J Ind Microbiol Biotechnol (2013) 40:517–521
DOI 10.1007/s10295-013-1244-z

SHORT COMMUNICATION

Detection of a novel active transposable element in *Caldicellulosiruptor hydrothermalis* and a new search for elements in this genus

Daehwan Chung · Joel Farkas · Janet Westpheling

Received: 11 January 2013 / Accepted: 12 February 2013 / Published online: 10 March 2013

Abstract We show that a previously annotated hypothetical protein is the transposase of a new and active IS element, IS*Cahy1*, widespread in *Caldicellulosiruptor* species. Transposition generated an 11-bp direct repeat at the insertion site in *Caldicellulosiruptor hydrothermalis*, suggesting a cut-and-paste mechanism. The discovery of an active insertion sequence in *Caldicellulosiruptor* species led to a survey of potential IS elements in the genome sequences of eight *Caldicellulosiruptor* species that identified several new elements, including one novel to this genus.

Keywords Transposition · Thermophilic anaerobes · *Caldicellulosiruptor* · IS element · IS*Cahy1* · IS*Cbe4*

Introduction

Members of the genus *Caldicellulosiruptor* are the most thermophilic cellulolytic bacteria known and have become the target of intense interest because of their potential for biomass conversion to biofuels and bioproducts [2, 20]. We recently reported the development of the first method for genetic manipulation of members of this genus and showed that restriction of heterologous DNA is a major barrier to DNA transformation [6]. During the course of isolating mutants that facilitate genetic analysis, we discovered an active insertion sequence (IS) element in *Caldicellulosiruptor hydrothermalis*. The element, designated IS*Cahy1*, is widespread among *Caldicellulosiruptor* species and this is the first demonstrated to be an active element in a member of this genus.

IS elements are mobile genetic elements that may mediate their own transposition. They are widely distributed phylogenetically, and occur in nearly all prokaryotic genera but they are not always obvious, especially in relatively uncharacterized bacteria. They have been shown to be involved in genomic rearrangement and horizontal gene transfer in prokaryotes and recent genome sequencing projects of cellulolytic thermophilic bacteria have identified new IS elements in Gram-positive bacteria [13, 19]. *Clostridium thermocellum* contains nearly 100 full-length IS elements [9], and at least one, IS*1447*, has been shown to be active [22]. *Clostridium cellulolyticum* also has an active IS element [12]. Other species, including *Thermoanaerobacter tengcongensis*, *Thermotoga maritima*, and *Caldicellulosiruptor*, have predicted IS elements in their genomes, though they have not been shown to be active. Here we show that an IS element, IS*Cahy1*, is an active element and that a previously annotated hypothetical protein in *Caldicellulosiruptor hydrothermalis* is the transposase of this active element. This finding prompted us to revisit the analysis of the genomes of several sequenced *Caldicellulosiruptor* species using several bioinformatics tools including ISsaga, a Web-based computational tool for IS annotation [18], and we identified a novel IS element apparently unique to this genus.

Electronic supplementary material The online version of this article (doi:10.1007/s10295-013-1244-z) contains supplementary material, which is available to authorized users.

D. Chung · J. Farkas · J. Westpheling (✉)
Department of Genetics, University of Georgia, Athens, GA 30602, USA
e-mail: janwest@uga.edu

D. Chung · J. Farkas · J. Westpheling
The BioEnergy Science Center, Department of Energy, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Table 1 *Caldicellulosiruptor hydrothermalis* strains used in this study

Strains	Description/phenotype	Source
JWCH001	*C. hydrothermalis 108* wild type (Ura^+/5-FOA^S)	DSMZ[a]
JWCH003	IS*Cahy1* insertion at nucleotide +396 in *pyrF* (Calhy_1352)/ (Ura^-/5-FOA^R)	This study [5]
JWCH004	Revertant of JWCH003 to uracil prototrophy (Ura^+/5-FOA^S)	This study

[a] German Collection of Microorganisms and Cell Cultures (DSMZ) (Braunschweig, Germany)

Materials and methods

Caldicellulosiruptor hydrothermalis strains used in this study are listed in Table 1. All strains were grown anaerobically in liquid or solid medium in low osmolarity defined (LOD) growth medium [8] with maltose as the sole carbon source at 68 °C. For growth of auxotrophic mutant JWCH003, the defined medium contained 40 μM uracil. Chromosomal DNA from *C. hydrothermalis* strains was extracted using the Quick-gDNA MiniPrep (Zymo) according to the manufacturer's instructions. The spontaneous uracil auxotrophic mutant JWCH003 and its revertant JWCH004 (Table 1) were characterized using PCR amplification and DNA sequencing. JWCH004 was isolated by spreading the overnight JWCH003 cultures onto LOD solid medium and selecting uracil prototrophy at 68 °C [8]. The reversion rate was calculated as the number of uracil prototrophic colonies per 10^9 cells. The insertion and excision of IS*Cahy1* was verified by DNA sequencing (Macrogen, Rockville, MD, USA) of products generated using primers FJ298 and JH020. DNA sequences of the primers used in this study are listed in Table S1. To produce an alignment and phylogenetic tree of 33 IS elements amino acid sequences, we used ClustalW, version 2 [11], which is based on the neighbor-joining (NJ) method. The tree was visualized with TreeView [14]. Bioinformatic analysis was performed using BLASTn [21], BLASTx [1], ISfinder (http://www-is.biotoul.fr/) [16], ISSaga (http://issaga.biotoul.fr/ISsaga/issaga_index.php) [18], and RepeatScout (http://bix.ucsd.edu/repeatscout/) [15].

Results and discussion

In experiments to select spontaneous mutants of *C. hydrothermalis* resistant to 5-Fluoroorotic acid (5-FOA), we isolated a mutant, JWCH003 (Table 1) [5], which was a uracil auxotroph (loss of uracil biosynthesis results in resistance to 5-FOA) and showed some reversion when plated on LOD medium lacking uracil. PCR amplification and sequencing of the *pyrF* gene from this mutant revealed an 1,832-bp insertion encoding a single open reading frame (Calhy0044) of 476 amino acids (Fig. 1, Figure S1). A BLASTn [21] search of the *C. hydrothermalis* genome showed that this sequence occurs 23 times in nearly identical copies, annotated as a hypothetical protein. A BLASTx [1] search of the IS finder database [16] revealed that the protein has 52 % amino acid similarity to the transposase of IS*Cb1* from *Clostridium beijerinckii*

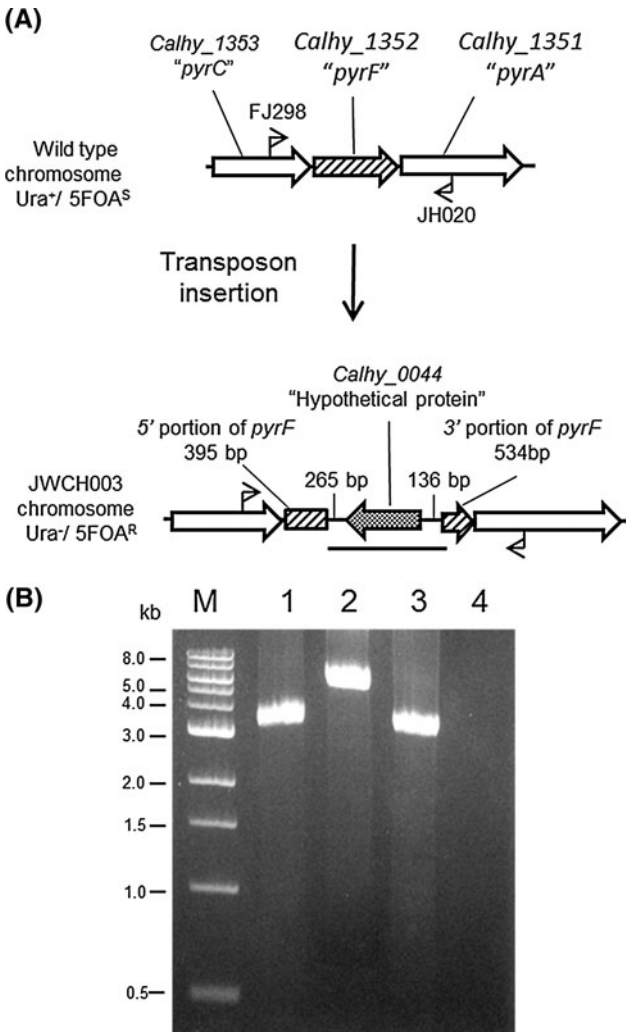

Fig. 1 Confirmation of IS element insertion and excision within the *pyrF* (Calhy_1352) open reading frame in *C. hydrothermalis*. **a** Diagram of the *pyrF* locus in the *C. hydrothermalis* chromosome. *The line below the diagram* depicts the extent of the IS element insertion (1,832 bp) in JWCH003. The insertion sequences are shown in Figure S1. *Bent arrows* depict primers used for amplification of the *pyrF* region in the wild-type, IS element insertion mutant, and revertant strains. **b** Gel separation of PCR products of the *pyrF* genome region in wild-type, 3.4 kb (*lane 1*), the IS element insertion strain JWCH003, 5.3 kb (*lane 2*) and the revertant strain JWCH004, 3.4 kb (*lane 3*) using primers FJ298 and JH020. *M* 1-kb DNA ladder (NEB), *lane 4* no template control

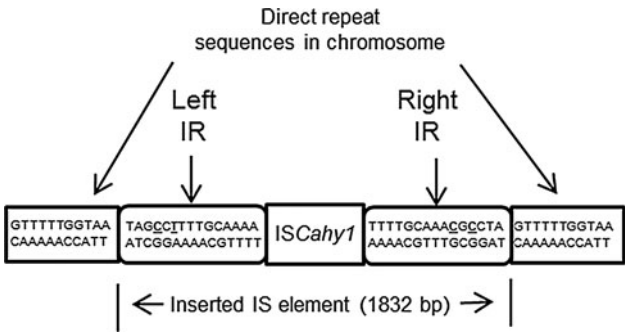

Fig. 2 Diagram of the insertion of IS*Cahy1* into the *C. hydrothermalis* chromosome. Inverted repeat (IR) sequences contain two mismatches (*underlined*). The duplicated target site is indicated at the ends. The sequence of the entire IS element is shown in Figure S1

suggesting that it encodes a transposase. The *pyrF* insertion is flanked by perfect 11-bp direct repeats (DRs) and is delineated by 15-bp inverted repeat (IR) sequences displaying two mismatches (Fig. 2). We named this element IS*Cahy1*. It also shows significant amino acid sequence homology (~35 %) to several known transposases in *Thermoanaerobacter* and *Clostridium* species (Fig. 3). All of the putative IS*Cahy1* elements contain similar length IRs at their ends and were flanked by DRs ranging from 10 to 21 bp in length, likely resulting from the insertion event (Table S2). It is classified as a member of the "ISNCY" family (Table 2), and contains a DDE domain at C-terminus, typical of some families of transposases. BLAST and IS finder searches showed that this element is widely distributed in all eight sequenced *Caldicellulosiruptor* species (Table 2).

One uracil prototrophic revertant of JWCH003 (JWCH004) was analyzed by PCR amplification revealing that the IS element had excised from the *pyrF* gene, restoring wild-type function and uracil prototrophy (Fig. 1; Table 1). PCR amplification of the *pyrF* region and sequencing of the PCR product from the revertant showed an insertion within the *pyrF* gene of the 15 bp left IR sequence contained in the IS element (Fig. 2). This is consistent with a cut-and-paste type mechanism typical of DNA transposable elements. We analyzed two additional independent JWCH003 revertants that showed the same excision site sequence as JWCH004, and the reversion rate was calculated as $\sim 10^{-7}$. Attempts to use Southern hybridization analysis of the mutant to detect IS movement with the transposase as probe (data not shown) were uninformative, as there are 23 copies of the transposase in the genome in *C. hydrothermalis*.

The identification of an active, unannotated transposable element in the *C. hydrothermalis* genome prompted us to scan other *Caldicellulosiruptor* genomes for transposable elements. Annotation of the eight published

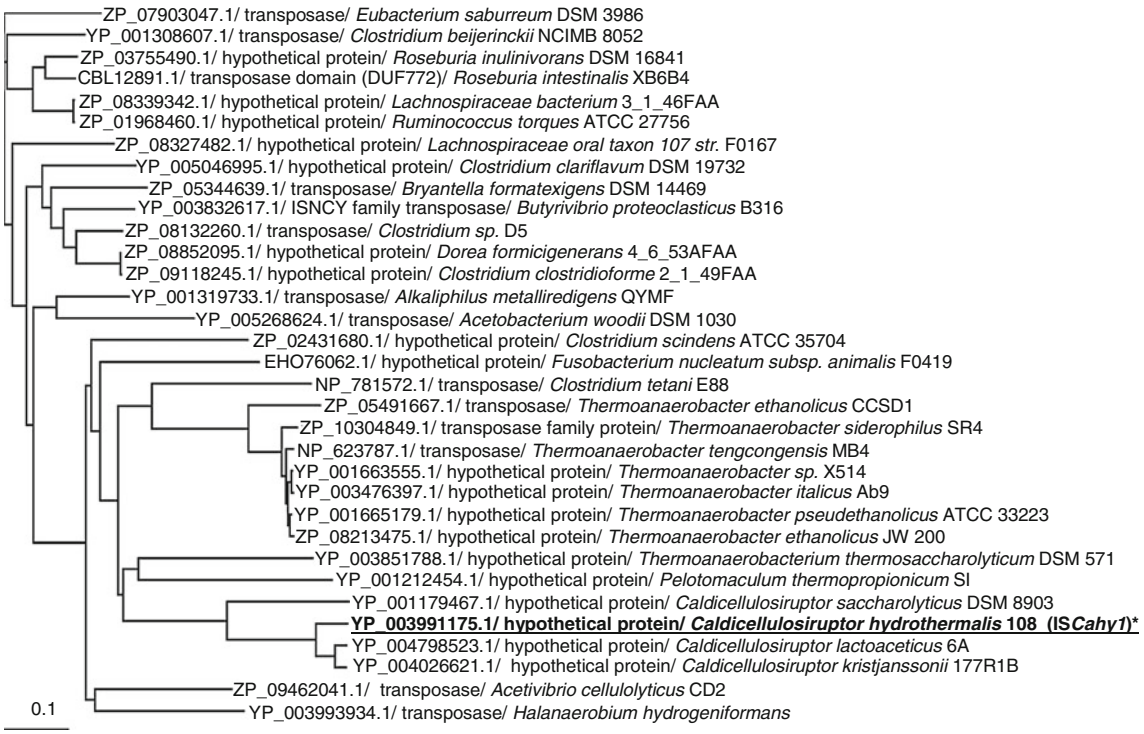

Fig. 3 Phylogram alignment of 33 IS elements sharing significant amino-acid sequence homology with IS*Cahy1*. The GenBank accession number, annotation, and host organism for each IS element are indicated. The distance scale is indicated by a *bar* defining the distance for 0.1-amino-acid substitution per site. The *bold underline* represents the IS*Cahy1* in *C. hydrothermalis*

Table 2 Predicted IS elements in *Caldicellulosiruptor* species

IS families or groups	*C. bescii*	*C. hydrothermalis*	*C. kristjanssonii*	*C. kronotsyensis*	*C. lactoaceticus*	*C. obsidiansis*	*C. owensensis*	*C. saccharolyticus*
IS110	0/1/6/1/4/8	5/0/1/0/2/6	8/0/4/2/6/14	0/0/3/0/3/3	22/1/20/0/7/43	5/1/6/0/3/12	0/0/2/0/2/2	3/10/3/3/4/19
IS1182	0/1/0/0/1/1	0/0/1/0/1/1	4/1/2/0/4/7	1/0/0/0/1/1	1/1/4/0/3/6	0/0/2/1/1/3	4/0/3/3/2/10	0/3/1/0/4/4
IS1634	0/0/3/0/1/3	0/1/0/0/1/1	7/0/5/2/3/14	3/1/1/0/2/5	2/2/3/1/3/8	0/0/0/1/1/1	0/0/1/1/2/2	2/1/0/0/1/3
IS200_IS605	2/10/3/1/9/16	2/0/1/4/6/7	3/1/0/2/4/6	0/1/2/4/5/7	5/1/0/4/4/10	3/2/0/5/5/10	1/0/8/0/4/9	2/1/5/3/7/11
IS200_IS605_ssgr_IS1341	1/0/0/2/3/3	0/0/0/1/1/1	1/0/0/0/1/1	0/0/0/1/1/1	0/0/1/0/1/1	0/0/1/1/1/2	0/0/0/0/0/0	0/0/1/2/2/3
IS21	5/0/3/5/2/5	0/0/0/0/0/0	0/0/0/0/0/0	2/0/3/2/1/5	5/0/2/1/2/4	1/1/2/1/3/4	0/0/0/0/0/0	11/10/2/1/2/17
IS256	3/1/20/0/4/24	0/2/9/0/9/11	6/4/24/0/8/34	0/1/14/0/9/15	3/2/25/0/5/30	2/3/14/0/2/19	2/1/4/0/3/7	27/15/13/2/19/57
IS3_ssgr_IS150	0/0/0/0/0/0	0/0/0/0/0/0	0/0/0/0/0/0	0/0/0/0/0/0	0/0/0/0/0/0	0/0/0/0/0/0	0/0/0/0/0/0	0/0/1/0/1/1
IS30	0/1/0/0/1/1	0/0/2/0/2/2	11/0/5/0/1/16	1/0/2/0/2/3	5/1/2/2/1/10	0/1/1/0/2/2	0/0/1/0/1/1	0/1/1/1/3/3
IS481	5/5/1/0/5/11	0/0/1/0/1/1	3/3/3/0/2/9	0/1/2/0/3/3	3/3/3/0/3/9	3/1/2/0/3/6	0/0/4/0/3/4	18/9/6/1/6/34
IS607	1/0/1/0/1/2	2/0/0/0/1/1	5/0/1/0/1/4	1/0/0/0/1/1	1/0/1/0/1/1	0/0/1/0/1/1	0/0/0/0/0/0	1/2/3/0/4/4
ISL3	0/2/3/0/5/5	0/0/2/0/1/2	0/1/5/0/5/6	0/1/0/3/1/4	0/0/5/0/4/5	0/1/0/2/3/3	0/0/2/1/2/3	1/1/0/1/3/3
ISNCY[a]	1/4/11/1/5/17	9/0/2/0/2/11	2/0/2/0/3/4	0/1/6/0/4/7	4/0/2/0/4/6	4/0/9/0/3/13	0/0/4/0/2/4	13/7/2/0/6/22
ISNCY_ssgr_ISPlu15	0/0/0/10/9/10	0/0/0/6/6/6	0/0/0/2/2/2	0/0/2/3/5/5	0/0/0/2/2/2	0/0/0/4/3/4	0/0/1/5/5/6	0/0/0/10/8/10
Total IS elements	18/25/51/20/50/106	18/3/19/11/33/50	50/10/51/8/40/117	8/6/35/13/38/60	51/11/68/10/40/135	18/10/38/15/31/80	7/1/30/10/26/48	78/60/38/24/70/191

Complete copy/partial copy/pseudogene/unknown/# IS types/total IS elements

complete copy, when the DNA sequence is full length with the two ends of the IS and a complete transposase; partial copy IS, when a part of the sequence is missing or when the transposase gene is mutated; pseudogene, when the transposase gene is mutated (STOP codon in frame—unexpected frameshift); unknown, when the ISsaga predictor is not able to determine if the ORF is complete or partial; # IS types, the number of specific IS element types in the classified IS family

[a] IS*CahyI* belongs to this IS family/group. ISNCY stands for IS Not Classified Yet

Caldicellulosiruptor genomes [3, 7, 10, 17] with ISSaga [18] predicted that all genomes contain several families of IS elements. Most IS element types are common to several, if not all, *Caldicellulosiruptor* species (Table 2). *C. saccharolyticus* and *C. owensensis* appear to be the most unique in terms of number and type of IS elements present. *C. saccharolyticus* contains the most IS types, complete elements, and total number of IS elements, including one IS3_ssgr_IS150 family element (apparently a pseudogene) that is not found in any other species. In contrast, *C. owensensis* has the fewest, and is completely lacking three families of IS elements (IS200_IS605_ssgr_IS1341, IS21, and IS3_ssgr_IS150) that are common in the genus (Table 2).

To further identify transposable elements in *Caldicellulosiruptor* genomes, we used RepeatScout [15], which identifies repetitive sequences. This method identified several IS elements already known or predicted to be transposases, as well as several repetitive protein domains such as cellulases and carbohydrate binding domains. We also identified a putative transposable element (IS*Cbe4*, 1,610 bp), present in *C. bescii* and *C. kronotskyensis* with partial copies in *C. owensensis* and *C. saccharolyticus*. This element shares little DNA sequence identity, but the 479-amino-acid protein shares 62 % similarity to IS*Tte3* from *Thermoanaerobacter tengcongensis*. There are no terminal IRs. The element also showed some similarity to a number of proteins from the newly defined IS*Lre2* family of IS elements [4]. The protein seems to be widespread in a number of Gram-positive bacteria, and is frequently annotated as Uncharacterized Protein Family 0236 (UPF0236). This sequence occurs seven times as complete, identical copies, and five times as partial copies in the genome of *C. bescii*. Most complete copies are flanked by 9-bp DRs. Taken together, these data support the notion that this is a novel IS element belonging to the IS*Lre2* family. The potential activity of this element is the subject of ongoing study.

The identification of an active transposable element in *C. hydrothermalis*, and a putative novel element in other *Caldicellulosiruptor* species, represents an opportunity for the development of new genetic tools as well as the study of new transposable elements types. It also emphasizes the need to be aware that transposition may be induced when generating and characterizing mutants in this genus as well as many others. Given the prevalence of IS elements in several other Gram-positive thermophiles, this caution also applies to a number of cellulolytic clostridia. As shown in Table 2, *C. saccharolyticus* contains nearly 200 IS elements. *C. hydrothermalis* and *C. owensensis* have fewer elements and depending on how many of these elements are active, the number may be a consideration for genetic manipulation and its affect on genome stability.

The number and diversity of IS elements with *Caldicellulosiruptor* species may also have implications for genome dynamics and lateral gene transfer. As shown in Table 2, most of the IS element types found in *Caldicellulosiruptor* species are shared in the genus. The identification of so many IS elements, annotated as hypothetical

proteins, also shows the need for genome annotation methods that give special consideration to repetitive DNA sequences. The transposase is often predicted as an ORF, but intervening sequence and the IRs and DRs are ignored [16]. Partial IS element sequences are often annotated based on protein homology, which can also be misleading when deciphering genomic data.

Acknowledgments We thank Lee Lynd, Dan Olson, and Adam Guss for sharing unpublished results, Jenna Oberstaller for assistance in genome sequence analysis, and Minseok Cha and Jennifer Copeland for helpful discussions during the course of the work. This work was supported by The BioEnergy Science Center supported by a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. JF was supported in part by a predoctoral Graduate Training In Genetics grant (NIH 5T32GM007103-30) to the Genetics Department of the University of Georgia.

References

1. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389–3402
2. Blumer-Schuette SE, Kataeva I, Westpheling J, Adams MW, Kelly RM (2008) Extremely thermophilic microorganisms for biomass conversion: status and prospects. Curr Opin Biotechnol 19(3):210–217. doi:10.1016/j.copbio.2008.04.007
3. Blumer-Schuette SE, Ozdemir I, Mistry D, Lucas S, Lapidus A, Cheng JF, Goodwin LA, Pitluck S, Land ML, Hauser LJ, Woyke T, Mikhailova N, Pati A, Kyrpides NC, Ivanova N, Detter JC, Walston-Davenport K, Han S, Adams MW, Kelly RM (2011) Complete genome sequences for the anaerobic, extremely thermophilic plant biomass-degrading bacteria *Caldicellulosiruptor hydrothermalis*, *Caldicellulosiruptor kristjanssonii*, *Caldicellulosiruptor kronotskyensis*, *Caldicellulosiruptor owensensis*, and *Caldicellulosiruptor lactoaceticus*. J Bacteriol 193(6):1483–1484. doi:10.1128/JB.01515-10
4. Brochet M, Da Cunha V, Couve E, Rusniok C, Trieu-Cuot P, Glaser P (2009) Atypical association of DDE transposition with conjugation specifies a new family of mobile elements. Mol Microbiol 71(4):948–959. doi:10.1111/j.1365-2958.2008.06579.x
5. Chung D, Cha M, Farkas J, Westpheling J (2013) Construction of a stable replicating shuttle vector for *Caldicellulosiruptor* species: use for extending genetic methodologies to other members of this genus. Plos One (in review)
6. Chung D, Farkas J, Huddleston JR, Olivar E, Westpheling J (2012) Methylation by a unique α-class N4-cytosine methyltransferase is required for DNA transformation of *Caldicellulosiruptor bescii* DSM6725. PLoS One 7(8):e43844. doi:10.1371/journal.pone.0043844
7. Elkins JG, Lochner A, Hamilton-Brehm SD, Davenport KW, Podar M, Brown SD, Land ML, Hauser LJ, Klingeman DM, Raman B, Goodwin LA, Tapia R, Meincke LJ, Detter JC, Bruce DC, Han CS, Palumbo AV, Cottingham RW, Keller M, Graham DE (2010) Complete genome sequence of the cellulolytic thermophile *Caldicellulosiruptor obsidiansis* OB47T. J Bacteriol 192(22):6099–6100. doi:10.1128/JB.00950-10
8. Farkas J, Chung D, Cha M, Copeland J, Grayeski P, Westpheling J (2013) Improved growth media and culture techniques for genetic analysis and assessment of biomass utilization by *Caldicellulosiruptor bescii*. J Ind Microbiol Biotechnol 40(1):41–49. doi:10.1007/s10295-012-1202-1
9. Feinberg L, Foden J, Barrett T, Davenport KW, Bruce D, Detter C, Tapia R, Han C, Lapidus A, Lucas S, Cheng JF, Pitluck S, Woyke T, Ivanova N, Mikhailova N, Land M, Hauser L, Argyros DA, Goodwin L, Hogsett D, Caiazza N (2011) Complete genome sequence of the cellulolytic thermophile *Clostridium thermocellum* DSM1313. J Bacteriol 193(11):2906–2907. doi:10.1128/JB.00322-11
10. Kataeva IA, Yang SJ, Dam P, Poole FL 2nd, Yin Y, Zhou F, Chou WC, Xu Y, Goodwin L, Sims DR, Detter JC, Hauser LJ, Westpheling J, Adams MW (2009) Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "*Anaerocellum thermophilum*" DSM 6725. J Bacteriol 191(11):3760–3761. doi:10.1128/JB.00256-09
11. Larkin MA, Larkin MA, Blackshields G, Brown NP, Chenna R, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG, Higgins DG (2007) ClustalW and ClustalX version 2.0. Bioinformatics 23(21):2947–2948. doi:10.1093/bioinformatics/btm404
12. Maamar H, de Philip P, Belaich JP, Tardif C (2003) ISCce1 and ISCce2, two novel insertion sequences in *Clostridium cellulolyticum*. J Bacteriol 185(3):714–725
13. Mahillon J, Chandler M (1998) Insertion sequences. Microbiol Mol Biol Rev 62(3):725–774
14. Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12(4):357–358
15. Price AL, Jones NC, Pevzner PA (2005) De novo identification of repeat families in large genomes. Bioinformatics 21(Suppl 1):i351–i358. doi:10.1093/bioinformatics/bti1018
16. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M (2006) ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res 34(Database issue):D32–D36. doi:10.1093/nar/gkj014
17. van de Werken HJ, Verhaart MR, VanFossen AL, Willquist K, Lewis DL, Nichols JD, Goorissen HP, Mongodin EF, Nelson KE, van Niel EW, Stams AJ, Ward DE, de Vos WM, van der Oost J, Kelly RM, Kengen SW (2008) Hydrogenomics of the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. Appl Environ Microbiol 74(21):6720–6729. doi:10.1128/AEM.00968-08
18. Varani AM, Siguier P, Gourbeyre E, Charneau V, Chandler M (2011) ISsaga is an ensemble of Web-based methods for high-throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. Genome Biol 12(3):R30. doi:10.1186/gb-2011-12-3-r30
19. Wagner A, Lewis C, Bichsel M (2007) A survey of bacterial insertion sequences using IScan. Nucleic Acids Res 35(16):5284–5293. doi:10.1093/nar/gkm597
20. Yang SJ, Kataeva I, Hamilton-Brehm SD, Engle NL, Tschaplinski TJ, Doeppke C, Davis M, Westpheling J, Adams MW (2009) Efficient degradation of lignocellulosic plant biomass, without pretreatment, by the thermophilic anaerobe "*Anaerocellum thermophilum*" DSM 6725. Appl Environ Microbiol 75(14):4762–4769. doi:10.1128/AEM.00236-09
21. Zhang Z, Schwartz S, Wagner L, Miller W (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7(1–2):203–214. doi:10.1089/10665270050081478
22. Zverlov VV, Klupp M, Krauss J, Schwarz WH (2008) Mutations in the scaffoldin gene, cipA, of *Clostridium thermocellum* with impaired cellulosome formation and cellulose hydrolysis: insertions of a new transposable element, IS*1447*, and implications for cellulase synergism on crystalline cellulose. J Bacteriol 190(12):4321–4327. doi:10.1128/JB.00097-08