

# SCIENTIFIC REPORTS



OPEN

## Adaptive engineering of a hyperthermophilic archaeon on CO and discovering the underlying mechanism by multi-omics analysis

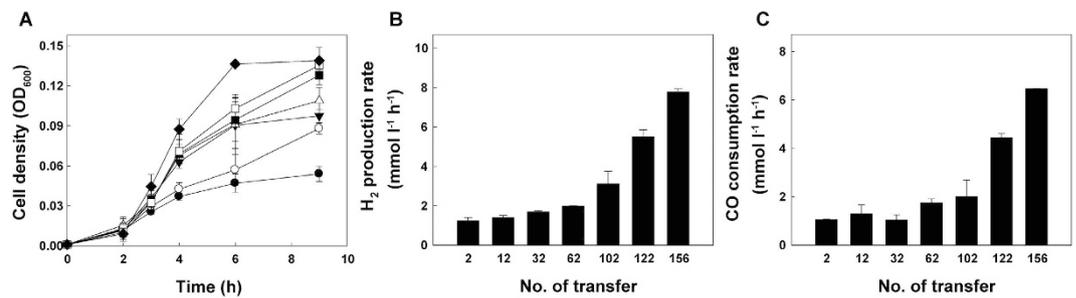
Seong Hyuk Lee<sup>1,2,\*</sup>, Min-Sik Kim<sup>3,\*</sup>, Jae-Hak Lee<sup>1,\*</sup>, Tae Wan Kim<sup>1,2</sup>, Seung Seob Bae<sup>1</sup>, Sung-Mok Lee<sup>1</sup>, Hae Chang Jung<sup>1,2</sup>, Tae-Jun Yang<sup>1</sup>, Ae Ran Choi<sup>1</sup>, Yong-Jun Cho<sup>4</sup>, Jung-Hyun Lee<sup>1,2</sup>, Kae Kyoung Kwon<sup>1,2</sup>, Hyun Sook Lee<sup>1,2</sup> & Sung Gyun Kang<sup>1,2</sup>

The hyperthermophilic archaeon *Thermococcus onnurineus* NA1 can grow and produce H<sub>2</sub> on carbon monoxide (CO) and its H<sub>2</sub> production rates have been improved through metabolic engineering. In this study, we applied adaptive evolution to enhance H<sub>2</sub> productivity. After over 150 serial transfers onto CO medium, cell density, CO consumption rate and H<sub>2</sub> production rate increased. The underlying mechanism for those physiological changes could be explained by using multi-omics approaches including genomic, transcriptomic and epigenomic analyses. A putative transcriptional regulator was newly identified to regulate the expression levels of genes related to CO oxidation. Transcriptome analysis revealed significant changes in the transcript levels of genes belonging to the categories of transcription, translation and energy metabolism. Our study presents the first genome-scale methylation pattern of hyperthermophilic archaea. Adaptive evolution led to highly enhanced H<sub>2</sub> productivity at high CO flow rates using synthesis gas produced from coal gasification.

Microorganisms have been metabolically engineered, employing integrated strategies of systems biology, synthetic biology and evolutionary engineering, to enhance their output of valuable products<sup>1</sup>. Generally, metabolic engineering via knowledge-based rational design has been used to enhance the output of target products; however, this approach requires *a priori* genetic or biochemical information, and the complexity of cellular physiological responses must be considered<sup>2</sup>. In this regard, an evolutionary engineering approach may serve as an alternative for obtaining suitable phenotypes in cases where there is a lack of prior knowledge<sup>3–6</sup>. This approach has been applied to many organisms<sup>7</sup>, including *Escherichia coli*, *Myxococcus xanthus*, *Pseudomonas fluorescens*, *Saccharomyces cerevisiae* and *Thermotoga maritima*<sup>8</sup>. However, it has never been reported for hyperthermophilic archaea.

Biohydrogen has the potential to considerably reduce costs and environmental impact because it can be produced using sunlight, minimal nutrients or organic waste effluents<sup>9</sup>. Microbial H<sub>2</sub> production by dark fermentation can produce H<sub>2</sub> at a faster rate than photosynthesis and can upcycle renewable, cheap and abundant resources<sup>10</sup>. *T. onnurineus* NA1 is a carboxydrotrophic hydrogenogen, capable of utilizing CO and producing H<sub>2</sub> in the following reaction:  $\text{CO} + \text{H}_2\text{O} \rightarrow \text{H}_2 + \text{CO}_2$  ( $\Delta G^\circ = -20 \text{ kJ/mol}$ )<sup>11</sup>. CO-driven H<sub>2</sub> production is mediated by an enzymatic system encoded in the CODH gene cluster, composed of a carbon monoxide dehydrogenase (CODH), a hydrogenase and a Na<sup>+</sup>/H<sup>+</sup> antiporter<sup>12–14</sup>. We have enhanced the H<sub>2</sub> production rate of the strain through knowledge-based rational design, such as by overexpressing genes in the CODH gene cluster by promoter engineering or by CO regulator engineering<sup>14–17</sup>. Although the design was successful, the resulting mutant strains could not grow in a fermentor supplied with a CO flow rate above 800 ml min<sup>-1</sup>. Therefore, the development of a mutant with improved performance at high CO flow rates is desirable.

<sup>1</sup>Korea Institute of Ocean Science and Technology, Ansan, Republic of Korea. <sup>2</sup>Department of Marine Biotechnology, Korea University of Science and Technology, Daejeon, Republic of Korea. <sup>3</sup>Korea Institute of Energy Research, Daejeon, Republic of Korea. <sup>4</sup>Chunlab, Inc., Seoul, Republic of Korea. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to H.S.L. (email: leeh522@kiost.ac.kr) or S.G.K. (email: sgkang@kiost.ac.kr)



**Figure 1. Physiological changes in *T. onnurineus* NA1 following serial transfers into fresh MM1 medium containing 100% CO.** After 2 (closed circle), 12 (open circle), 32 (closed inverted triangle), 62 (open triangle), 102 (closed square), 122 (open square) and 156 (closed diamond) transfers, (a) cell density (expressed as optical density at 600 nm) was determined at the indicated timepoints. (b)  $H_2$  production rates and (c) CO consumption rates were determined during exponential phase. All experiments were conducted independently in duplicate.

Locus_tag	Genome position	Timeline of the mutation	Mutation type <sup>a</sup>	Codon change	Description
TON_0536–0541	490749–496535	after 12 <sup>th</sup> transfer	5786-bp deletion	frame shift	cytosolic NiFe-hydrogenase, formate transporter, formate dehydrogenase, iron-sulfur binding proteins
TON_0820	760634	after 12 <sup>th</sup> transfer	substitution	Leu to Arg	aromatic amino acid permease
TON_1525	1400830	after 12 <sup>th</sup> transfer	substitution	Thr to Ile	putative transcriptional regulator
TON_1544	1420154	after 62 <sup>nd</sup> transfer	substitution	Pro to Leu	membrane protein
TON_0982	906462	after 32 <sup>nd</sup> transfer	T insertion	frame shift	aminotransferase
TON_1548	1422235	after 102 <sup>nd</sup> transfer	T deletion	frame shift	hypothetical protein
TON_1664	1527982	after 156 <sup>th</sup> transfer	A deletion	frame shift	cation transporter
TON_1694	1559247	after 102 <sup>nd</sup> transfer	12-bp insertion	frame shift	membrane-associated metalloprotease
TON_0544	499427–499765	after 156 <sup>th</sup> transfer	338-bp deletion	frame shift	alcohol dehydrogenase
TON_1475–1476	1351559–1351609	after 62 <sup>nd</sup> transfer	51-bp deletion	–	intergenic region

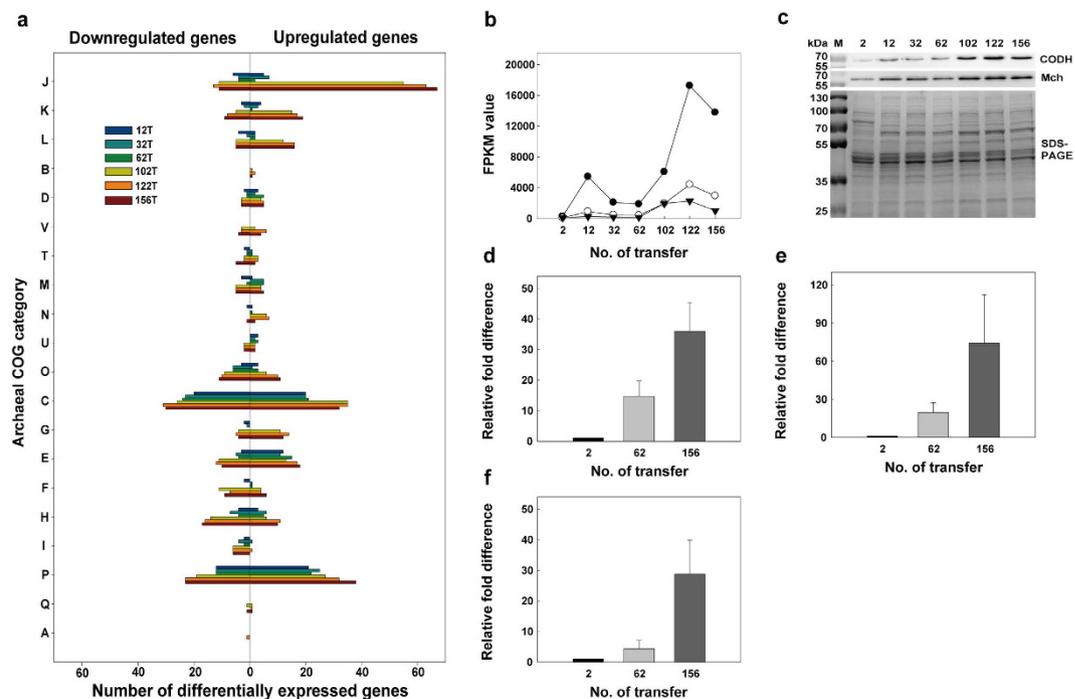
**Table 1. Mutations found in the 156 T strain.** <sup>a</sup>All mutations were confirmed by PCR verification and Sanger sequencing.

In the current study, we utilized evolutionary engineering to optimize *T. onnurineus* NA1 under CO conditions and investigated the underlying mechanism of optimization by employing multi-omics analysis. Genetic variations and transcriptomic and epigenomic changes were analyzed using next generation sequencing. The methylation profile of NA1 was monitored via SMRT DNA sequencing<sup>18</sup>. DNA methylation has been reported to be involved in the regulation of genome replication<sup>19</sup>. Genome-scale methylation patterns of some bacteria and archaea have been published<sup>20–22</sup>. However, genomic scale interpretation of base modifications has not yet been reported for hyperthermophiles. Kinetic analysis of  $H_2$  production in the adapted strain was conducted and compared with other genetically engineered strains.

## Results

**Physiological changes in *T. onnurineus* NA1 during adaptive evolution on CO.** We applied an evolutionary approach to enable *T. onnurineus* NA1 to adapt to 100% pure CO. *T. onnurineus* NA1 was grown to stationary phase to trigger spontaneous mutation<sup>23–26</sup> and then transferred to fresh medium containing CO as an energy source; the strain was transferred over 150 times. Through these serial transfers, physiological changes were monitored, and gradual increases in cell density,  $H_2$  production rate and CO consumption rate were observed (Fig. 1). After 156 transfers, the evolved strain, designated 156 T, showed 2.8-, 6.5- and 5.9-fold higher values in maximum cell density,  $H_2$  production rate and CO consumption rate, respectively, than the parental strain (2 T) (Fig. 1). The correlation between the increases in CO consumption rate and  $H_2$  production rate seemed to reflect that  $H_2$  production occurred by hydrogenogenic carboxydutrophy. The enhanced cell growth was associated with the increases in CO consumption rate and  $H_2$  production, concomitant with energy conservation via the conversion of CO oxidation.

**Genomic analysis.** The genomic DNA of 156 T was isolated and sequenced using Illumina Hi-seq2000 and SMRT DNA sequencing. In comparison with the sequence of the parental strain, several mutations, including base substitutions, deletions and insertions, were identified and verified by Sanger sequencing (Table 1, Supplementary Fig. 1). The mutation list includes base substitutions, leading to amino acid changes, in aromatic amino acid permease (TON\_0820), transcriptional regulator (TON\_1525) and membrane protein (TON\_1544) and single- or multiple-base deletions/insertions in aminotransferase (TON\_0982), hypothetical protein (TON\_1548), cation transporter (TON\_1664), metalloprotease (12 bp inserted) (TON\_1694), and alcohol



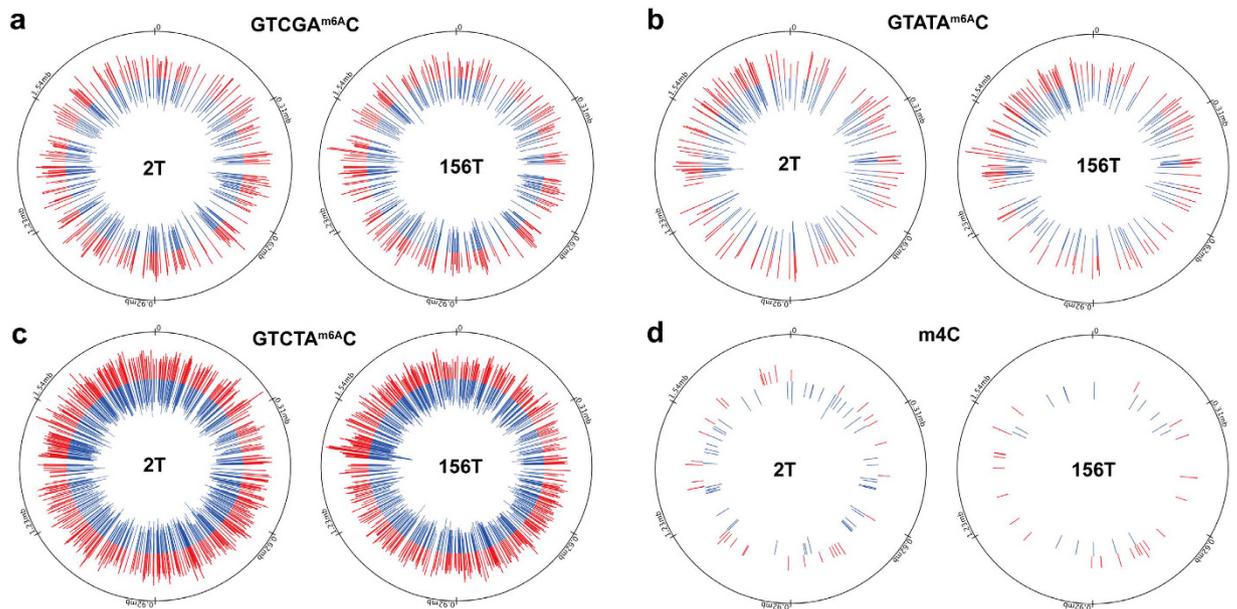
**Figure 2. Archaeal COG classification of differentially expressed genes and changes in the expression levels of genes in the CODH gene cluster during adaptation.** (a) The following COG categories were included: J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; B, Chromatin structure and dynamics; D, Cell cycle control, cell division and chromosome partitioning; V, Defense mechanisms; T, Signal transduction mechanisms; M, Cell wall/membrane/envelope biogenesis; N, Cell motility; U, Intracellular trafficking, secretion and vesicular transport; O, Posttranslational modification, protein turnover and chaperones; C, Energy production and conversion; G, Carbohydrate transport and metabolism; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; P, Inorganic ion transport and metabolism; Q, Secondary metabolite biosynthesis, transport and catabolism; A, RNA processing and modification. The numbers of upregulated genes and downregulated genes are indicated as bars in the right and left panels of the histogram, respectively. (b) RNA-seq analysis of genes encoding CODH (TON\_1018) (closed circle), Mch (TON\_1023) (open circle) and Mnh (TON\_1031) (closed inverted triangle). (c) Western blot analysis of TON\_1018 (67.7 kDa) and TON\_1023 (61.7 kDa). RT-qPCR analysis of mRNA abundance of TON\_1018 (d), TON\_1023 (e) and TON\_1031 (f). Error bars indicate standard deviations from duplicate experiments. FPKM, Fragments per kilobase per million mapped reads; M, molecular mass marker.

dehydrogenase (338 bp deleted) (TON\_0544), as well as in the intergenic region (51 bp deleted) between sodium/phosphate symporter (TON\_1475) and pyruvate/ketoglutarate ferredoxin oxidoreductase (TON\_1476). A long deletion (5786 bp deleted) between cytosolic NiFe-hydrogenase subunit gamma (TON\_0536) and 4Fe-4S binding protein (TON\_0541) also occurred (Supplementary Fig. 2).

**Genome-wide transcriptional analysis.** Transcriptional changes were monitored by RNA-seq using RNA samples from the strains produced after 2, 12, 32, 62, 102, 122, and 156 transfers. The differentially expressed genes in each transferred strain were selected using the DESeq and edgeR packages after being compared against the transcript levels in the 2 T strain. As the transfer process continued, the number of differentially expressed genes (DEGs) gradually increased. Approximately 200 and 500 DEGs were found in the early stages of transfer (12 T, 32 T, 62 T) and late stages of transfer (102 T, 122 T, 156 T), respectively. To interpret the biological and functional meanings of these transcriptional changes, a total of 819 DEGs were categorized according to archaeal clusters of orthologous genes (arCOGs)<sup>27</sup> (Fig. 2a).

As the transfer process progressed, we observed remarkable changes in the number of upregulated DEGs in the COG J (translation, ribosomal structure and biogenesis), COG K (transcription) and COG L (replication, recombination and repair) groupings, together belonging to the category ‘information storage and processing’ (Fig. 2a and Supplementary Data 1). The gradual increase in DEG expression concomitant with the adaptation process seems reasonable because the overexpression of genes encoding ribosomal proteins, RNA polymerase subunits, DNA polymerase II subunits, translation/transcription initiation or elongation factors and DNA topoisomerase subunits supports the observed faster growth and increased growth yield.

Gene clusters encoding seven hydrogenases, Sulfl, SulflI, Mbh, Mfh2, Mch and Frh, and one oxidoreductase, Mbx, homologous to hydrogenase<sup>28</sup>, were included in the DEG list; they fell under the COG C (energy production and conservation) grouping. The expression levels of the ATP synthase genes (TON\_1749–1753) gradually increased during the adaptation. In particular, genes encoding CODH-Mch subunits (TON\_1017–1024) that

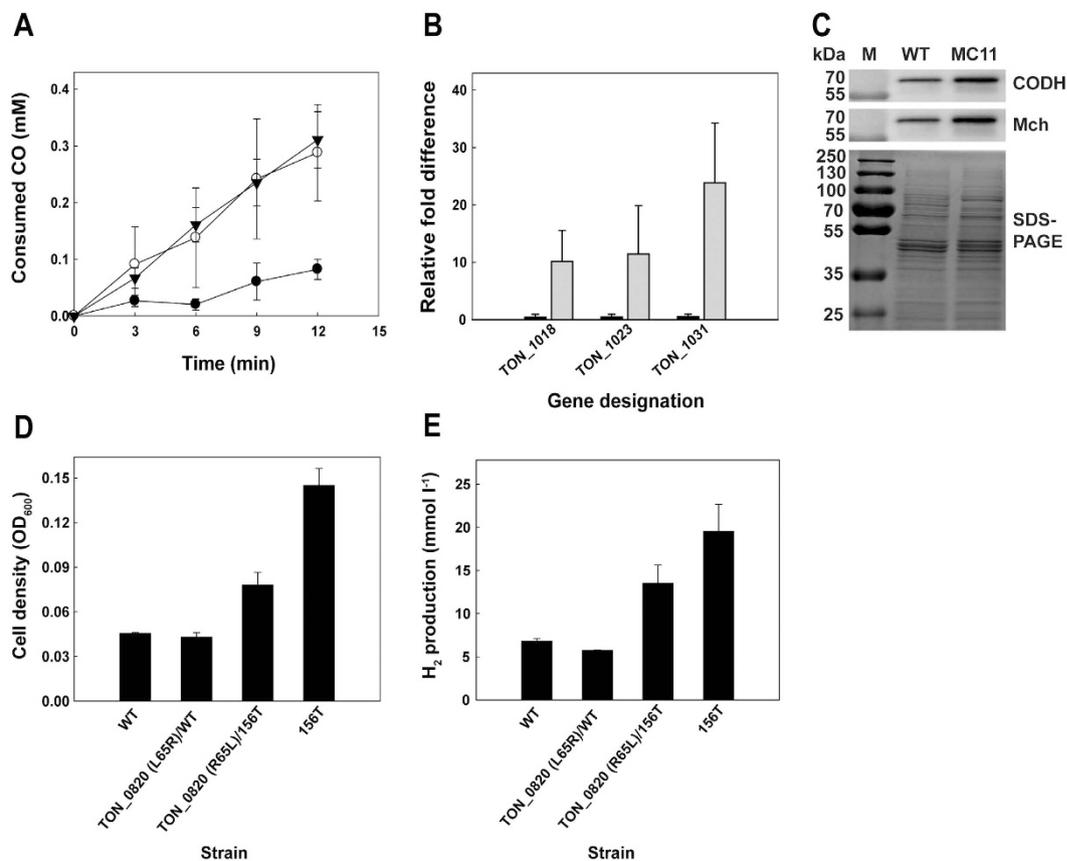


**Figure 3. Comparison of the methylation profiles of adenine and cytosine residues between the 2T and 156T strains.** Only positions with  $p$ -values lower than  $10^{-4}$  (score  $\leq 40$ ) and larger than 25 (coverage  $> 25$ ) were selected as modified bases. The locations of methylated adenines within GTCGA<sup>m6A</sup> (a), GTATA<sup>m6A</sup> (b) and GTCTA<sup>m6A</sup> (c) and of methylated cytosine, m4C (d), are displayed in the 2 T (left panel) and 156 T genomes (right panel) as bars. The length of each bar represents the level of confidence in the methylation ( $-10^4 \log[P\text{-value}]$ ). Methylations in the + and - strands are indicated by red and blue bars, respectively. In the outermost circles, the locations dividing the genome into 6 equal parts are indicated as ticks.

are essential for hydrogenogenic carboxydrotrophy were significantly upregulated at both the transcriptional and translational level as the transfer process progressed (Fig. 2, Supplementary Fig. 3a). The CODH-Mch complex has been proposed to oxidize CO, thereby conserving energy by generating an electrochemical gradient<sup>14</sup>. The electrochemical gradient is eventually coupled to ATP generation by ATP synthase; therefore, the overexpression of CODH-Mch and ATP synthase could allow cells to generate more ATP. The expression of other gene clusters did not persistently increase during the adaption. The Mbh genes (TON\_1583–1595) were strongly upregulated during the early stages of adaptation, but their transcript levels during the late stages of adaption were similar to those in the 2 T strain. The Mbx genes (TON\_0486–0498) were only upregulated at late stages. The Mbh and Mbx might be necessary to dispose of or recycle reduced ferredoxin generated from amino acid degradation. Conversely, the genes encoding Sulfl (TON\_0534–0537), SulflII (TON\_0052), Mfh2 (TON\_1564–1580) and Frh (TON\_1560–1561) showed significant downregulation throughout the entire adaptation period (Supplementary Fig. 3b). Mfh2 has been reported to participate in energy conservation using electrons generated from formate oxidation, and two soluble hydrogenases, Sulfl and SulflII, have been proposed to uptake H<sub>2</sub> to enhance reducing power<sup>29,30</sup>. During the adaptation period, the cells seemed likely to upregulate the genes necessary for growth on CO and to downregulate the genes not needed for this growth.

The upregulated genes in the COG P (inorganic ion transport and metabolism) grouping included two gene clusters encoding Na<sup>+</sup>/H<sup>+</sup> antiporters (TON\_1021–1031, TON\_1582–1589) in proximity to the Mch and Mbh hydrogenases (Supplementary Fig. 3a). These antiporters have been proposed to form a complex with hydrogenase subunits that exchanges proton gradients into sodium gradients; these gradients are eventually coupled with ATP generation by a sodium-specific ATP synthase<sup>31</sup>. The transcript patterns of these antiporter genes were correlated with those of their cognate hydrogenase gene clusters. Genes encoding iron transporters (TON\_0255–0256, 0984, 1873) and a peptide transporter (TON\_1765–1768) exhibited increased transcript levels, likely to enhance cellular uptake of iron and peptides to facilitate cell growth. The gene encoding the cell envelope S-layer protein (TON\_0413) was also upregulated together with genes (TON\_1837, 1842–1843, 1848, 1851) involved in the synthesis of dTDP<sub>T</sub>-rhamnose, a substrate used for the glycosylation of S-layer protein<sup>32</sup> (Supplementary Fig. 4).

**Genome-wide epigenomic analysis.** Adenine and cytosine methylation patterns in the genomes of the 2T and 156T strains were analyzed by SMRT DNA sequencing<sup>18</sup> (Fig. 3). In both strands of the 2T and 156T genomes, 2190 and 2105 sites of N<sup>6</sup>-methyl adenine (m6A) were respectively detected. Adenine methylation remained virtually the same, and more than 95% of the methylated sites (2017 out of 2190) in the 2T strain were also detected in the 156T strain, while the number of hemi-methylated sites increased as the adaptation process progressed. After analyzing the sequences surrounding the modified sites by MEME (Supplementary Table 1), more than 95% of the adenine methylation sequences were identified to contain a (GT(C/A)(G/T)AC) motif, which corresponds to the *AccI* restriction enzyme site. The *AccI* restriction enzyme (TON\_1382) and its cognate methyltransferase (TON\_1383) are encoded in the *T. onnurineus* NA1 genome.



**Figure 4.** The effects of the T55I mutation in TON\_1525 and the L65R mutation in TON\_0820. (a) CO consumption was analyzed by myoglobin assay for wild-type (WT) (closed circle), MC11 (open circle) and 156T (closed inverted triangle) strains. (b) RT-qPCR analysis of TON\_1018, TON\_1023 and TON\_1031 in WT (black bar) and MC11 (gray bar) strains. (c) Western blot analysis of TON\_1018 (67.7 kDa) and TON\_1023 (61.7 kDa) in WT and MC11 strains. (d) Comparison of cell density (expressed as optical density at 600 nm) and (e) H<sub>2</sub> production of the TON\_0820 mutants in comparison with the WT and 156T strains. Error bars indicate the standard deviations of independent triplicate (a,b) or duplicate (d,e) experiments.

Conversely, the cytosine residue methylation pattern significantly changed. In the 2T and 156T strains, 83 and 40 cytosine residues methylated at the N4 site of the cytosine base (N<sup>4</sup>-methylcytosine, m4C) were respectively identified (Supplementary Data 2). Among the 40 methylated cytosines in the 156T strain, 36 of the methylated cytosines were not modified in the 2T strain. The target sequence for cytosine methylation was not as clear as that for adenine methylation, but G<sup>m4C</sup>C was likely to be preferentially methylated over other possible sequences, such as T<sup>m4C</sup>C, A<sup>m4C</sup>C and C<sup>m4C</sup>C. More than 50% of the cytosine methylation sites possessed G<sup>m4C</sup>C sequences in both the 2T (53 sites) and 156T (24 sites) strains.

In the promoter region of the Mfh2 hydrogenase gene cluster, two distinct methylation sites were identified only in the 156T strain. The methylated cytosine residues were adjacent to the DNA-binding motif (GTTNNNAAC) for a putative sulfur response regulator, SurR (Supplementary Fig. 6a), that has been reported in *Pyrococcus furiosus*<sup>33</sup>. One methylated adenine residue was detected inside an inverted repeat. Although it was not experimentally proven, the methylation of this residue might affect the binding of SurR and an unknown transcriptional regulator to the Mfh2 promoter, resulting in the downregulation of the gene cluster. After a few transfers, the transcript levels of the genes in the Mfh2 gene cluster were drastically downregulated (Supplementary Fig. 6b–d).

**Effects of mutations on adaptive physiological changes.** To determine the mutations critical to physiological change, we tracked the distributions of seven mutations within the coding regions of selected transferred strains. The timepoint when each mutation initiated was scattered from 12th transfer to 156th transfer, and the distribution of each mutation increased as the number of transfers increased, except for the mutation in the aminotransferase gene (TON\_0982) (Supplementary Fig. 7). Considering the pattern of changes in H<sub>2</sub> production rate and CO consumption rate, that is, small increases during early stages and then big increases after 102nd transfer, the distributions of mutations of the genes encoding putative transcriptional regulator (TON\_1525) and aromatic amino acid permease (TON\_0820) among total population seemed well matched with those physiological changes. Therefore, the TON\_1525 and TON\_0820 mutations found in strain 156T were introduced into the parental strain. The resulting mutant, TON\_1525 (T55I)/WT (designated as MC11), showed a higher CO consumption rate than the wild-type strain; this rate was comparable to that of the 156T strain (Fig. 4a).

Organism	Strategy	H <sub>2</sub> production rate (mmol l <sup>-1</sup> h <sup>-1</sup> )	Specific H <sub>2</sub> production rate (mmol g <sup>-1</sup> h <sup>-1</sup> )	Reference
156T	Evolutionary engineering	220.8 <sup>c</sup>	334.6 <sup>c</sup>	This study
ΔCorR/corQR <sup>1</sup>	Transcriptional regulator engineering	191.9 <sup>a</sup>	249.6 <sup>a</sup>	15
KS0510	Promoter engineering	155.1 <sup>b</sup>	245.1 <sup>b</sup>	16
MC01	Promoter engineering	123.5 <sup>b</sup>	194.7 <sup>b</sup>	14
wild-type	–	32.9 <sup>a</sup>	151.3 <sup>a</sup>	14

**Table 2. Comparison of H<sub>2</sub> production rates between wild-type and mutant strains of *T. onnurineus* NA1.** <sup>a–c</sup>100% CO was continuously fed at flow rates of 240 (MC01 and KS0510 strains), 400 (wild-type and ΔCorR/corQR<sup>1</sup> strains) or 800 ml min<sup>-1</sup> (156T strain).

The increased CO consumption in the MC11 strain was likely related to the increased transcript levels of the genes in the CODH gene cluster. The transcript levels of the TON\_1018, TON\_1023, and TON\_1031 genes were respectively 20.6-, 25.4-, and 47.7-fold higher than the levels in the parental strain (Fig. 4b). Consistent with the transcriptional changes, CODH (TON\_1018) and Mch (TON\_1023) protein expression significantly increased in the MC11 mutant (Fig. 4c). These results indicate that the mutation in the TON\_1525 gene played a crucial role in causing physiological change by enhancing the transcript levels of the CODH gene cluster.

To investigate whether the TON\_1525 protein could act like a DNA binding protein on the CODH promoter and whether mutation of TON\_1525 affected its binding affinity, wild-type and T55I recombinant TON\_1525 proteins were expressed and purified from *E. coli*. The unmodified recombinant TON\_1525 protein specifically bound to the CODH promoter, while the binding affinity of the T55I mutant protein was less than that of the wild-type protein, as shown by competition electrophoretic mobility shift assay (EMSA) (Supplementary Fig. 5). At present, TON\_1525 can be predicted to function as a transcriptional repressor on the CODH promoter, as mutation of TON\_1525 decreased its binding affinity toward the regulator of the promoter and elevated the transcript level of the downstream gene cluster.

The effect of the mutation of TON\_0820 was not clearly shown when we assessed a TON\_0820 (L65R)/WT mutant. However, when we reverted the mutation (L65R) in the 156 T strain back to the wild-type sequence (R65L), creating TON\_0820 (R65L)/156 T, significant decreases in cell density and H<sub>2</sub> production were observed (Fig. 4d,e). This result suggests that the mutation of TON\_0820 is another critical factor underlying the observed physiological changes that occurred during the adaptive evolution process, probably in combination with other factors.

**Kinetic measurements of CO consumption and H<sub>2</sub> production.** Based on the phenotype developed by the 156 T strain following serial transfers under CO conditions, the H<sub>2</sub> production potential of the strain was tested in a bioreactor where 100% CO was continuously fed at a high flow rate. The 156 T strain showed a 7- and 5.2-fold higher maximum biomass yield and H<sub>2</sub> production rate, respectively, when grown under CO at a flow rate of 400 ml min<sup>-1</sup> than those of the wild-type strain (Supplementary Fig. 8a,b). The growth yield based on the optical cell density measured at 600 nm was 5.5, which is very exceptional for a hyperthermophilic archaeon. The 156 T strain also showed growth and produced H<sub>2</sub> at flow rates of 800 ml min<sup>-1</sup> and above, whereas the rationally designed mutants could not grow, indicating that adaptation rendered the 156 T strain less sensitive to high CO levels (Supplementary Fig. 8c,d). It is noteworthy that the 156 T strain displayed the highest H<sub>2</sub> production rate and specific H<sub>2</sub> production rate among the previously reported mutants (Table 2).

H<sub>2</sub> production by the 156 T strain was investigated using synthetic gas (syngas) derived from coal gasification. The H<sub>2</sub> production rate and specific H<sub>2</sub> production rate linearly correlated with the flow rate of syngas (Supplementary Table 2). Considering the composition of the syngas, which is comprised of a low CO level and contaminated with CH<sub>4</sub>, N<sub>2</sub>, CO<sub>2</sub> and H<sub>2</sub>, the H<sub>2</sub> production rates, calculated by subtracting the H<sub>2</sub> content of the syngas, were considerable in comparison with the values obtained under conditions with 100% CO.

Taken together, these results illustrate that adaptive evolution is a powerful method of developing a carboxydotrophic hydrogenogen with enhanced H<sub>2</sub> production from CO.

## Discussion

In the current study, we performed evolutionary engineering of a hyperthermophilic archaeon, *T. onnurineus* NA1, known as a carboxydotrophic hydrogenogen<sup>14</sup> or formate-oxidizing archaeon<sup>34</sup>, to enhance H<sub>2</sub> production from CO. Over 150 serial transfers of cell cultures led *T. onnurineus* NA1 to adapt to possess better utilization of CO and higher production of H<sub>2</sub>, associated with increased cell density.

Genomic analysis revealed a total of ten mutations over 2000 generations. The mutation rates of *T. onnurineus* NA1 were estimated to be approximately  $0.00097 \pm 0.00052$  per replication<sup>35</sup> and 0.0018 per genome per replication<sup>36</sup>. Considering that the mean mutation rates of *E. coli* and *S. cerevisiae* are 1.06 and 1.39 per 100 generations, respectively<sup>7</sup>, the mutation rate of *T. onnurineus* NA1 was quite low, reflecting that thermophilic microbes possess mechanisms to ensure genetic fidelity, a crucial component of the ability to sustain life under extreme environments<sup>35,36</sup>. It is noteworthy that a long deletion, encompassing approximately 5.8 kb of coding region, occurred very early after the 12th transfer.

Our study unveiled two mutations in the TON\_1525 and TON\_0820 genes that played pivotal roles in the adaptive evolution of *T. onnurineus* NA1 on CO. TON\_1525 belongs to a family of Tfx DNA-binding proteins that has not been well studied. The Tfx protein that has been identified in *Methanobacterium thermoautotrophicum*

consists of two domains, a basic DNA-binding domain with a helix-turn-helix motif and an acidic domain possibly used for transcriptional activation<sup>37</sup>. Multiple sequence alignment revealed that TON\_1525 showed only 22.9% sequence identity to Tfx from *M. thermoautotrophicum*. A BLAST search revealed that TON\_1525 showed sequence identities of over 30% toward a set of genes annotated as a transcriptional regulator or, a DNA-binding protein of *Thermococcales* species and methanogens. Based on EMSA data, TON\_1525 seems likely to play a role as a transcriptional repressor on the CODH promoter. Previously, a CO-responsive regulator, CorQR, was reported to activate the expression of the CODH gene cluster<sup>15</sup>. We did not detect any mutations in CorQR during the adaptation process. Another crucial mutation at TON\_0820 seemed contributed to enhancing final cell density (Fig. 4d,e). L65R alteration of wild-type strain did not show much effect, however, R65L mutation in 156 T significantly decreased final cell density. The TON\_0820 mutation might contribute to enhancing cell density probably by supplying more aromatic amino acids. The genome of *T. onnurineus* NA1 does not encode biosynthetic pathways for aromatic amino acids such as Thr, Phe and Trp<sup>12</sup>. These results indicated that two mutations contribute to enhance CO consumption and fitness, however, we need consider to effect of other mutations and further experiments are required to understand overall mechanism under genomic variations on CO-containing medium. Although the distribution pattern is not so well match with physiological changes, TON\_0982 or TON\_1694 encoding amino transferase or a membrane associated-protease, respectively can affect cell growth by modulating the rate of amino acid/protein hydrolysis. The functionality of TON\_1544 and TON\_1548 could not be predicted or understood at this moment and the role of TON\_1664 encoding cation transporter in archaea is unclear yet and awaits further analysis.

Epigenomic analysis revealed cytosine methylation at N4 sites in *T. onnurineus* NA1. N<sup>4</sup>-methylcytosine (m4C) has been discovered in both thermophilic and mesophilic bacteria<sup>38,39</sup>. Considering the high incidence of deamination at high temperatures<sup>40</sup>, the m4C modification may help avoid high mutation rates in organisms living at high temperatures. As deamination of m5C residues increases at high temperatures, thermophiles are unlikely to possess m5C residues in their genomes<sup>38</sup>. Aside from the existence of several other genes annotated as restriction enzymes and putative methyltransferases in the genome, no other consensus sequences, except the sequence encoding the *AccI* restriction site, were discovered in proximity to methylated adenine sites. Nonetheless, it is noteworthy that the methylation profile produced in this study relied on the detection ability of SMRT DNA sequencing technology, and we cannot completely exclude the possibility that other types of methylation could exist in *T. onnurineus* NA1.

In conclusion, we performed evolutionary engineering of a hyperthermophilic archaeon using CO and demonstrated that this approach was very effective in enhancing H<sub>2</sub> productivity. The underlying mechanism was unique in that a novel transcriptional regulator was identified as a crucial adaptation factor, in addition to the potential involvement of distinctive epigenomic changes. The evolved strain was superior to previously reported strains engineered by rational design. In addition to by-product gases from the steel-making process<sup>41</sup>, syngas from a coal gasifier was shown to be an alternative source of CO for H<sub>2</sub> production by *T. onnurineus* NA1.

## Methods

**Strains, media and culture conditions.** The *T. onnurineus* NA1 (KCTC10859) strain was routinely cultured in yeast extract-peptone-sulfur (YPS) medium as previously reported<sup>42</sup>. The modified medium 1 (MM1)<sup>34,43</sup> was prepared with 1 g liter<sup>-1</sup> yeast extract, 35 g liter<sup>-1</sup> NaCl, 0.7 g liter<sup>-1</sup> KCl, 3.9 g liter<sup>-1</sup> MgSO<sub>4</sub>, 0.4 g liter<sup>-1</sup> CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.3 g liter<sup>-1</sup> NH<sub>4</sub>Cl, 0.15 g liter<sup>-1</sup> Na<sub>2</sub>HPO<sub>4</sub>, 0.03 g liter<sup>-1</sup> Na<sub>2</sub>SiO<sub>3</sub>, 0.5 g liter<sup>-1</sup> NaHCO<sub>3</sub>, 0.5 g liter<sup>-1</sup> cysteine-HCl, and 0.001 g liter<sup>-1</sup> resazurin. One milliliter liter<sup>-1</sup> of Holden's trace element/Fe-EDTA solution<sup>44</sup> and 1 ml liter<sup>-1</sup> of Balch's vitamin solution<sup>45</sup> were added as a supplement to the medium. After autoclaving, the medium was kept in an anaerobic chamber (Coy Laboratory Products, Grass Lake, MI) filled with an anoxic gas mixture (N<sub>2</sub>:H<sub>2</sub>:CO<sub>2</sub>, 90:5:5) to equilibrate, and the final pH of the medium was adjusted to 6.5 with 2 N HCl. For the cultures in serum bottles, the media were reduced with 0.005% Na<sub>2</sub>S·9H<sub>2</sub>O and the headspaces were filled with 100% CO (MM1-CO). The serum bottles were sealed with bromobutyl rubber stoppers and aluminum crimp caps. For the adaptive evolution, the parental strain was cultured on MM1-CO medium at 80 °C for 20 h and then a 2% (v/v) cells were inoculated into a fresh medium using a sterile syringe. At every tenth transfer, cell cultures were stoked at 4 °C for later use. For batch culture, the strain was cultured in a 3 L bioreactor (Fermentec, Cheongwon, Korea), whose working volume was 2 L at 300 rpm agitation speed. The batch culture was maintained at pH 6.1–6.2 using 0.2 M NaOH in 3.5% NaCl. Using a mass flow controller, 100% CO was supplied at feeding rates of 400, 800 or 1000 ml min<sup>-1</sup> (MKPrecision, Seoul, Korea). Coal-gasified syngas (30.9% CO, 20.4% H<sub>2</sub>, 6.95% CO<sub>2</sub>, 1.42% CH<sub>4</sub>, and 32.1% N<sub>2</sub>) was provided by the Institute for Advanced Engineering in Korea and supplied at feeding rates of 160, 320, 480 or 640 ml min<sup>-1</sup>. The cultures were maintained under anaerobic conditions by preparing the media in an anaerobic chamber (Coy Laboratory Products, Grass Lake, USA) filled with an anoxic gas mixture (N<sub>2</sub>:H<sub>2</sub>:CO<sub>2</sub>, 90:5:5), and the bioreactor was sparged with pure argon gas (99.999%) through a microsparger.

**Construction of mutants.** Mutants of the TON\_1525 and TON\_0820 genes were made by applying a gene recombination system. Briefly, we designed primer sets for base-pair substitutions and mutated the TON\_1525 and TON\_0820 genes by site-directed mutagenesis. Each mutated gene and its flanking regions were ligated by one-step sequence- and ligation-independent cloning (SLIC)<sup>46</sup>, and subsequent mutants were generated through homologous recombination using an unmarked in-frame deletion<sup>15</sup> method and a modified gene disruption system that has been previously used for *Thermococcus kodakarensis* KOD1<sup>47</sup>. The mutations were verified by PCR using the primers listed in Supplementary Table 3 for Sanger sequencing.

**Preparation of cell suspensions and myoglobin assay.** To analyze CO consumption rate, a myoglobin assay was conducted by modifying a previously reported protocol<sup>48,49</sup>. To prepare cell suspensions, cells were cultured on MM1-CO medium and then harvested by centrifugation at  $7500 \times g$  for 20 min at room temperature under anaerobic conditions. Cell pellets were washed twice and resuspended in MM1 medium. Final cell density (expressed as optical density at 600 nm) was measured with a BioPhotometer plus UV-Vis spectrophotometer (Eppendorf, Hamburg, Germany) and adjusted to 0.3. Aliquots of 100  $\mu$ l of the cell suspension were mixed with 0.9 ml of CO-saturated MM1 medium and incubated at 80 °C. During the reaction, 10  $\mu$ l of the reaction solution was collected every 3 mins and immediately added to a quartz cuvette containing a myoglobin reaction mixture that included 56  $\mu$ M myoglobin (Sigma-Aldrich, St. Louis, UAS) from equine heart dissolved in 10 mM sodium thionite in 0.1 M potassium phosphate buffer (pH 7). After a 3 min incubation, the absorbance at 423 nm was measured to analyze consumed CO.

**Purification of the proteins encoded by TON\_1525(WT) and TON\_1525(T55I).** TON\_1525 and TON\_1525(T55I) plasmids were transformed into an *E. coli* BL21 Rosetta (DE3) pLysS strain. Gene expression was induced by IPTG at the mid-exponential growth phase, followed by a 3 h incubation at 37 °C. The cells were harvested by centrifugation ( $6,000 \times g$  at 4 °C for 20 min) and resuspended in 50 mM Tris-HCl buffer (pH 8.0) containing 0.1 M KCl and 10% (v/v) glycerol. The cells were disrupted by sonication and then centrifuged at  $15,000 \times g$  at 4 °C for 30 min. The resulting supernatant was applied to a column loaded with TALON metal affinity resin (BD Clontech, Mountain View, CA) and washed with 10 mM imidazole in 50 mM Tris-HCl buffer (pH 8.0) containing 0.1 M KCl and 10% (v/v) glycerol. The enzyme was eluted with the same buffer but with the addition of 300 mM imidazole. Protein concentration was determined by Bradford assay, and protein purity was evaluated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

**Electrophoretic mobility shift assay (EMSA).** DNA probes for EMSA were produced from *T. onnurineus* NA1 genomic DNA. A 150-bp probe corresponding to the CODH promoter was amplified using the primer pair labeled\_FAM\_1017\_150\_F and 1017\_150\_R. For competition analysis, a specific 150-bp cold competitor probe was derived from the same CODH promoter region as above via amplification using the primer pair unlabeled\_FAM\_1017\_150\_F and 1017\_150\_R. Poly(deoxyinosinic-deoxycytidylic acid) (Sigma-Aldrich, St. Louis, USA) and pUC18 plasmid DNA (New England BioLabs, Ipswich, MA) digested with *Hpa* II were used as non-specific competitors in 100 $\times$  molar excess. The PCR products were purified using a Qiagen PCR purification kit (Qiagen, GmbH, Hilden, Germany). EMSA reactions including 100 nM DNA with varying amounts of protein were set up in 20  $\mu$ l of EMSA buffer (20 mM Tris-HCl, 200 mM KCl, 1 mM EDTA, 5% glycerol, pH 7.5). After reaction at 80 °C for 10 min, the samples were immediately loaded onto a non-denaturing 5% polyacrylamide gel. The resulting gel was stained with ethidium bromide (EtBr) and analyzed to visualize the presence of proteins in the shifted bands. All of the primers that were used are listed in Supplementary Table 3.

**Reverse transcription-quantitative PCR (RT-qPCR) and western blotting.** To eliminate genomic DNA from total RNA preparations, 8  $\mu$ g of RNA was incubated with 8 units of RNase-free DNase I (Thermo Scientific Fermentas, St. Leon-Rot, Germany) at 37 °C for 30 min and purified via chloroform extraction and ethanol precipitation. RNA was quantified with a NanoDrop 2000 UV-Vis spectrophotometer (Thermo Scientific, West Palm Beach, USA), and cDNA was created from 1  $\mu$ g of RNA by incubation with 40 units of Moloney murine leukemia virus reverse transcriptase (Thermo Scientific Fermentas, St. Leon-Rot, Germany), 5  $\mu$ M random hexamers, and 1 mM deoxynucleoside triphosphate (dNTP) at 37 °C for 1 h in reverse transcription buffer. The reaction products were serially diluted to find the adequate concentration for real-time PCR analysis, and the samples were amplified with SYBR green real-time PCR master mix (Toyobo, Osaka, Japan). Amplified signals were detected using the StepOnePlus system (Applied Biosystems, Foster City, USA), and all primers that were used are listed in Supplementary Table 3. The relative amount of each gene was calculated based on cycle threshold (CT) values using a relative standard curve after normalization against the corresponding quantity of 16 S rRNA (TON\_1979). Western blots were prepared and analyzed using a chemiluminescent dye from an Immun-Star horseradish peroxidase (HRP) chemiluminescent kit (Bio-Rad, Hercules, USA). Polyclonal antibodies were produced by immunizing rabbits with recombinant TON\_1018 or TON\_1023 proteins according to established procedures for producing custom rabbit antibodies (Young In Frontier Co., Seoul, Korea) and purified through His-Bind resin (Novagen, Madison, USA).

**Genome sequencing.** For genome re-sequencing, we extracted genomic DNA from cultures of the 156 T strain without single-colony isolation. Genome sequencing was performed using Illumina HiSeq-2000 (Macrogen, Seoul, Korea) and PacBio Single Molecule Real-Time (SMRT) sequencing (Pacific Biosciences, Menlo Park, USA)<sup>50</sup>. HiSeq-2000 sequencing was performed using barcoded adaptors for multiplexed paired-end 101 bp sequencing and provided 100 $\times$  coverage of the 156 T genome. The resulting genomic sequence data were processed using CASAVA-1.8.2 and BWA v0.5.9-r16 software (Illumina, San Diego, USA) for alignment. PCR duplicates were removed by Picard v1.79, and variants were detected using SAMtools v0.1.18. PacBio SMRT sequencing of a 10-kb insert library provided approximately 100 $\times$  coverage. Assembly and consensus polishing were performed using SMRTpipe HGAP and SMRTpipe Quiver, respectively. All mutations were verified by PCR and Sanger sequencing, and all primers are listed in Supplementary Table 3.

**RNA extraction and RNA sequencing.** RNA was prepared from exponentially growing cells that were transferred 2, 12, 32, 62, 102, 122 or 156 times into CO-containing media. RNA extraction was performed using TRIzol reagent (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions with some modifications, and the quantity and quality of the total RNA were evaluated using RNA electropherograms (Agilent 2100 Bioanalyzer, Palo Alto, USA) and RNA integrity numbers (RINs)<sup>51</sup>. A total of 10  $\mu$ g of total RNA over an RIN

value of 8.0 from each sample was used as a starting material and treated with a Ribo-Zero rRNA Removal Kit for Bacteria (Epicentre, Madison, USA). The resulting mRNA samples were processed into sequencing libraries using a TruSeq Stranded Total RNA Sample Prep Kit (Illumina, San Diego, USA) following the manufacturer's protocol. Sequencing was performed using a HiSeq 2500 (Illumina, San Diego, USA) to generate directional, paired-end 100 base pair reads. Quality-filtered reads were mapped to a National Center for Biotechnology Information (NCBI) reference genome sequence (BioProject ID PRJNA59043) using CLC Genomics Workbench 6.5 (CLC bio, Aarhus, Denmark). The raw RNA-seq data have been deposited in the NCBI Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) under the accession code GSE73031.

**Identification of differentially expressed genes (DEGs).** To evaluate DEGs, both the DESeq and edgeR packages were used<sup>52,53</sup>. Replicate experiments were performed only for the initial culture, which was cultured only 2 times. Fold changes in gene expression in a selected series of adapted culture samples were calculated based on comparisons to the initial culture sample. A false discovery rate (FDR) of less than 0.001 and a fold change of either more than 2 or less than 0.5 were used as the criteria for selection, and only genes that passed the thresholds from both algorithms were identified as DEGs. To interpret the functions of the identified DEGs, 819 genes were classified according to the archaeal cluster of orthologous genes (arCOGs)<sup>27</sup>. To analyze changes in transcript patterns of pathway-related genes, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (<http://www.genome.jp/kegg/>) was also used.

**Methylation site analysis.** Using a Covaris G-tube (Covaris, Inc., Woburn, MA, USA), we generated 20-kb DNA fragments by shearing genomic DNA and removed small fragments using an AMPureXP bead purification system (Beckman Coulter Inc, Brae, CA, USA) according to the manufacturer's protocol. A total of 5 µg of each sample was used as input for library preparation. A SMRTbell library was constructed using a SMRTbell™ Template Prep Kit 1.0 (Pacific Biosciences, Menlo Park, USA). Using the BluePippin Size selection system (Sage Science, Inc., Beverly, MA, USA), we removed the small fragments to create a large-insert library. After a sequencing primer annealed to the SMRTbell template, the SMRTbell libraries were bound to polymerases using a DNA/Polymerase Binding kit P4 (Pacific Biosciences, Menlo Park, USA). Following the polymerase binding reaction, a MagBead Kit (Pacific Biosciences, Menlo Park, USA) was used to bind the library complex to MagBeads before sequencing. MagBead-bound complexes provide more reads per SMRT Cell. The polymerase-SMRTbell-adaptor complex was then loaded into zero-mode waveguides (ZMWs). The SMRTbell library was sequenced using 2 SMRT cells (Pacific Biosciences, Menlo Park, USA) with C2 chemistry (DNA sequencing Reagent 2.0), and 1 × 180 min movies were captured for each SMRT cell using a PacBio RS sequencing platform (Pacific Biosciences, Menlo Park, USA). To determine methylation sites, Pacific Biosciences SMRTanalysis software (version 2.1.1) was used, and the method depended on the sensitivity of the polymerase kinetics to the DNA template structure, as DNA synthesis was recorded in real time. An *in silico* kinetic reference and t-test based kinetic score were calculated based on interpulse duration (IPD) for detection of modified base positions.

**Bioinformatics analysis.** A homology search was performed using the Basic Local Alignment Search Tool (BLAST; <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) against a nonredundant protein database maintained by the National Center for Biotechnology Information (NCBI)<sup>54</sup>. Multiple sequence alignment was performed using a pairwise sequence alignment (<http://www.ebi.ac.uk/Tools/psa/>) tool from the European Bioinformatics Institute (EBI).

## References

- Lee, J. W., Kim, T. Y., Jang, Y. S., Choi, S. & Lee, S. Y. Systems metabolic engineering for chemicals and materials. *Trends in Biotechnol.* **29**, 370–378 (2011).
- Çakar, Z. P., Turanlı-Yıldız, B., Alkim, C. & Yılmaz, Ü. Evolutionary engineering of *Saccharomyces cerevisiae* for improved industrially important properties. *FEMS Yeast Res.* **12**, 171–182 (2012).
- Winkler, J. D. & Kao, K. C. Recent advances in the evolutionary engineering of industrial biocatalysts. *Genomics* **104**, 406–411 (2014).
- Adams, J. & Rosenzweig, F. Experimental microbial evolution: history and conceptual underpinnings. *Genomics* **104**, 393–398 (2014).
- Conrad, T. M., Lewis, N. E. & Palsson, B. Ø. Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol.* **7**, 509 (2011).
- Hindré, T., Knibbe, C., Beslon, G. & Schneider, D. New insights into bacterial adaptation through *in vivo* and *in silico* experimental evolution. *Nat Rev Microbiol.* **10**, 352–365 (2012).
- Dettman, J. R. *et al.* Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Mol Ecol.* **21**, 2058–2077 (2012).
- Latif, H. *et al.* Adaptive evolution of *Thermotoga maritima* reveals plasticity of the ABC transporter network. *Appl. Environ. Microbiol.* **81**, 5477–5485 (2015).
- Hallenbeck, P. C., Abo-Hashesh, M. & Ghosh, D. Strategies for improving biological hydrogen production. *Bioresour Technol.* **110**, 1–9 (2012).
- Nath, K. & Das, D. Improvement of fermentative hydrogen production: various approaches. *Appl. Microbiol. Biotechnol.* **65**, 520–529 (2004).
- Amend, J. P. & Shock, E. L. Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiol Rev.* **25**, 175–243 (2001).
- Lee H. S. *et al.* The complete genome sequence of *Thermococcus onnurineus* NA1 reveals a mixed heterotrophic and carboxydrotrophic metabolism. *J. Bacteriol.* **190**, 7491–7499 (2008).
- Rittmann, S. K.-M. R. *et al.* One-carbon substrate-based biohydrogen production: microbes, mechanism, and productivity. *Biotechnol. Adv.* **33**, 165–177 (2015).
- Kim, M. S. *et al.* CO-dependent H<sub>2</sub> production by genetically engineered *Thermococcus onnurineus* NA1. *Appl. Environ. Microbiol.* **79**, 2048–2053 (2013).
- Kim, M.-S. *et al.* A novel CO-responsive transcriptional regulator and enhanced H<sub>2</sub> production by an engineered *Thermococcus onnurineus* NA1 strain. *Appl. Environ. Microbiol.* **81**, 1708–1714 (2015).

16. Lee, S. H. *et al.* Screening of a novel strong promoter by RNA sequencing and its application to H<sub>2</sub> production in a hyperthermophilic archaeon. *Appl. Microbiol. Biotechnol.* **99**, 4085–4092 (2015).
17. Lee, S. H. *et al.* Comparison of CO-dependent H<sub>2</sub> production with strong promoters in *Thermococcus onnurineus* NA1. *Appl. Microbiol. Biotechnol.* **98**, 979–986 (2014).
18. Flusberg, B. A. *et al.* Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods.* **7**, 461–465 (2010).
19. Casadesús, J. & Low, D. Epigenetic gene regulation in the bacterial world. *Microbiol. Mol. Biol. Rev.* **70**, 830–856 (2006).
20. Fang, G. *et al.* Genome-wide mapping of methylated adenine residues in pathogenic *Escherichia coli* using single-molecule real-time sequencing. *Nat. Biotechnol.* **30**, 1232–1239 (2012).
21. Bendall, M. L. *et al.* Exploring the roles of DNA methylation in the metal-reducing bacterium *Shewanella oneidensis* MR-1. *J. Bacteriol.* **195**, 4966–4974 (2013).
22. Ouellette, M., Jackson, L., Chimileski, S. & Papke, R. T. Genome-wide DNA methylation analysis of *Haloferax volcanii* H26 and identification of DNA methyltransferase related PD-(D/E)XK nuclease family protein HVO\_A0006. *Front. Microbiol.* **6**, 1–11 (2015).
23. Torkelson, J. *et al.* Genome-wide hypermutation in a subpopulation of stationary-phase cells underlies recombination-dependent adaptive mutation. *EMBO J.* **16**, 3303–3311 (1997).
24. Weiner, L. & Model, P. Role of an *Escherichia coli* stress-response operon in stationary-phase survival. *Proc. Natl. Acad. Sci. USA* **91**, 2191–2195 (1994).
25. Kivisaar, M. Stationary phase mutagenesis: Mechanisms that accelerate adaptation of microbial populations under environmental stress. *Environ Microbiol* **5**, 814–827 (2003).
26. Loewe, L., Textor, V. & Scherer, S. High deleterious genomic mutation rate in stationary phase of *Escherichia coli*. *Science* **302**, 1558–1560 (2003).
27. Wolf, Y. I., Makarova, K. S., Yutin, N. & Koonin, E. V. Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. *Biol. Direct* **7**, 46 (2012).
28. Nad, D., Oxidoreductase, P. H. S., Schut, G. J., Bridger, S. L. & Adams, M. W. W. Insights into the metabolism of elemental sulfur by the hyperthermophilic archaeon *Pyrococcus furiosus*: characterization of a coenzyme A-dependent NAD(P)H sulfur oxidoreductase. *J. Bacteriol.* **189**, 4431–4441 (2007).
29. Santangelo, T. J., Čuboňová, L. & Reeve, J. N. Deletion of alternative pathways for reductant recycling in *Thermococcus kodakarensis* increases hydrogen production. *Mol. Microbiol.* **81**, 897–911 (2011).
30. Kanai, T. *et al.* Distinct physiological roles of the three [NiFe]-hydrogenase orthologs in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *J. Bacteriol.* **193**, 3109–3116 (2011).
31. Lim, J. K., Mayer, F., Kang, S. G. & Müller, V. Energy conservation by oxidation of formate to carbon dioxide and hydrogen via a sodium ion current in a hyperthermophilic archaeon. *Proc. Natl. Acad. Sci. USA* **111**, 11497–11502 (2014).
32. Kaminski, L. & Eichler, J. *Haloferax volcanii* N-glycosylation: Delineating the pathway of dTDP-rhamnose biosynthesis. *Plos One.* **9**, e97441 (2014).
33. Lipscomb, G. L. *et al.* SurR: A transcriptional activator and repressor controlling hydrogen and elemental sulphur metabolism in *Pyrococcus furiosus*. *Mol. Microbiol.* **71**, 332–349 (2009).
34. Kim, Y. J. *et al.* Formate-driven growth coupled with H<sub>2</sub> production. *Nature* **467**, 352–355 (2010).
35. Mackwan, R. R., Carver, G. T., Kissling, G. E., Drake, J. W. & Grogan, D. W. The rate and character of spontaneous mutation in *Thermus thermophilus*. *Genetics* **180**, 17–25 (2008).
36. Grogan, D. W., Carver, G. T. & Drake, J. W. Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon *Sulfolobus acidocaldarius*. *Proc. Natl. Acad. Sci. USA* **98**, 7928–7933 (2001).
37. Hochheimer, A., Hedderich, R. & Thauer, R. K. The DNA binding protein Tfx from *Methanobacterium thermoautotrophicum*: structure, DNA binding properties and transcriptional regulation. *Mol. Microbiol.* **31**, 641–650 (1999).
38. Ehrlich, M. *et al.* DNA methylation in thermophilic bacteria: N4-methylcytosine, 5-methylcytosine, and N5methyladenine. *Nucleic Acids Res.* **13**, 1399–1412 (1985).
39. Ehrlich, M., Wilson, G. G., Kuo, K. C. & Gehrke, C. W. N4-methylcytosine as a minor base in bacterial DNA. *J. Bacteriol.* **169**, 939–943 (1987).
40. Connolly, B. A., Fogg, M. J., Shuttleworth, G. & Wilson, B. T. Uracil recognition by archaeal family B DNA polymerases. *Biochem. Soc. Trans.* **31**, 699–702 (2003).
41. Cho, W. *et al.* Optimal design and operation of a natural gas tri-reforming reactor for DME synthesis. *Catal. Today* **139**, 261–267 (2009).
42. Bae, S. S. *et al.* *Thermococcus onnurineus* sp. nov., a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent area at the PACMANUS field. *J. Microbiol. Biotechnol.* **16**, 1826–1831 (2006).
43. Holden, J. F. *et al.* Diversity among three novel groups of hyperthermophilic deep-sea *Thermococcus* species from three sites in the northeastern Pacific Ocean. *FEMS Microbiol. Ecol.* **36**, 51–60 (2001).
44. Sokolova, T. G. *et al.* The first evidence of anaerobic CO oxidation coupled with H<sub>2</sub> production by a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *Extremophiles* **8**, 317–323 (2004).
45. Balch, W. E. & Wolfe, R. S. New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere. *Appl. Environ. Microbiol.* **32**, 781–791 (1976).
46. Jeong, J. Y. *et al.* One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies. *Appl. Environ. Microbiol.* **78**, 5440–5443 (2012).
47. Matsumi, R., Manabe, K., Fukui, T., Atomi, H. & Imanaka, T. Disruption of a sugar transporter gene cluster in a hyperthermophilic archaeon using a host-marker system based on antibiotic resistance. *J. Bacteriol.* **189**, 2683–2691 (2007).
48. Riggs, S. S. & Heindel, T. J. Measuring carbon monoxide gas-liquid mass transfer in a stirred tank reactor for syngas fermentation. *Biotechnol. Prog.* **22**, 903–906 (2006).
49. Kundu, S., Premer, S. A., Hoy, J. A., Trent, J. T. & Hargrove, M. S. Direct measurement of equilibrium constants for high-affinity hemoglobins. *Biophys. J.* **84**, 3931–3940 (2003).
50. Eid, J. *et al.* Real-time DNA sequencing from single polymerase molecules. *Science* **323**, 133–138 (2009).
51. Schroeder, A. *et al.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol. Biol.* **7**, 3 (2006).
52. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
53. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* **26**, 139–140 (2010).
54. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).

## Acknowledgements

This work was supported by the KIOST in-house program (PE99413), the Development of Biohydrogen Production Technology Using the Hyperthermophilic Archaea program of the Ministry of Oceans and Fisheries in the Republic of Korea, and the Marine Genomics 100+ Korea program of the Ministry of Oceans and Fisheries in the Republic of Korea.

### Author Contributions

H.S.L., S.G.K. and T.W.K. conceptualized and designed the experiments. S.H.L. and M.-S.K. performed most of the experiments and analysed the data. J.-H.L. and Y.-J.C. performed bioinformatics analyses. S.S.B. and S.-M.L. carried out kinetic analysis using synthetic gas. H.C.J., T.-J.Y. and A.R.C. performed construction of mutants. J.-H.L. and K.K.K. contributed critical comments on the manuscript. S.H.L., M.-S.K., H.S.L. and S.G.K. wrote the paper with input from the co-authors.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Lee, S. H. *et al.* Adaptive engineering of a hyperthermophilic archaeon on CO and discovering the underlying mechanism by multi-omics analysis. *Sci. Rep.* **6**, 22896; doi: 10.1038/srep22896 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>