are in the same operon with phaA,

separating phaA and phaC into two operons. Downstream of the phaA-containing operon, a different phaC gene encoding class I PHA synthase forms a separate transcriptional unit in the opposite direction to phaA. These operon structures are very interesting although the implications for expression in Accumulibacter are not clear. In *Aeromonas caviae*, a similar MaoC domain was also found downstream of phaC, and exhibited (*R*)-specific enoyl-CoA hydratase activity and played a role in channeling (*R*)-3-hydroxyacyl-CoA monomer from β-oxidation of fatty acid to PHA synthesis (Fukui *et al.*, 1998). This might suggest that the phaA investigated in our study might also be involved in fatty acid metabolism, in addition to PHA biosynthesis, thus subjecting to a separate regulation from phaC. The potential relationship between fatty acid and PHA metabolisms is of particular interest since Wilmes *et al.* (2008) proposed a role of fatty acid β-oxidation in EBPR metabolism. Further experimental investigation is needed to elucidate this.

Table S1. Summary of genes with the closest match to clade IIA

Gene	IMG/M OID of IIA	Closest match to IIA in sludge metagenome		Closest match to IIA in Genbank (non-Accumulibacter)	
		OID of non-IIA homolog	ID[1] to IIA (%)	Organism	ID[1] to IIA (%)
Pyruvate dehydrogenase complex	2001099980	2001086930	92.31	Dechloromonas aromatica RCB	81.76
Pyruvate synthase (pyruvate flavodoxin/ferredoxin oxidoreductase)	2001062970	2001144560	92.96	Rhodoferax ferrireducens T118	71.21
Citrate synthase	2001061560	2001103000	94.66	Azoarcus sp. EbN1	84.85
Isocitrate lyase	2001006860	2000000540 2000224990	100 99.22	Dechloromonas aromatica RCB	89.81
Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	2001061540	2000037440	98.3	Dechloromonas aromatica RCB	85.65
Fumarate reductase subunit C	2001024420	2001036520	80.3	Chromobacterium violaceum ATCC 12472	64.39
Methylmalonyl-CoA mutase, N-terminal domain/subunit	2001062840	2000234230 2000052210	90.55 84.09	Nocardioides sp. JS614	65.12
Acetyl-CoA acetyltransferase	2001070320	2001139020	91.11	Chromobacterium violaceum ATCC 12472	64.19
Poly(3-hydroxyalkanoate) synthetase	2001070280	2000319730	99.73	Cyanothece sp. PCC 7425	47.45
Heme/copper-type cytochrome/quinol oxidases, subunit 1	2001010230	2000086280	99.41	Dechloromonas aromatica RCB	87.29
Cbb3-type cytochrome oxidase, subunit 1	2001063240	2001042560	97.4	Azoarcus sp. BH72	87.32
Nitrous oxide reductase	2001002170	2001100970 2000131850	90.79 97.76	Dechloromonas aromatica RCB	90.05

[1] Percentage of DNA sequence identity

Trina McMahon 6/14/10 9:54 PM
Formatted ... [1]
Trina McMahon 6/14/10 9:54 PM
Formatted Table ... [2]
Trina McMahon 6/14/10 9:54 PM
Deleted: Fumarate reductase subunit C
Trina McMahon 6/14/10 9:54 PM
Formatted ... [3]
Trina McMahon 6/14/10 9:54 PM
Deleted: 2001024420… ... [4]
Trina McMahon 6/14/10 9:54 PM
Formatted ... [5]
Trina McMahon 6/14/10 9:54 PM
Formatted Table ... [6]
Trina McMahon 6/14/10 9:54 PM
Formatted ... [7]
Trina McMahon 6/14/10 9:54 PM
Formatted Table ... [8]
Trina McMahon 6/14/10 9:54 PM
Formatted ... [9]
Trina McMahon 6/14/10 9:54 PM
Deleted: Acetyl-CoA acetyltransferase
Trina McMahon 6/14/10 9:54 PM
Formatted ... [10]
Trina McMahon 6/14/10 9:54 PM
Deleted: 2001070320
Trina McMahon 6/14/10 9:54 PM
Formatted ... [11]
Trina McMahon 6/14/10 9:54 PM
Deleted: 2001139020
Trina McMahon 6/14/10 9:54 PM
Formatted ... [12]
Trina McMahon 6/14/10 9:54 PM
Deleted: 91.11
Trina McMahon 6/14/10 9:54 PM
Formatted ... [13]
Trina McMahon 6/14/10 9:54 PM
Deleted: Chromobacterium violaceu ... [14]
Trina McMahon 6/14/10 9:54 PM
Formatted ... [15]
Trina McMahon 6/14/10 9:54 PM
Deleted: 64.19
Trina McMahon 6/14/10 9:54 PM
Formatted ... [16]
Trina McMahon 6/14/10 9:54 PM
Formatted ... [17]
Trina McMahon 6/14/10 9:54 PM
... [18]
Trina McMahon 6/14/10 9:54 PM
... [19]
Trina McMahon 6/14/10 9:54 PM
Formatted Table ... [20]
Trina McMahon 6/14/10 9:54 PM

Supplementary Figure Legend

Figure S1. Transcription profiles during a normal EBPR cycle. The *y*-axis shows the normalized transcript levels by comparison with the first time point. The maximal fold change (MFC) is the ratio of the maximal to minimal cDNA copy number during a cycle.

Figure S2. A heatmap showing the clustering pattern of gene expression during an EBPR cycle, generated from the heatmap function using R software package (R Development Core Team, 2008). The expression levels of each gene were indicated as the ratio to its maximal expression during a cycle in order to be at the same scale.

Supplementary Reference

Fukui, T., Shiomi, N., and Doi, Y. (1998) Expression and characterization of (R)-specific enoyl coenzyme A hydratase involved in polyhydroxyalkanoate biosynthesis by Aeromonas caviae. *J. Bacteriol.* **180**: 667-673.

Garcia Martin, H., Ivanova, N., Kunin, V., Warnecke, F., Barry, K.W., McHardy, A.C., et al. (2006) Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. *Nat. Biotechnol.* **24**: 1263-1269.

Kunin, V., He, S., Warnecke, F., Peterson, S.B., Garcia Martin, H., Haynes, M., et al. (2008) A bacterial metapopulation adapts locally to phage predation despite global dispersal. *Genome Res.* **18**: 293-297.

Nomura, M., Gourse, R., and Baughman, G. (1984) Regulation of the synthesis of ribosomes and ribosomal components. *Annu. Rev. Biochem.* **53**: 75-117.

R Development Core Team (2008) *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria.

Savli H, Karadenizli A, Kolayli F, Gundes S, Ozbek U, Vahaboglu H (2003). Expression stability of six housekeeping genes: a proposal for resistance gene quantification studies of Pseudomonas aeruginosa by real-time quantitative RT-PCR. *J Med Microbiol* **52:** 403-408.

Vandecasteele, S.J., Peetermans, W.E., Merckx, R., and Van Eldere, J. (2001) Quantification of expression of Staphylococcus epidermidis housekeeping genes with Taqman quantitative PCR during in vitro growth and under different conditions. *J. Bacteriol.* **183**: 7094-7101.

Trina McMahon 6/14/10 9:54 PM
Deleted: Garcia Martin H, Ivanova N,

Trina McMahon 6/14/10 9:54 PM
Deleted: ,…,… KW… AC … [21]

Trina McMahon 6/14/10 9:54 PM
Formatted: Font:Not Italic

Trina McMahon 6/14/10 9:54 PM
Formatted: Justified, Level 2

Trina McMahon 6/14/10 9:54 PM
Deleted: …. … [22]

Trina McMahon 6/14/10 9:54 PM
Formatted: Font:Not Bold

Trina McMahon 6/14/10 9:54 PM
Deleted: ,…,…,… SB…, … [23]

Trina McMahon 6/14/10 9:54 PM
Formatted: Font:Not Italic

Trina McMahon 6/14/10 9:54 PM
Deleted: .

Trina McMahon 6/14/10 9:54 PM
Formatted: Font:Not Bold

Trina McMahon 6/14/10 9:54 PM
Deleted: ,…,…. … [24]

Trina McMahon 6/14/10 9:54 PM
Formatted: Font:Not Bold

Trina McMahon 6/14/10 9:54 PM
Deleted: ….: … [25]

Trina McMahon 6/14/10 9:54 PM
Formatted: Justified

Trina McMahon 6/14/10 9:54 PM
Formatted: Justified, Level 2

Trina McMahon 6/14/10 9:54 PM
Deleted: SJ… WE…, …. … [26]

Trina McMahon 6/14/10 9:54 PM
Formatted: Font:Not Bold

Wilmes, P., Andersson, A.F., Lefsrud, M.G., Wexler, M., Shah, M., Zhang, B. et al. (2008) Community proteogenomics highlights microbial strain-variant protein expression within activated sludge performing enhanced biological phosphorus removal. *ISME J.* **2**: 853-864.

Trina McMahon 6/14/10 9:54 PM
Deleted: .

Trina McMahon 6/14/10 9:54 PM
Formatted: Justified, Line spacing: single

Page 7: [1] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Line spacing: 1.5 lines

Page 7: [2] Change	**Trina McMahon**	**6/14/10 9:54 PM**

Formatted Table

Page 7: [3] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Portuguese

Page 7: [4] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

2001024420

Page 7: [4] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

2001036520

Page 7: [4] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

80.3

Page 7: [4] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

Chromobacterium violaceum ATCC 12472

Page 7: [4] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

64.39

Page 7: [5] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:10 pt

Page 7: [6] Change	**Trina McMahon**	**6/14/10 9:54 PM**

Formatted Table

Page 7: [7] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:10 pt

Page 7: [8] Change	**Trina McMahon**	**6/14/10 9:54 PM**

Formatted Table

Page 7: [9] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:10 pt

Page 7: [10] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:3 pt

Page 7: [11] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:3 pt

Page 7: [12] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:3 pt

Page 7: [13] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:3 pt

Page 7: [14] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

Chromobacterium violaceum ATCC 12472

Page 7: [15] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:3 pt

Page 7: [16] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:3 pt

Page 7: [17] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:10 pt

Page 7: [17] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:10 pt

Page 7: [17] Formatted	**Trina McMahon**	**6/14/10 9:54 PM**

Font:10 pt

Page 7: [18] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

Nitrous oxide reductase

Page 7: [18] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

2001002170

Page 7: [18] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

2001100970

Page 7: [18] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

90.79

Page 7: [18] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

Dechloromonas aromatica RCB

Page 7: [18] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

90.05

Page 7: [19] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

2000131850	97.76

Page 7: [20] Change	**Trina McMahon**	**6/14/10 9:54 PM**

Formatted Table

Page 8: [21] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

,

Page 8: [21] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

,

Page 8: [21] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

KW

Page 8: [21] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

AC

Page 8: [22] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

.

Page 8: [22] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

Page 8: [23] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

,

Page 8: [23] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

,

Page 8: [23] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

,

Page 8: [23] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

SB

Page 8: [23] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

,

Page 8: [24] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

,

Page 8: [24] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

,

Page 8: [24] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

.

Page 8: [25] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

.

Page 8: [25] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

:

Page 8: [26] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

SJ

Page 8: [26] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

WE

Page 8: [26] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

,

Page 8: [26] Deleted	**Trina McMahon**	**6/14/10 9:54 PM**

.