Deletion of Cel48S from *Clostridium thermocellum*

Daniel G. Olson\(^{b,c}\), Shital A. Tripathi\(^{a,c}\), Richard J. Giannone\(^{d}\), Jonathan Lo\(^{b,c}\), Nicky C. Caiazza\(^{a,c}\), David A. Hogsett\(^{a,c}\), Robert Hettich\(^{c,d}\), Adam M. Guiss\(^{b,c}\), Genia Dubrovsky\(^{b,c}\), and Lee R. Lynd\(^{b,c,e}\)

\(^{a}\)Mascoma Corporation, NH 03766; \(^{b}\)Hayer School of Engineering and \(^{c}\)Department of Biological Sciences, Dartmouth College, NH 03755; and \(^{d}\)BioEnergy Science Center, \(^{e}\)Oak Ridge National Laboratory, TN 37830

Edited\(^{*}\) by Lonnie O’Neal Ingram, University of Florida, Gainesville, FL, and approved August 16, 2010 (received for review April 9, 2010)

*Clostridium thermocellum* is a thermophilic anaerobic bacterium that rapidly solubilizes cellulose with the aid of a multienzyme cellulose complex. Creation of knockout mutants for Cel48S (also known as CelS, S\(_5\), and S\(_8\)), the most abundant cellulose subunit, was undertaken to gain insight into its role in enzymatic and microbial cellulose solubilization. Cultures of the Cel48S deletion mutant (S mutant) were able to completely solubilize 10 g/L crystalline cellulose. The cellulose hydrolysis rate of the S mutant strain was 60% lower than the parent strain, with the S mutant strain also exhibiting a 40% reduction in cell yield. The cellulose produced by the S mutant strain was purified by affinity digestion, characterized enzymatically, and found to have a 35% lower specific activity on Avicel. The composition of the purified cellulose was analyzed by tandem mass spectrometry with APEX quantification and no significant changes in abundance were observed in any of the major (>1% of cellulosomal protein) enzymatic subunits. Although most cellulolytic bacteria have one family 48 cellulase, *C. thermocellum* has two, Cel48S and Cel48Y. Cellulose solubilization by a Cel48S and Cel48Y double knockout was essentially the same as that of the Cel48S single knockout. Our results indicate that solubilization of crystalline cellulose by *C. thermocellum* can proceed to completion without expression of a family 48 cellulase.

**Results**

**S Mutant Strain Construction.** Plasmid pDGO-01 (*SI Appendix, Fig. S1*) was transformed into the parent strain (*SI Appendix, Table S1*) by electroporation using recently described methods, and, after overnight recovery, cells harboring the plasmid were selected by the addition of thiamphenicol (Tm) (*Fig. L4*). Cells were subcultured into media containing both 5-fluoroorotic acid (FOA) and Tm to select for cells where integration of the chloramphenicol acetyl transferase (cat) gene had replaced the cel48S gene on the chromosome (*Fig. 1B*). This also selected for loss of the pDGO-01 plasmid. To confirm deletion of cel48S, clones resistant to Tm and FOA were screened at the cel48S locus using diagnostic PCR, showing a 4.6-kb amplicon for the cel48S region and a 3.6-kb amplicon for the Δcel48S::PspFH-cat region (*Fig. 1A*). Additional PCR reactions showed the presence of the cat gene and the absence of two internal fragments of cel48S in the S mutant strain (*SI Appendix, Figs. S2 and S3*). The amplicon in the S mutant strain (*SI Appendix, Table S1*) was sequenced and the cel48S gene was found to have been replaced by the cat cassette from plasmid pDGO-01 (*SI Appendix, Dataset S1*).

**SY Mutant Strain Construction.** Plasmid pJL2 was transformed into the S mutant strain by electroporation as above, but selection was performed using neomycin (Neo). Selection on media containing Neo and FOA resulted in a strain in which integration of the neomycin resistance (neo) gene had replaced the cel48Y gene.

**Author contributions:** D.G.O., S.A.T., R.J.G., J.L., N.C.C., D.A.H., R.H., A.M.G., and L.R.L. analyzed data; and D.G.O. and L.R.L. wrote the paper.

**Conflict of interest statement:** Several of the authors are employees or hold a consulting position with the Mascoma Corporation, which has a financial interest in the organism described here.

**Data deposition:** The sequence reported in this paper has been deposited in the GenBank database (accession no. [S1](#)).

**Funding:** This Direct Submission article had a prearranged editor.

**Freely available online through the PNAS open access option.**

\(^{*}\)To whom correspondence should be addressed: E-mail: lee.lynd@dartmouth.edu.

This article contains supporting information online at www.pnas.org/cgi/doi/10.1073/pnas.1003584107.

\(^{1}\)Embassy of the United States, Beirut, Lebanon.

\(^{2}\)Embassy of the United States, Tunis, Tunisia.
The slope of the linear part of the graph (0.05–0.2 mg) shows that the cellulose from the S mutant strain has a substantially lower specific activity (0.60 ± 0.04 U/mg protein) than purified cellulosomes from either the WT (1.00 ± 0.07 U/mg protein) or parent strain (0.95 ± 0.09 U/mg protein). Extrapolating from the saturation region of the graph (0.4–0.8 mg), the saturation rate can be estimated to be 0.4 µmol/min for the parent and WT strains and 0.2 µmol/min for the S mutant strain (Fig. 3).

Proteomic Characterization. Affinity digestion (18) was used to purify cellulosomes collected at the end of Avicel fermentations: total protein concentration was 1.47 ± 0.06 mg/mL for the WT, 1.56 ± 0.09 mg/mL for the parent, and 1.22 ± 0.14 mg/mL for the S mutant. The WT and parent strains exhibited a reproducible banding pattern typical of the C. thermocellum cellulosome (Fig. 4) (13), with 14 major bands denoted S1–S14 (6). As expected, the S mutant strain was similar to the WT and parent strains except for the S8 band (~80 kDa), which was much fainter and of a slightly higher molecular weight (Fig. 4). The S8 band has been shown to correspond to the protein Cel48S (19). To determine the identity of the ~80-kDa band, it was excised from a gel and characterized by mass spectrometry (NextGen Sciences). The band from the WT and parent strains was found to be a mixture of Cel48S and Cel9Q at an approximately 2:1 ratio (SI Appendix, Table S2). The band from the S mutant strain was exclusively Cel9Q and no Cel48S was detected (SI Appendix, Table S2).

The purified cellulose was further analyzed by tandem mass spectrometry to determine if the composition of the cellulosome had changed in response to the deletion of Cel48S. The most significant change detected was the disappearance of Cel48S in the S mutant strain. In addition to Cel48S, there were four proteins whose abundance had changed significantly (P < 0.01) between the parent and S mutant strains (SI Appendix, Table S3): Cth_0452 (OlpC) (20), Cth_2761 (annotated as GH9-CBM3c-Doc1) (15), Cth_3079 (Orf2p), and Cth_3132 (annotated as Unknown-Doc1) (15). All of these showed changes of ≥twofold with the notable exception of Cth_0452, which showed a decrease of approximately eightfold (Fig. 5).

Discussion

Deletion of Cel48S from C. thermocellum led to a decrease in the enzymatic hydrolysis rate, a decrease in microbial hydrolysis rate, and a decrease in biomass formation during growth on Avicel. The similarity of enzyme saturation curves for the WT and parent strains suggests that the ΔgyrF mutation in the parent strain has no effect on cellulose function, as expected. The S mutant strain, however, exhibited a reduction in both specific activity and saturation rate. A reduction in specific activity is indicative of impaired function and consistent with decreased synergy among components of the cellulosome in the absence of Cel48S (3).

The role of GH families in cellulose solubilization is a topic of much debate. Family 48 cellulases are a prominent component of many bacterial cellulase systems and, due to their ubiquity, are thought to play an important role in cellulose solubilization (21). On one hand, disruption of the single family 9 GH in C. phytofermentans eliminated its ability to grow on filter paper (11), suggesting that some GH families are essential. On the other hand, down-regulation of cel48F in C. cellulolyticum gave only a modest effect, with no reduction in growth rate and a 30% reduction in specific activity of the purified cellulosome (22). The deletion of Cel48S from C. thermocellum reduced growth rate and specific activity by about half, which was more substantial than the cel48F phenotype observed in C. cellulolyticum but not the complete elimination of cellulolytic ability observed in C. phytofermentans. Cel48Y expression was either not detected or detected at extremely low levels in several recent studies of the C. thermocellum cellulosome (12, 13, 15) and was not detected in our samples (SI Appendix, Table S3). Nevertheless, we
deleted it to show that it was not responsible for the residual cellulolytic activity observed in the S mutant strain and to further support our claim that family 48 GHs are not necessary for growth of \textit{C. thermocellum} on crystalline cellulose. This raises the question: what is the role of family 48 GH enzymes in bacterial cellulase systems? Based on the enzymatic and microbial data, \textit{Cel48S} appears to be a rate-limiting enzyme in cellulose solubilization, but, even in the absence of \textit{Cel48S}, \textit{C. thermocellum} produces a cellulosome with the ability to completely solubilize crystalline cellulose. Understanding the mechanism behind this residual activity is a promising direction for future work.

Mutant characterization was undertaken at a microbial as well as enzymatic level because there is strong technological interest in microbial conversion systems and many additional fundamental phenomena are operative when cellulose solubilization is mediated by microbial cultures as compared with enzyme preparations. Neither the \(\Delta\text{pyrF}\) modification of the parent strain nor the \(\Delta\text{cel48S}:\text{PgapDH-cat}\) modification in the S mutant strain exhibited a deleterious effect on cellulose growth (Fig. 2). Avicel consumption, on the other hand, was much slower for the S mutant strain than for either the WT or parent strains with

<table>
<thead>
<tr>
<th>Strain</th>
<th>Maximum rate, g/L/h</th>
<th>Substrate</th>
<th>Biomass</th>
</tr>
</thead>
<tbody>
<tr>
<td>CELLOBIOSE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>1.1 ± 0.04</td>
<td>0.06 ± 0.008</td>
<td>0.9 ± 0.01</td>
</tr>
<tr>
<td>Parent</td>
<td>1.1 ± 0.02</td>
<td>0.06 ± 0.006</td>
<td>1.0 ± 0.01</td>
</tr>
<tr>
<td>S mutant</td>
<td>1.0 ± 0.09</td>
<td>0.06 ± 0.006</td>
<td>0.4 ± 0.01</td>
</tr>
</tbody>
</table>

Biomass production was inferred from pellet nitrogen measurements. The data represent the averages of the results from duplicate measurements. Error represents one SD.

Table 1. Maximum rate of substrate consumption and biomass production in the WT, parent, and S mutant strains

Fig. 2. Batch fermentations of WT, parent, and S mutant strains growing on either Avicel or cellobiose. (Left) Substrate consumption; (Right) biomass production. Biomass production was inferred based on pellet nitrogen measurements. Growth of all three strains was similar on cellobiose, whereas on Avicel, the S mutant strain consumed the Avicel more slowly and made less biomass. The data at each time point represent the averages of the results from duplicate measurements. Error bars represent SD.

Fig. 3. Enzymatic activity of purified cellulosomes against 0.6 g/L Avicel. Activity is measured in SGE. The data represent the averages of the results from triplicate experiments. Error bars represent SD.
twofold more time required to achieve complete cellulose solubilization under comparable conditions. Slower Avicel consumption by the mutant culture is consistent with the lower activity of the mutant cellulosome observed during in vitro experiments. However, the reduced pellet biomass observed in the S mutant may also be a factor. The relative importance of reduced cellulosome efficiency and reduced cellulosome production in determining the slower utilization of Avicel by the mutant is unclear at this time. The reduction in cellulosomal protein was partially compensated by an increase in supernatant protein (SI Appendix, Figs. S4 and S5). The identity of these proteins is currently unknown but might point to a regulatory effect. The absence of Cel48S did not significantly impact the abundance of any other major component of the cellulosome. Densitometry analysis of the denaturing gel (SI Appendix, Fig. S4) shows very little change for bands other than Cel48S. Tandem mass spectrometry revealed changes in four cellulosomal proteins; two of these were noncatalytic (OlpC and Orf2p) and two were minor catalytic components (Cthe_2761 and Cthe_3132, <1% of cellulosomal protein). Both of the catalytic components that showed significant changes in abundance between the S mutant and parent strains also showed significant changes in abundance between the WT and parent strains, suggesting that these changes do not have a large effect on cellulose function. The two other proteins that showed significant changes in abundance, OlpC and Orf2p, contain cohesin domains and are therefore thought to be structural components of the cellulosome. The twofold increase in Orf2p abundance and eightfold decrease in OlpC abundance might reflect the action of regulatory mechanisms. C. thermocellum has been shown to regulate protein expression by substrate sensing (23) and changes in growth rate (24, 25), both of which may have occurred in the S mutant strain.

The development of gene knockout capability for C. thermocellum has provided an opportunity to improve understanding of it's microbial cellulosome utilization and the action of the cellulosome. The creation and characterization of a cel48S mutant strain and its cellulosome underscore the value of this approach.

Materials and Methods
Strains and Media. C. thermocellum strain DSM 1313 (WT) was grown in modified DSM 122 broth (26) with the addition of 50 mM 3-(N-morpholino)propanesulfonic acid (Mops) sodium salt and 3 g/L trisodium citrate (Na₃C₆H₅O₇·2H₂O). The parent strain (SI Appendix, Table S1) is a deletion of pyrF, which exhibits auxotrophy for uracil and needs to be supplemented with uracil at 40 μg/mL. The S mutant strain was derived from the parent strain by replacing the cel48S gene with a thiamphenicol antibiotic resistance marker. Unless otherwise noted, cells were grown at 55 °C with gentle stirring using 5 g/L cellobiose as the primary carbon source. All manipulations were carried out inside an anaerobic chamber (Coy Laboratory Products Inc.) with an atmosphere of 85% nitrogen, 10% carbon dioxide, 5% hydrogen, and <5 parts per million oxygen.

Molecular Biological Methods. Plasmids were constructed using yeast-mediated ligation (27) or standard cloning techniques (28). Plasmid pMQ87 was a gift from Robert Shanks (27). Plasmids were maintained in Escherichia coli TOP10 cells (Invitrogen Corporation) and prepared using QIAGEN Plasmid Mini kit (QIAGEN Inc.). Sequences of chromosomal DNA were obtained by PCR using genomic DNA from C. thermocellum strain DSM 1313 as the template and primers designed using the C. thermocellum ATCC 27405 genome published by the Joint Genome Institute (http://www.jgi.doe.gov/). Plasmid pGO-01 (SI Appendix, Fig. S1 and Table S1) was based on the C. thermocellum, E. coli, Saccharomyces cerevisiae shuttle plasmid pMU749. Approximately 1-kb regions of homology flanking the cel48S gene on the C. thermocellum chromosome were added upstream and downstream of a thiamphenicol resistance cassette. The native pyrF gene, under control of the C. thermocellum cellulosome phosphorylase (cbp) promoter, was cloned outside the homologous flanks (SI Appendix, Fig. S1). Plasmid pL2 was constructed in a similar manner, with ~1-kb regions of homology flanking the cel48S gene on the C. thermocellum chromosome and neomycin resistance provided by the kan marker from plasmid pKm1 (29). PCR was performed using either Taq or Phusion DNA polymerase (New England Biolabs Inc.) according to the directions provided by the manufacturer. When using whole cells as the PCR template, a 10-min heating step was included at the beginning of the thermocycling protocol to lyse the cells. When using Taq DNA polymerase, the lysing temperature was 95 °C. When using Phusion DNA polymerase, the lysing temperature was 98 °C. DNA sequencing was performed using standard techniques with an ABI Model 3100 genetic analyzer (Applied Biosystems).

**Fig. 4.** SDS/PAGE of proteins from the WT, parent, and S mutant strains. The left three lanes are supernatant from 10 g/L Avicel fermentations. The right three lanes are culture supernatant purified by affinity digestion and diluted eightfold. The Cel48S band is indicated by the triangle. The left lane shows molecular mass markers (kDa).

**Fig. 5.** Protein abundance determined by APEX. Dashed lines represent a twofold change. The limit of detection for Cel48S was $5 \times 10^{-5}$. 

4 of 6 | www.pnas.org/cgi/doi/10.1073/pnas.1003584107

Olson et al.
Transformation. Transformation was performed according to protocol. Cells were prepared for transformation by inoculation from a freezer stock into one of several cell media with 5 μg cell membrane analysis to the primary culture source. The culture was incubated at 55 °C until the optical density at 600 nm (OD600) reached 0.4–0.8. Cells were washed in reverse-osmosis purified (18 MΩ) water that had been autoclaved to remove oxygen. Twenty microliters of cell suspension were added to each cuvette along with 1–8 μL of DNA (10–2,000 ng) diluted in water. Standard 0.1-cm gap electrophoresis cuvettes were used. A series of 60 square pulses, each of 30-μs duration, were applied to the sample. The period of the pulses was 300 μs and the amplitude was 1.9 kV/cm, resulting in an applied field strength of 19 kV/cm. After pulsing, cells were allowed to recover at 51 °C in 3–5 mL uracil-supplemented media overnight (15–18 h) before being subjected to selective pressure. Selection for the cat marker was performed by the addition of trim at a final concentration of 48 μg/mL to the culture medium. Selection for neo was performed by the addition of Neo at a final concentration of 250 μg/mL to the culture medium. Selection against the pyrF gene was performed by the addition of FOA at a final concentration of 500 μg/mL. When used in conjunction, the FOA concentration was 500 μg/mL and the thiamphenicol concentration was 6 μg/mL.

Microbial Growth and Hydrolysis Analysis. Microbial growth and hydrolysis analysis were determined by batch fermentation in uracil-supplemented media with 10 g/L Avicel PH-105 microcrystalline cellulose (Sigma-Aldrich) as the primary carbon source. Fermentations were performed in a 1-L volume at 55 °C in a Sartorius Q+ fermentation system with pH control provided by the addition of 5N potassium hydroxide. Fermentations were determined to be complete when no further base addition occurred. The culture was stirred at 100 rpm, which was sufficient to keep Avicel particles suspended. Thirty-milliliter aliquots were drawn at intervals throughout the fermentation to determine substrate consumption and biomass formation. Cellulosomal protein concentration was measured by SDS-PAGE analysis. Pellet nitrogen was measured with a Shimadzu TEC-CPH elemental analyzer with TNM-1 and ASI-V modules (Shimadzu Corp.) on 1 mL aliquots that had been washed twice with water and then centrifuged. Dry weight was measured by washing an 8-mL aliquot twice with water, followed by centrifugation. The washed sample was then dried at 60 °C to constant weight. For cellobiose-grown cultures, dry weight was expressed exclusively of biomass. For Avicel-grown cultures, the dry weight represents the sum of Avicel and biomass. Biomass was calculated from pellet nitrogen data assuming that nitrogen makes up a constant 10.6% of cell mass (30). Residual Avicel was determined by subtracting biomass from dry weight.

Enzymatic Analysis. After fermentation, celluloisomes were purified from 200 mL of broth, using the affinity purification protocol from Morgenstern et al. (16). Total protein was measured with Bio-Rad Bradford protein assay with BSA (BSA) as a standard. Initial hydrolysis rate measurements were performed at 55 °C in a 50-mL volume in a 150-mL serum bottle with constant shaking and 50 mM sodium acetate (pH 5.0), 5 mM cysteine-HCl, 12 mM CaCl2, 55 °C in a 50-mL volume in a 150-mL serum bottle with constant shaking and following the protocol of Fernandez et al. (17) The activity buffer contained 50 mM sodium acetate (pH 5.0), 5 mM cysteine-HCl, 12 mM CaCl2, 40 μg/mL tetracycline (to prevent microbial growth), and 0.02% vol/vol Novozym 188 (to convert cellobiose to glucose for later analysis and to prevent product inhibition from affecting initial rate measurements [Sigma-Aldrich]). The concentration of Novozym 188 was >10-fold in excess of what would have been necessary to convert all of the cellobiose generated by the cellulosome into glucose. One-microliter aliquots were taken hourly during the first 3 h for SGE analysis. Glucose was measured using the Hexokinase Glucose Assay kit (Sigma-Aldrich). The slope of these measurements was used to determine the initial hydrolysis rate (μmol/min) for a range of different cellulosome loadings from 0.05 to 0.8 mg. One unit (U) of enzyme activity releases 1 μmol SGE/min. Specific activity is measured in U/mg protein.

Denaturing Gel Electrophoresis. Ten-microliter aliquots of the fermentation supernatant and 1.25–μL aliquots of purified cellulosomes were analyzed on 2.5% Tris-HCl SDS/PAGE gel (Bio-Rad) to visualize the individual subunits (Fig. 4). Bands from the three purified lanes were excised, and the major components were analyzed by mass spectrometry (NextGen Science) (SI Appendix, Table S2).

Protein Measurements. Samples were prepared for protein measurement following the method of Zhang and Lynd (31). Supernatant protein was measured with the Bradford assay (Thermo Scientific), and pellet protein was measured with the BCA assay (Thermo Scientific). BSA was used as the standard.

Protein Data Collection for Tandem Mass Spectrometry. Purified cellulosomes were processed for 2D LC-MS/MS analysis as follows. Proteins were precipitated overnight at 20 °C by adding trichloroacetic acid (TCA) to a final concentration of 20%. The resulting protein pellets were washed with ice-cold acetone, resolubilized in denaturing buffer (8 M urea, 100 mM Tris, pH 8.0), and reduced with 20 mM DTT. Samples were diluted to 4 M urea with 100 mM Tris, 10 mM CaCl2, pH 8.0, and digested via two additions of modified trypsin (Promega) at a 1:75 enzyme to protein ratio (w/w). Peptides were protonated with 0.1% formic acid, spun filtered (Ultrafree-MC, Millipore), and 25 μg was loaded onto a Mudpit (32, 33) back column paired with cation exchange (C18, Luna; Phenomenex) and C18 reversed phase (RP, Aqua; Phenomenex) resins, as previously described (34), and separated by charge (salt pulses of 0, 10, 25, and 100 of 500 mM ammonium acetate) and hydrophobicity (100-mn aqueous to organic gradient) using an HPLC pump (u3000; Dionex) coupled to an LTO XL mass spectrometer (Thermo Scientific). Eluting peptides were measured, isolated, and fragmented by the LTO XL operating in data-dependent mode. Each sample was analyzed in triplicate.

The resulting tandem mass spectra were searched with SEQUEST (35) against the C. thermocellum ATCC 27405 proteome concatenated with reversed FASTA protein entries to assess false-discovery rates (FDR), common contaminants, and the cat gene to assess the fidelity of the Cel48S deletion and determine the composition of the resulting cellulosome. As urea was used as the denaturant, searches were performed with the inclusion of carbamylation (+43 Da) as a dynamic modification potentially occurring on lysine, arginine, and peptide N-termini.

Protein Data Analysis. The DTASelect data were converted into pepXML format with the freely available software program Out2XML (Institute for Systems Biology). Protein abundance was analyzed with the APEX Quantitative Proteomics Tool (36) (U. Craig Venter Institute) using a 1% false-detection threshold. After quantification, proteins that were not identified in all nine samples (three replicates per strain, three strains) were eliminated. Only cellulosomal proteins (12, 15) were included in abundance normalization. APEX is a technique for the acquisition of protein expression data from mass spectrometry experiments that uses machine learning to correct for sequence-specific detection bias (37). The 40 most abundant proteins (SI Appendix, Table 53) were selected for training the APEX classifier. Pair-wise comparisons were made between data sets (WT, parent, and S mutant) and proteins whose abundance had changed significantly were identified by t test (p < 0.01) of log-transformed data. Log-transformed data were found to be normal by the Shapiro–Wilks test at the 0.05 level.

ACKNOWLEDGMENTS. We thank Dr. Erin Wiswall for assistance with protein purification and Meryl B. Olson for her assistance with statistical calculations. This research was supported by Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Department of Energy Grant DE-FG02-02ER15353 and grants from the Mascoma Corporation and the BioEnergy Science Center, Oak Ridge National Laboratory, a Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the Department of Energy Office of Science.

shows that cellulose degradation requires the family 9 hydrolase.


Olson et al.
Q: 1. Please provide modifier for Cel48S in title.

Q: 2. Please contact PNAS_Specialist@dartmouthjournals.com if you have questions about the editorial changes, this list of queries, or the figures in your article. Please include your manuscript number in the subject line of all e-mail correspondence; your manuscript number is 201003584. Please (i) review the author affiliation and footnote symbols carefully, (ii) check the order of the author names, and (iii) check the spelling of all author names and affiliations. Please indicate that the author and affiliation lines are correct by adding the comment “OK” next to the author line. Please note that this is your opportunity to correct errors in your article prior to publication. Corrections requested after online publication will be considered and processed as errata.

Q: 3. Please provide cities for all author affiliations.

Q: 4. PNAS allows up to five key terms that (i) do not repeat terms present in the title (which is searchable online) and (ii) do not include nonstandard abbreviations. You may add two terms. Also, please check the order of your key terms and approve or reorder them as necessary.

Q: 5. Please review the information in the author contribution footnote carefully. Please make sure that the information is correct and that the correct author initials are listed. Note that the order of author initials matches the order of the author line per journal style. You may add contributions to the list in the footnote; however, funding should not be an author’s only contribution to the work.

Q: 6. You will receive a notification from the PNAS eBill system in 1-2 days. Each corresponding author is required to log in to the system and provide payment information for applicable publication charges (purchase order number or credit card information) upon receipt of the notification. You will have the opportunity to order reprints through the eBill system if desired, as well. Failure to log in and provide the required information may result in publication delays.

Q: 7. Reminder: You have chosen the Open Access option for your paper and have agreed to pay an additional $1275 (or $950 if your institution has a site license).

Q: 8. Please indicate whether the sequences have been deposited in GenBank or another publicly accessible database before your page proofs are returned. It is PNAS policy that the data be deposited BEFORE the paper can appear in print. Please ensure that correct database is listed (delete GenBank is necessary) and add accession numbers.

Q: 9. PNAS no longer allows citations of unsupported data; the citation for unpublished data (Tripathi et al. attached manuscript) has thus been removed from throughout the text. Please (a) provide the data as Supporting Information, (b) provide an “in press” reference if the article has been accepted for publication, or (c) omit the statement from the text.
Q: 10_PNAS italicizes the names of genes and alleles. Please check cel48S and cel48Y throughout the manuscript and correct as necessary. If, by “XXX gene,” you mean “the gene that encodes protein XXX,” then italic type is not necessary.

Q: 11_Please verify that all supporting information (SI) citations are correct. Note, however, that the hyperlinks for SI citations will not work until the article is published online. In addition, SI that is not composed in the main SI PDF (appendices, datasets, movies, and “Other Supporting Information Files”) have not been changed from your originally submitted file and so are not included in this set of proofs. The proofs for any composed portion of your SI are included in this proof as subsequent pages following the last page of the main text. If you did not receive the proofs for your SI, please contact PNAS_Specialist@dartmouthjournals.com.

Q: 12_Can you cite the reference in which the methods appear?

Q: 13_PNAS italicizes the names of genes and alleles. Please check cat, Δcel48S::PgapDH-cat, and neo throughout the manuscript and correct as necessary. If, by “XXX gene,” you mean “the gene that encodes protein XXX,” then italic type is not necessary.

Q: 14_Did the clones shows these amplicons or did the PCR? Please clarify.

Q: 15_PNAS italicizes the names of genes and alleles. Please check ΔpyrF throughout the manuscript and correct as necessary. If, by “XXX gene,” you mean “the gene that encodes protein XXX,” then italic type is not necessary.

Q: 16_Is each component less than 1% or are the two together less than 1%? Please clarify.

Q: 17_PNAS does not allow statements of novelty or priority. Please rewrite/approve edit.

Q: 18_Please provide the institutional affiliation, city and state or country for Robert Shanks.

Q: 19_PNAS italicizes the names of genes and alleles. Please check cbp and kan throughout the manuscript and correct as necessary. If, by “XXX gene,” you mean “the gene that encodes protein XXX,” then italic type is not necessary.

Q: 20_Please provide a unit of measure for 0.4–0.8.

Q: 21.Does 60 need a unit of measure (Hz, ms, etc)?