SUPPLEMENTARY INFORMATION

Identification of a haloalkaliphilic and thermostable cellulase with improved ionic liquid tolerance

Tao Zhang[1], Supratim Datta[2], Jerry Eichler[3], Natalia Ivanova[1], Seth D. Axen[1], Cheryl A. Kerfeld[1, 4], Feng Chen[1], Nikos Kyrpides[1], Philip Hugenholtz[1,2,5], Jan-Fang Cheng[1], Kenneth L. Sale[2], Blake Simmons[2*] and Eddy Rubin[1*]

(1) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598; (2) The Joint BioEnergy Institute, Emeryville, CA 94608; (3) Dept. of Life Sciences, Ben Gurion University of the Negev, P.O. Box 653, Beersheva 84105, Israel; (4) Department of Plant and Microbial Biology, UC Berkeley, CA 94704; (5) Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Queensland 4072, Australia..

*Correspondences should be addressed to Eddy Rubin (EMRubin@lbl.gov) and Blake Simmons (basimmo@sandia.gov).

Methods

Assay for determine substrate specificity

The substrate specificity of *Hu*-CBH1 was determined by assaying against pNPC, pNPX, pNPG, pNPM and Mannan (Sigma, MO). The enzymatic reactions (100 µl), containing 4 mM final substrate concentration in 100 mM phosphate buffer (pH 9.5) and 2 M sodium chloride, were incubated for 30 minutes at 37°C. The amount of p-nitrophenol released was measured at 405 nm, after the reaction had been quenched by

adding 2 M sodium carbonate, using a molar extinction coefficient of 18,000 $M^{-1}cm^{-1}$. The amount of sugars released when using mannan as substrate was determined by the DNS assay.

Sequence alignment

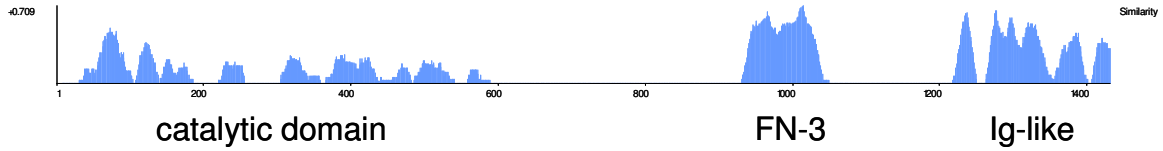
Protein sequences of the deduced products of the cellulaseenriched gene cluster were aligned by ClustalW2 using default parameters (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

Supplementary Table 1.

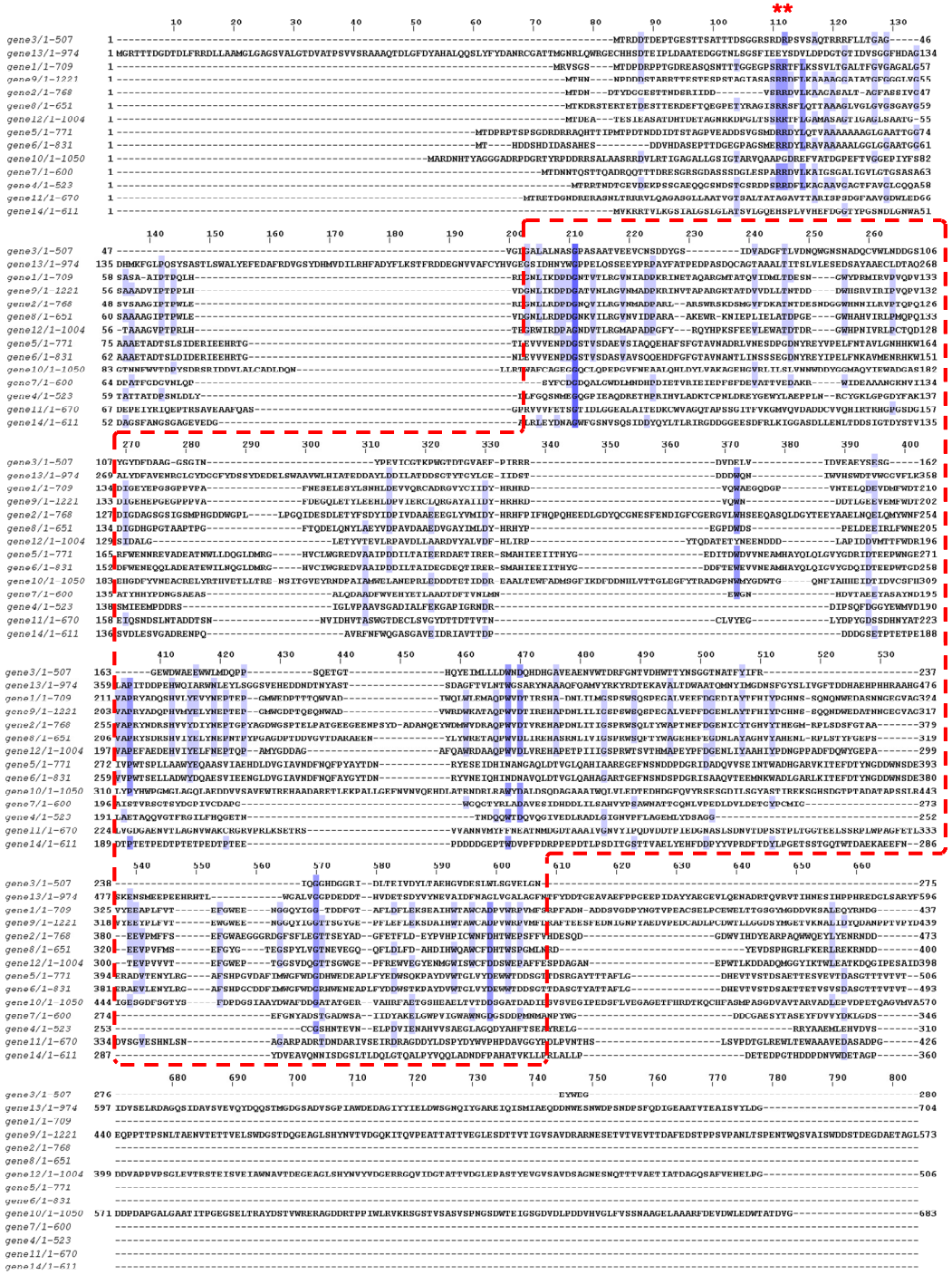
Polysaccharide substrates	Total Activity (mM of glucose released)
Mannan	ND
Carboxymethyl Cellulose	0.85
p-nitrophenyl-beta-D-glucoside	ND
p-nitrophenyl-beta-D-cellobioside	0.69
p-nitrophenyl- beta-D-xyloside	ND
p-nitrophenyl- beta-D-mannoside	ND

Different polysaccharide substrates were used to determine substrate specificity of the enzyme. *Hu*-BCH1 was only reactive to carboxymethyl cellulose and p-nitrophenyl-beta-D-cellobioside, suggesting that the enzyme is a cellobiohydrolase.

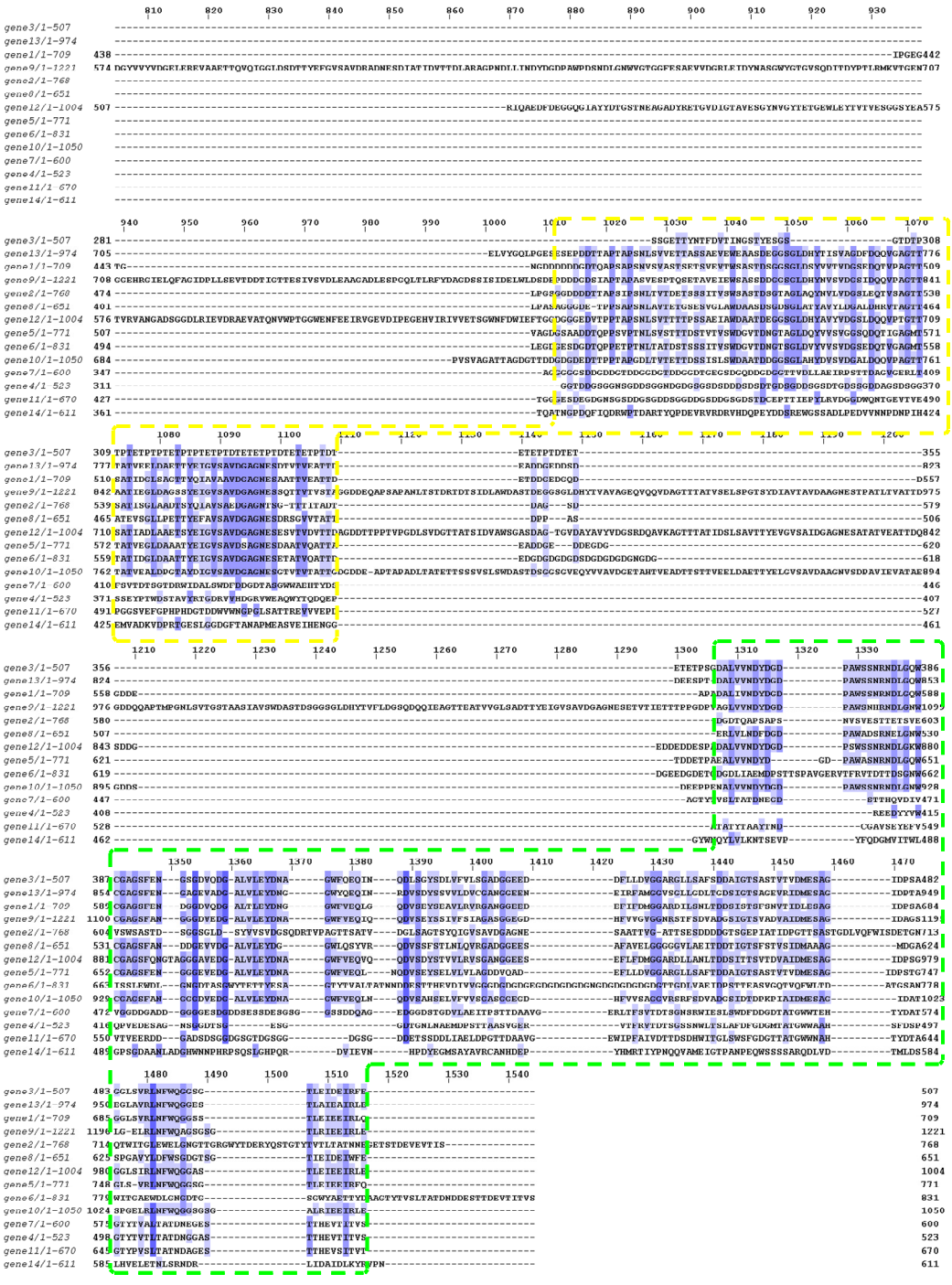
Supplementary Figure 1. There are 3 conserved domains among the 14 genes of the cellulolytic gene cluster in *Halorhabdus utahensis*. Deduced protein sequences of genes 1 to 14 were aligned by ClustalW. Gaps were allowed. The similarity between the aligned sequences was calculated using a 14 amino acid sliding windows across the entire gene product sequences. Three conserved regions were identified, namely, a catalytic domain, a fibronectin-3 domain (FN-3) and an Ig-like domain.



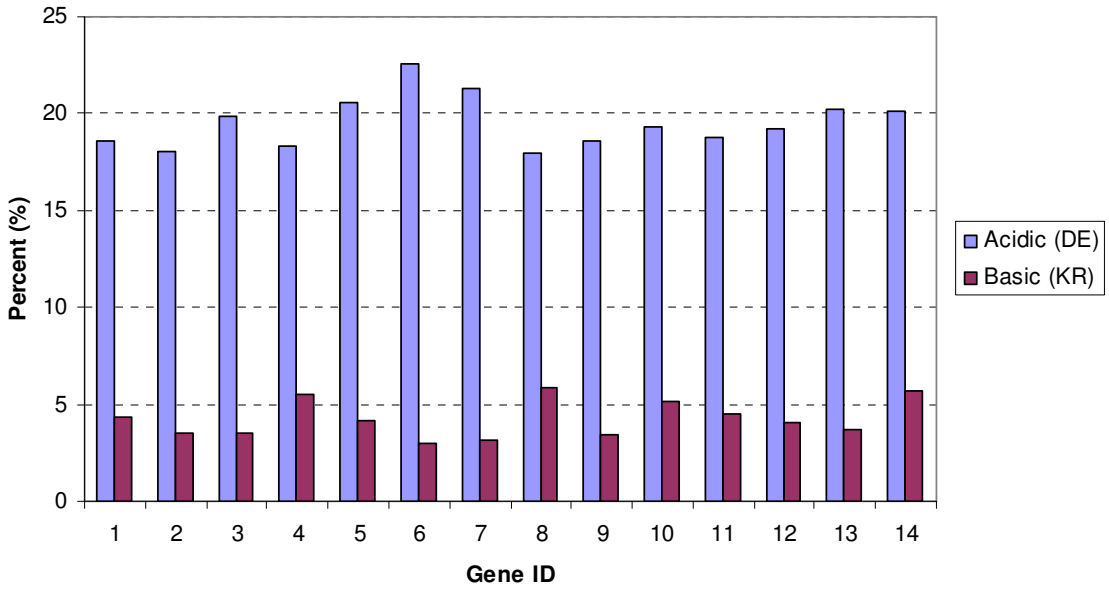
Supplementary Figure 2. The alignment of amino acids sequences of putative glycosylhydrolase gene products in the cellulolytic gene cluster of *Halorhabdus utahensis* revealed conserved motif and domains. Conserved amino acids are highlighted with blue shading. The location of the double arginine signature of a Tat secretion pathway signal motif is marked by double *asterisk*s above the sequences. Catalytic, Fibronectin III (FN3) and Ig-like domains are highlighted by red, yellow and green coloured dash line enclosed regions, respectively.



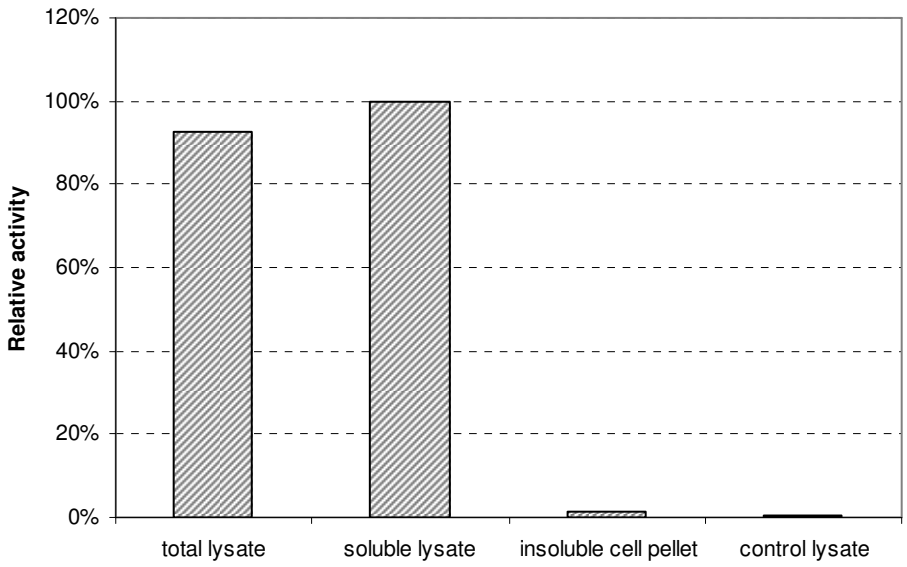
Supplementary Figure 2 (continued).



Supplementary Figure 3. Halophilic proteins encoded by the cellulolytic gene cluster are highly enriched with acidic amino acids (Aspartate or D, and Glutamate or E), but significantly deprived of basic amino acids (Arginine or R, and Lysine or L). The percentages of acidic and basic amino acids are shown in blue and red columns, respectively. Gene-1 is identical to *Hu*-CBH1.



Supplementary Figure 4. *Hu*-CBH1 was present in soluble fraction of cell lysate. Recombinant protein was expressed in *Haloferax volcanii* cells. The cells were pelleted, resuspened in lysis buffer (2 M NaCl and 10 mM Tris-HCl (pH 7.0)) and lyzed by sonication. The soluble and insoluble lysate fractions were separated by centrifugation. Control lysate was prepared using cells transformed with the expression vector only. Equal amounts of total lysate from *Hu*-CBH1-expressing and control cells, and equal proportion of soluble and insoluble fractions of *Hu*-CBH1 lysate were used as enzyme source in a CMC assay performed in 2 M NaCl and 10 mM Tris-HCl (pH 7.0), at 37 °C for 1 hour. For comparison, the activity of the soluble fraction of *Hu*-CBH1 lysate was set as 100%.



Supplementary Figure 5. *Hu*-CBH1 and other 5 known alkaliphilic cellulases are enriched with acidic amino acids. The percentages of acidic and basic amino acids are shown in blue and red bars, respectively. Accession numbers for Cel5A of *B. agaradhaerens*, Cel5A of *Vibrio so. G21*, glucanase of *Bacillus sp.*, glucanase of *Bacillus sp. KSM-64* and alkaline cellulase of *Bacillus sp. KSM-S237* are O85465, ADJ93836, P19424, AAA73189 and JC7532, respectively.