A Novel CO-Responsive Transcriptional Regulator and Enhanced H_2 Production by an Engineered *Thermococcus onnurineus* NA1

Min-Sik Kim,[a] Ae Ran Choi,[a] Seong Hyuk Lee,[a,b] Hae-Chang Jung, [a,b] Seung Seob Bae,[a,b] Tae-Jun Yang,[a] Jeong Ho Jeon,[a] Jae Kyu Lim,[a] Hwan Youn, [c] Tae Wan Kim,[a,b] Hyun Sook Lee,[a,b] Sung Gyun Kang[a,b]

Korea Institute of Ocean Science and Technology, Ansan, South Korea,[a] and Department of Marine Biotechnology, University of Science and Technology, Daejeon, South Korea[b]

Department of Biology, California State University, Fresno, California, USA[c]

Address correspondence to Hyun Sook Lee, leeh522@kiost.ac, or Sung Gyun Kang sgkang@kiost.ac.

M.S.K, A.R.C. and S.H.L. contributed equally to this article.

Running title: A novel CO-dependent transcriptional regulator

Supplemental Material

Table S1 Primers used in this study

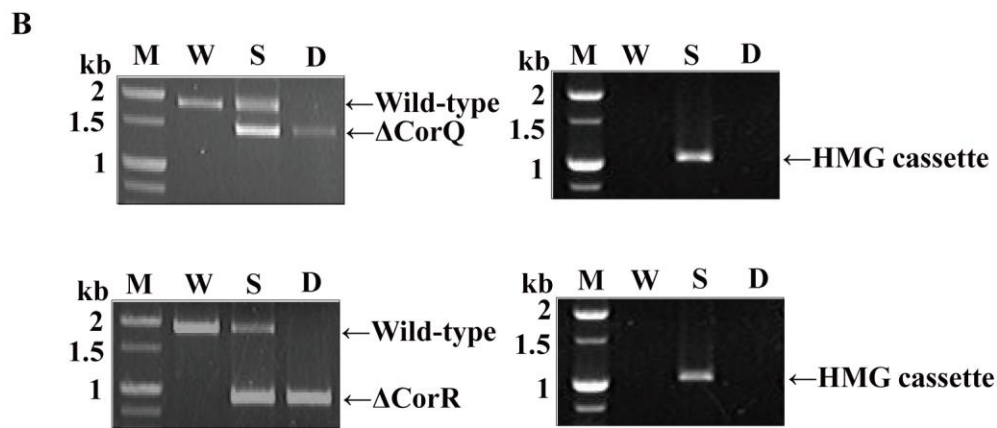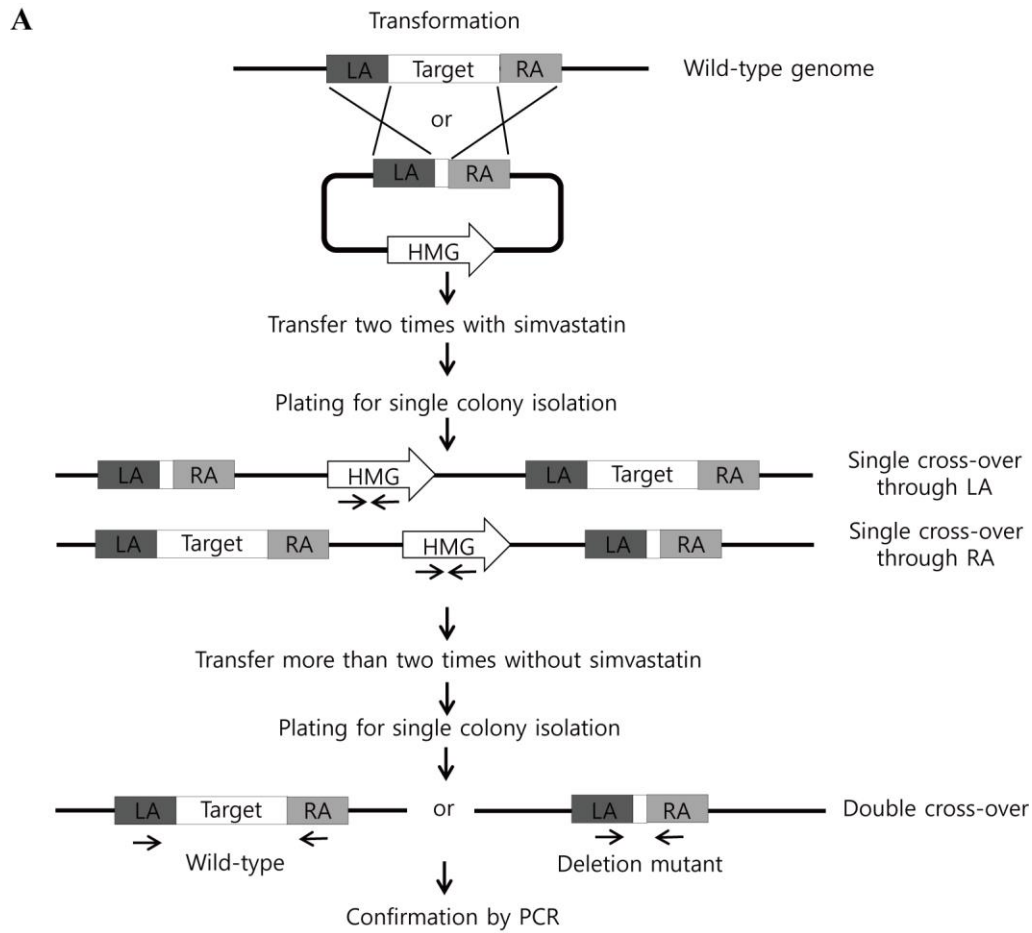| Primer	Oligonucleotide sequences	
RT-qPCR		
TON_1018	5'-gttcgagaatcctgctggtctt-3'	5'-agcaactggcaagtctgaaatg-3'
corR (TON_1016)	5'-agatggggttcagattcgatgaag-3'	5'-tgtccctcagctcatcgtagataa-3'
Construction of mutant		
TON_1016- inverse PCR	5'-catccccagggagcattact-3'	5'-aaaaggaataactgtactca-3'
TON_1016- LA and RA	5'-caggacactcgaagaggtagtccag-3'	5'-gttccttgctactattgtatctgca-3'
TON_1015-1-LA	5'-gccagtgccaagcttgcatgcaagttgagagggcttttctttaatc-3'	5'-tacgaattcgagctcggtaccaaggccagaccttgaagagctttac-3'
TON_1015-1-RA	5'-aagccctctcaacttgcatgcatagaagagctagtaggcagaaaca-3'	5'-gccagtgccaagcttgcatgctgatggagaaactcccataatctta-3'
Construction of pQRc		
pQRc-LA	5'-gggagatgactgcaggcatgcgtggc-3'	5'-gccagtgccaagcttgcatgcaatgg-3'
pQRc-RA	5'-ttacgaattggatccggtaccccagac-3'	5'-ctcactctcgttccaggtaccttatcc-3'
pQRc-inverse PCR	5'-tcatctcccaagcattttat-3'	5'-ctgcaggcatgcgtggcgtc-3'
corQR-insert	5'-tgtatgcatactctttcataggtgt-3'	5'-gattttaagatggaatcgtgttcgg-3'
Confirmation of mutant construction		
TON_1015-1-confirm	5'-cactttcatagcagaataaa-3'	5'-ctcgtatggcctgaacttct-3'
TON_1016-confirm	5'-aagacttgtagaaagtggag-3'	5'-ttgttggcgttaaggacaac-3'

HMG-confirm   5'-atgaacgtaaaagtcctttg-3'   5'-ttctttaggtagtttacttc-3'

TABLE S2 Kinetic parameters of wild-type, ΔCorQ/$corQR^{\uparrow}$, and ΔCorR/$corQR^{\uparrow}$ strains.

Kinetic parameter[a]	Wild-type	ΔCorQ/$corQR^{\uparrow a}$	ΔCorR/$corQR^{\uparrow a}$
μ_{max} (h^{-1})	0.4	0.3	1.4
r_{max} (mmol liter^{-1} h^{-1})	33.0	46.0	191.9
Biomass productivity (g liter^{-1} h^{-1})[b]	0.06	0.08	0.29
q_{max} (mmol g^{-1} h^{-1})	151.3	92.4	249.6
H_2 productivity (mmol liter^{-1} h^{-1})[b]	33.1	41.8	168.4
$Y_{p/x}$ (g H_2 g^{-1} biomass)	0.4	0.7	1.2

[a] Kinetic parameters were calculated with data from graphs in Fig. 4. μ_{max}, maximum specific growth rate; r_{max}, maximum H_2 production rate; q_{max}, maximum specific H_2 production rate; $Y_{p/x}$, product yield coefficient with respect to biomass.

[b] Productivity was determined by dividing total yield by time difference from 8 to 11 h for the wild-type strain and from 6.5 to 9.5 h for ΔCorQ/$corQR^{\uparrow}$ and from 4.5 to 11 h for ΔCorR/$corQR^{\uparrow}$.

A

Transformation

Wild-type genome

or

Transfer two times with simvastatin

Plating for single colony isolation

Single cross-over through LA

Single cross-over through RA

Transfer more than two times without simvastatin

Plating for single colony isolation

Double cross-over

Wild-type

Deletion mutant

Confirmation by PCR

B

kb M W S D
←Wild-type
←ΔCorQ

kb M W S D
←HMG cassette

kb M W S D
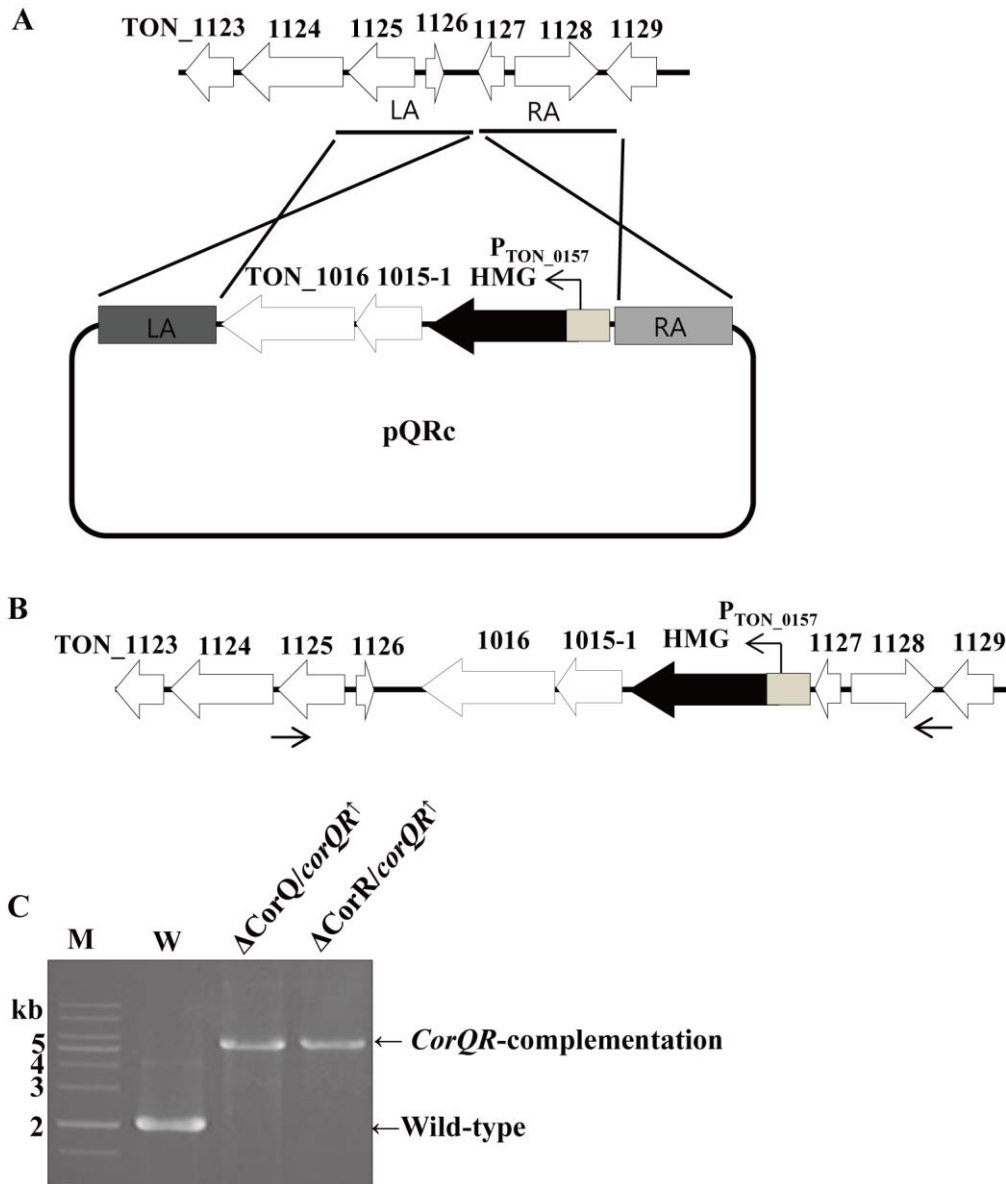←Wild-type
←ΔCorR

kb M W S D
←HMG cassette

1

2

3 **FIG S1** Gene disruption. (A) Schematic diagram showing the strategy for in-frame deletions

4 of *corR* and *corQ* genes. Cells were transformed with the vector containing a deleted target

5

5   gene flanked by two ~1 kb DNA fragments suitable for homologous recombination (LA, left

6   arm; RA, right arm). After two-times plating with 10 μM simvastatin, single colonies, where

7   single cross-over occurred, were selected and confirmed by PCR. After two-times more

8   plating without simvastatin, cells having mutated target gene were selected. Locations of

9   primers used for confirmation were marked by black arrows under corresponding genes. (B)

10   PCR confirmation of ΔCorQ (two upper pannels) and ΔCorQ (two lower pannels) mutants.

11   DNA bands specific for wild-type, ΔCorQ, ΔCorR, and HMG cassette were indicated by

12   arrows at the right sides of figures. M, DNA size marker; W, genomic DNA of the wild-type

13   strain; S, genomic DNA of cells harboring the integrated vector; D, genomic DNA of the

14   mutants.

15

A

TON_1123 1124 1125 1126 1127 1128 1129

LA   RA

TON_1016 1015-1 HMG P_{TON_0157}

LA   RA

pQRc

B

TON_1123 1124 1125 1126   1016 1015-1 HMG P_{TON_0157} 1127 1128 1129

C

M   W   ΔCorQ/*corQR*↑   ΔCorR/*corQR*↑

kb
5
4
3

← *CorQR*-complementation

2

← Wild-type

FIG S2 Construction of ΔCorQ/*corQR*↑ and ΔCorR/*corQR*↑ strains. (A) Schematic diagram showing the strategy for integrating *corQR* genes into the intergenic region between TON_1126 and TON_1127. pQRc vector was constructed to contain *corQR* and $P_{TON\_0157}HMG_{pfu}$ cassette genes flanked by 1 kb DNA regions for homologous recombination which were designated as LA (left arm) and RA (right arm). (B) A schematic diagram

7

23   showing the genome context of *corQR*-complemented strains after double cross-over. The

24   positions of primer used for confirmation of mutant construction were indicated by arrows.

25   (C) PCR confirmation of *corQR* integration. DNA bands corresponding to wild-type and

26   integrated sequences were indicated by arrows. M, DNA size marker; W, genomic DNA of

27   the wild-type strain; ΔCorQ/*corQR*$^\uparrow$, ΔCorQ mutant containing integrated *corQR* sequence;

28   ΔCorR/*corQR*$^\uparrow$, ΔCorR mutant containing integrated *corQR* sequence.

29

A

```
TERMP_01155  262  LELLLELISQVNSFRAACKALGVSPSTYWEKIRSLEEKLGISLIISVRGGRKKGITILTEFA  319
TES1_1213    262  LELLLELISQVNSFRAACKALGVSPSTYWERIKNLEEKLGVALLISVRGGRKKGITILTDFA  319
TAM4_1057    295  MELLLELISQLKSFRAACKAIGVSPSTYWERIKELEYKLGMSLLITVRGGRKKGITVLTGFA  355
TON_1016     285  MELLLELISQTRSFRAACKIMGVSPSTYWERIRDIEEKLGRRLIVSVKGGRKKGITVLTGVA  345
Kcr_0760     284  LELLSLIEELGSLSQAARTIGATPSSIFKRIRRMEGVLGLRLITSSRGGYLRGGVRLTPEC  344
```

```
TERMP_01155  320  KDLLKEYREIREKAI----VSLYEY---  340
TES1_1213    320  KELLREYKHIREKVL----VSLYEYK--  341
TAM4_1057    356  RDILEEYRRVREKVL----LSLYT----  375
TON_1016     346  LDLLKEYQRIRERVL----LSLNERF--  367
Kcr_0760     345  EGLVRRYRELKVSIINRYRLSLMEKLDG  372
```

B

```
TERMP_01156  122  MKVNIYECISCYRAPPIGRTLCDFEAGLIQGVMEELVGKNVT--REIYCWGLGNSFCGFEV  183
TES1_1214     58  MKVNIYECMSCYKAPLIGRTLCDFEAGLIQGVMEKLIGKNIT--RELYCWGLGNSFCGFEV  119
TON_1015-1    63  IKLNIYECMSCYHTVPIGRTLCDFEAGLIQGIIEELVGRNIT--REIYCQGLGYSFCGFEV  124
TAM4_1099     58  MKVNLYECISCYNIKPVGRTLCDFEAGFIQGVMEALIGKNIT--REVYCWGLGNHFCGFEV  119
Kcr_0759     117  IRITVKDSFEARGRR-ADAPVCHFVAGILSAIVEDVFSIRTGPLVEESCAATGNDFCTFSA  186
                                       *                    *         *
```

30

31   **FIG S3** Multiple sequence alignments of the helix-turn-helix (HTH) motifs of CorR proteins

32   (A) and the V4R domains of CorQ proteins (B). Shown from top to bottom (GeneBank

33   accession numbers in parentheses) are: TERMP_01155 (ADT84131) from *T. barophilus* MP,

34   TES1_1213 (AHF80595) from *Thermococcus* sp. ES1, TAM4_1057 (EEB72950) from

35   *Thermococcus* sp. AM4, TON_1016 (ACJ16504) from *T. onnurineus* NA1, Kcr_0760

36   (ACB07508) from *Ca.* K. cryptofilum OPF8, TERMP_01156 (ADT84132) from *T.*

37   *barophilus* MP, TES1_1214 (AHF80596) from *Thermococcus* sp. ES1, TON_1015-1-

38   (KM489057) from *T. onnurineus* NA1, TAM4_1099 (EEB72992) from *Thermococcus* sp.

39   AM4, Kcr_0759 (ACB07507) from *Ca.* K. cryptofilum OPF8. The conserved residues are

40   shown with black backgrounds (white letters) and three conserved cysteine residues are

41   indicated by asterisks.