The ISME Journal (2014) 8, 2353–2356

www.nature.com/ismej

COMMENTARY

Re-analysis of omics data indicates *Smithella* may degrade alkanes by addition to fumarate under methanogenic conditions

Boonfei Tan, Camilla Nesbø and Julia Foght

The ISME Journal (2014) **8,** 2353–2356; doi:10.1038/ismej.2014.87; published online 27 May 2014

Here we comment on three aspects of a recent report in *ISME J* (Embree *et al.*, 2013), which concluded that: the draft genome of *Smithella* ME-1 lacks genes encoding alkylsuccinate synthase (ASS) subunits; it therefore does not express those genes when grown on *n*-hexadecane in mixed culture under methanogenic conditions; and it may use an alternate mechanism to initiate anaerobic *n*-hexadecane degradation. We have re-analyzed the draft genome of *Smithella* ME-1 and three metatranscriptomes, and reached conclusions opposite to those of Embree *et al.* (2013).

Biodegradation of *n*-alkanes is well-understood under aerobic conditions and pathways are being elucidated under nitrate- and sulfate-reducing conditions. In contrast, under methanogenic conditions complete *n*-alkane degradation requires a consortium of Archaea and Bacteria; most of the latter have not yet been isolated, hence the bacterial metabolic pathways are largely unknown (Aitken *et al.*, 2013; Callaghan, 2013). Understanding methanogenic *n*-alkane degradation is crucial for optimizing bioremediation in anaerobic environments, modeling formation of heavily biodegraded petroleum in oil reservoirs and potentially for converting unrecovered residual hydrocarbons in petroleum reservoirs to methane for recovery as a valuable fuel.

The most widely reported mechanism for anaerobic *n*-alkane degradation is initiated by hydrocarbon addition to fumarate. The initial hydrocarbon-activating reaction under nitrate- and sulfate-reducing conditions is catalyzed by the glycyl radical enzyme ASS, encoded by the *assABC* genes (Callaghan, 2013). Owing to high sequence similarity and conservation of the *assA* gene encoding the α-subunit of ASS, its presence has been used as a diagnostic marker to indicate the genetic capability for alkane activation by addition to fumarate (Callaghan, 2013). Under methanogenic conditions several studies have implicated *Syntrophus* spp. and/or the phylogenetically related but largely uncultivated genus *Smithella* as primary *n*-alkane degraders owing to their abundance in enrichment cultures (Zengler *et al.*, 1999; Gray *et al.*, 2011; Cheng *et al.*, 2013) and environments impacted by petroleum hydrocarbons (Gray *et al.*, 2011). The alkane-activating mechanism(s), however, remain cryptic because signature alkylsuccinate metabolites have not been detected *in situ* or in cultures under methanogenic conditions, and *assA* genes have not been assigned to either genus (Aitken *et al.*, 2013; Callaghan, 2013). Thus, alternative alkane activation mechanisms have been proposed (Aitken *et al.*, 2013) but not demonstrated.

Embree *et al.* (2013) recently published the draft genome of *Smithella* sp. ME-1 (DDBJ/EMBL/GenBank accession number AWGX00000000) obtained from an *n*-hexadecane-degrading methanogenic enrichment culture. The authors used fluorescence-activated cell sorting to separate six bacterial cells related to *Smithella* from the community, amplified the total DNA using whole-genome multiple displacement amplification, sequenced using Illumina Hi-seq (Illumina, San Diego, CA, USA), assembled using a *de novo* co-assembler and annotated the draft genome using the RAST server (Embree *et al.*, 2013). Furthermore, they obtained metatranscriptomes from enrichment cultures grown on hexadecane, butyric acid or caprylic acid as the only organic carbon source. RAST failed to detect and annotate *assABC* genes in the draft *Smithella* genome; consequently, mapping of metatranscripts from the enrichment cultures to the draft genome did not reveal expression of *ass* genes in any of the three cultures. Embree *et al.* (2013) therefore concluded that *Smithella* is incapable of *n*-alkane activation by addition to fumarate, even though they observed transcription of genes homologous to those encoding glycyl radical activating enzymes in hexadecane-degraders and expression of fatty acid utilization genes required for β-oxidation of *n*-alkanes subsequent to activation.

These conclusions ran counter to hypotheses mentioned above implicating *Smithella/Syntrophus* in alkane degradation, as well as our own results, and prompted closer investigation of the *Smithella* draft genome and metatranscriptomes. We recently analyzed the metagenome of a methanogenic short-chain alkane-degrading enrichment culture (SCADC; Tan *et al.*, 2013) and recovered a partial

Smithella genome in which we detected a single copy of *assA* (KF824850; unpublished). We used this sequence plus several annotated *assA* genes (for example, Callaghan *et al.*, 2012) for tblastn screening of the *Smithella* ME-1 draft genome and therein found seven genes encoding glycyl radical enzymes, including a nearly full-length putative *assA* gene on contig 5325 (accession number AWGX01000974; gene length 2584 bp) that was not annotated in the *Smithella* ME-1 draft genome (Embree *et al.*, 2013); a putative *assA* gene fragment on contig 9960 (AWGX01000099; gene length 235 bp); and five putative pyruvate formate lyase (PFL) genes on contigs 6993, 13305, 13440, 7458 and 4758 (AWGX01000042, −380, −095, −531 and −777; gene lengths 1643–2543 bp). The truncated putative *assA* gene sequence (−099) was too short to confidently assign function and therefore was eliminated from further analysis.

We then translated the near full-length putative *assA* sequences from *Smithella* ME-1, *Smithella* SCADC and five reference *assA* sequences. Pairwise comparison (Supplementary Table S1) showed that the putative AssA proteins in *Smithella* ME-1 (AWGX01000974) and *Smithella* SCADC (KF824850) had high amino acid identity to each other (87%) and to known AssA in *Azoarcus sp.* HxN1 (CAO03074), *Aromatoleum* sp. OcN1 (CBK27727), *Desulfoglaeba alkanexedens* ALDC (ADJ51097) and to both assA copies in *Desulfatibacillum alkenivorans* Ak-01 (ABH11460 and ABH11461) (61%–69% over >800 amino acids). Sequence alignments of putative AssA from *Smithella* ME-1 (AWGX01000974) and *Smithella* SCADC (KF824850) were robust over the full length of the reference AssA sequences (Figure 1). Moreover, when the translated *assA* sequence from *Smithella* ME-1 was used in BLASTP searches against the NCBI nr-database, the top 100 hits yielded AssA homologs detected in bacteria and enrichment cultures, many of which were associated with methanogenic alkane degradation and all of which had significant bit scores and e-values; lower hits corresponded to benzyl succinate synthase α-subunit (BssA) but not to PFL (results not shown).

Because sequence alignment alone is inadequate for annotating gene sequences, we subjected the *assA* gene in *Smithella* ME-1 to phylogenetic analysis (Figure 2), revealing that it is closely related to known *assA* genes in cultivated bacteria and to the *assA* gene annotated in *Smithella* SCADC but distantly related to other glycyl radical genes including *bssA* and *pfl* (Figure 2). It is likely that Embree *et al.* (2013) inadvertently failed to detect *assA* because it was not annotated by the automated RAST server, possibly owing to the short contig length (<3 kb); however, *assA* was readily detected using manual sequence similarity searches. Thus, Embree *et al.* (2013) have actually provided the first genetic evidence directly linking *Smithella* to fumarate activation of *n*-alkanes, rather than documenting its absence.

We also re-analyzed transcription of this newly annotated *assA* gene in mixed cultures grown with

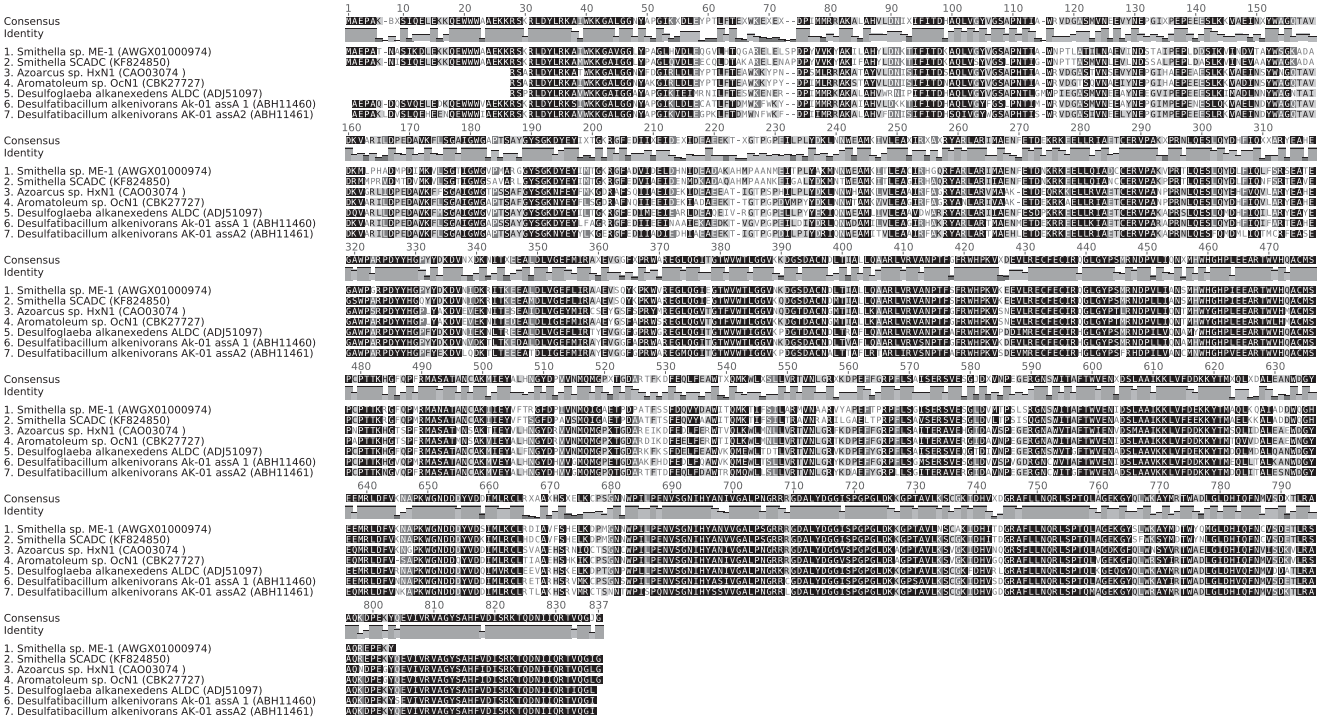

Figure 1 Sequence alignment of translated putative *assA* in *Smithella* spp. with reference AssA sequences. Sequences were aligned using Muscle 3.3 in Geneious R7, with conserved sequence motifs highlighted in black. Pairwise percentage identity of the nucleotide alignment is shown in Supplementary Table S1.

n-hexadecane, butyric acid or caprylic acid (Embree *et al.*, 2013). We mapped the metatranscripts (GSE49830) to the six near full-length glycyl radical gene homologs detected in *Smithella* ME-1 (Figure 2) using CLC Genomics Workbench (CLC Bio, Aarhus, Denmark), with two mismatches allowed per read alignment. In doing so, we observed that 1822 metatranscriptome reads mapped to the *assA* gene (AWGX01000974) under hexadecane conditions but only 232 reads mapped under caprylate growth conditions and 2 reads with growth on butyrate (see Supplementary Figure S1 for relative expression). This result contrasts with the report by Embree *et al.* (2013) that there was no expression of AWGX01000974 during growth on hexadecane, because it was not annotated by RAST. The very low expression of AWGX01000974 under caprylate conditions (Supplementary Figure S1) correlates with the low abundance of *Smithella* ME-1 in the caprylate culture versus overwhelming dominance in the hexadecane culture (Figure 4b in Embree *et al.*, 2013), supporting our contention that *assA* expression correlates with growth of *Smithella* ME-1 on hexadecane in a methanogenic mixed culture. Further supporting our proposal, Embree *et al.* (2013) detected genes encoding α-methylacyl-coA racemase and methyl-malonyl-coA that are proposed to be involved in epimerization and carbon skeleton rearrangement of metabolic intermediates, respectively, in the proposed fumarate activation pathway used under nitrate- and sulfate-reducing conditions (Callaghan *et al.*, 2012; Jarling *et al.*, 2012). Genes for β-oxidation of fatty acids were also highly transcribed during growth on hexadecane (Embree *et al.*, 2013), consistent with utilization of *n*-hexadecane via a fumarate activation pathway. Expression of putative PFL genes on contigs AWGX01000042, −531 and −777 (Supplementary Figure S1) during growth on hexadecane likely reflects conversion of pyruvate to acetyl-coA and formate (Lu *et al.*, 2012), a process common in methanogenic substrate degradation.

Based upon our re-analysis of the *Smithella* ME-1 draft genome and metatranscriptomes by manual curation using tblastn rather than automated RAST annotation of fumarate addition genes, we reach a conclusion opposite to that of Embree *et al.* (2013). We propose instead that *Smithella* is indeed genetically capable of activating and utilizing long-chain alkanes like *n*-hexadecane under methanogenic conditions via addition to fumarate, by virtue of possessing and expressing *ass* genes during methanogenic growth on *n*-hexadecane. We further note that automated annotation pipelines like RAST

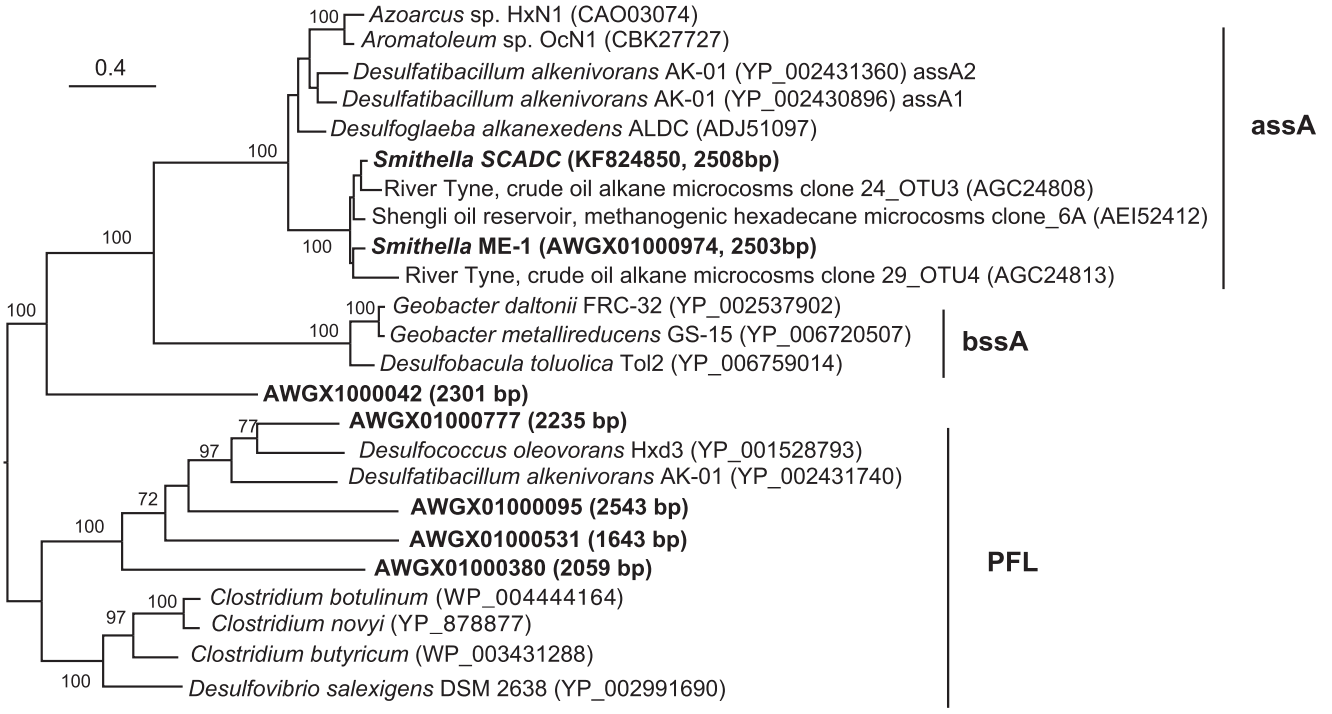

Figure 2 Maximum likelihood tree of six translated glycyl radical enzyme genes in the draft *Smithella* ME-1 genome (Embree *et al.*, 2013) recovered through tblastn searches (shown in boldface, with the SCADC *assA* sequence); sequence length is shown in parentheses. Closely related sequences were recovered from the NCBI nr-database through BLASTX searches and all sequences were aligned using MAFFT, followed by phylogenetic tree construction using PhyML (Guindon *et al.*, 2010) with LG model and 100 bootstrap replicates in Geneious R7 (www.geneious.com). Bootstrap support ⩾70% is indicated. The tree was rooted by midpoint rooting. The *assA* sequences from clones (indicated on tree) were not full length and ranged from 414–662 bp. All other sequences used in the tree were full length (>2400 bp). A tree with the same overall topology was obtained when including only full-length sequences and removing gaps (not shown).

have sometimes resulted in misannotation of fumarate addition genes even in well-characterized organisms. For example, the *bssA* gene sequence in *Geobacter daltonii FRC-32* (NC_011979.1) is annotated as formate C-acetyltransferase (Geob_2448). Similarly, in our own work the *Smithella* SCADC *assA* gene was misannotated by RAST as PFL (EC 2.3.1.54). Therefore, manual curation of sequences is necessary for accurate identification of genes such as those involved in anaerobic hydrocarbon activation.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

This work was supported by Genome Canada and Genome Alberta through the Hydrocarbon Metagenomics Project (*http://www.hydrocarbonmetagenomics.com/*).

B Tan, C Nesbø and J Foght are at Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada
C Nesbø is at CEES, Department of Biology, University of Oslo, Oslo, Norway
E-mail: julia.foght@ualberta.ca

References

Aitken CM, Jones DM, Maguire MJ, Gray ND, Sherry A, Bowler BFJ *et al.* (2013). Evidence that crude oil alkane activation proceeds by different mechanisms under sulfate-reducing and methanogenic conditions. *Geochim Cosmochim Acta* **109**: 162–174.

Callaghan AV, Morris BEL, Pereira IAC, McInerney MJ, Austin RN, Groves JT *et al.* (2012). The genome sequence of *Desulfatibacillum alkenivorans* AK-01: a blueprint for anaerobic alkane oxidation. *Environ Microbiol* **14**: 101–113.

Callaghan AV. (2013). Enzymes involved in the anaerobic oxidation of *n*-alkanes: from methane to long-chain paraffins. *Front Microbiol* **4**: 89.

Cheng L, Ding C, Li Q, He Q, Dai LR, Zhang H. (2013). DNA-SIP reveals that *Syntrophaceae* play an important role in methanogenic hexadecane degradation. *PLoS One* **8**: e66784.

Embree M, Nagarajan H, Movahedi N, Chitsaz H, Zengler K. (2013). Single-cell genome and metatranscriptome sequencing reveal metabolic interactions of an alkane-degrading methanogenic community. *ISME J* **8**: 757–767.

Gray ND, Sherry A, Grant RJ, Rowan AK, Hubert CRJ, Callbeck CM *et al.* (2011). The quantitative significance of Syntrophaceae and syntrophic partnerships in methanogenic degradation of crude oil alkanes. *Environ Microbiol* **13**: 2957–2975.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *System Biol* **59**: 307–321.

Jarling R, Sadeghi M, Drozdowska M, Lahme S, Buckel W, Rabus R *et al.* (2012). Stereochemical investigations reveal the mechanism of the bacterial activation of *n*-alkanes without oxygen. *Angew Chem Int Ed Engl* **51**: 1334–1338.

Lu W, Du J, Schwarzer NJ, Gerbig-Smentek E, Einsle O, Andrade SLA. (2012). The formate channel FocA exports the products of mixed-acid fermentation. *Proc Nat Acad Sci USA* **109**: 13254–13259.

Tan B, Dong XL, Sensen CW, Foght JM. (2013). Metagenomic analysis of an anaerobic alkane-degrading microbial culture: potential hydrocarbon-activating pathways and inferred roles of community members. *Genome* **56**: 599–611.

Zengler K, Richnow HH, Rossello-Mora R, Michaelis W, Widdel F. (1999). Methane formation from long-chain alkanes by anaerobic microorganisms. *Nature* **401**: 266–269.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)