

## Additional File 5. Phylogenetic footprinting of upstream regions of candidate HisR-regulated genes (see legend below).

fig|1196029.3.peg.4456  
fig|1006007.3.peg.213  
fig|412384.9.peg.5136

**Bacillus endophyticus 2102**  
**Bacillus megaterium WSH-002**  
**Bacillus aryabhattai strain T61**

```
fig|1196029.3.peg.4456 acagaaaatttgattatatttcatatattatatttgagta-tgaactgttgacgagggcgaaatgaaagatgtaaacactactaataactttaattcattatatttc
fig|1006007.3.peg.213 gaactaagctgtggtttttttgtaa-atttgaacagctggcgggtgacggctgactttatcgggtgtaaactatcaataactttaatttattatttcgc
fig|412384.9.peg.5136 gaactaagctgtggtttttttgtaa-atttgaacagctggcgggtgacggctgactttatcgggtgtaaactatcaataactttaattggtatttcgc
* ** * * *** ** ***** * * **** * * **** * **
```

YRCwTTAgYRWRnTAAAGtR

```
fig|1196029.3.peg.4456 taa ttcgttagtgtgataaagtt-----tttttatt---ttaaactaa-----aggatg-gttacgATGTCGAAGTTATTTATGT 4.09
fig|1006007.3.peg.213 taa tacattagtagtgataaagtaaaaacaacctatcaatggttcctaataagataaaaatcaagcttaaaggacgtgtcggaATGTCAAAGTTATTTATGT 5.44
fig|412384.9.peg.5136 taa tacattagcgtgataaagtaaaatcaacctatcaatggttcctaataagataaaaatcaagcttaaaggacgtgtcagaATGTCAAAGTTATTTATGT 5.44
**** * ***** ***** * ** * * * ***** * ** * **
```

fig|1230341.3.peg.778  
fig|866895.3.peg.864  
fig|2045014.3.peg.1092

**Salimicrobium sp. MJ3**  
**Halobacillus halophilus DSM 2266**  
**Halobacillus sp. Marseille-P3879**

YRCwTTAgYRWRnTAAAGtR

```
fig|1230341.3.peg.778 tgtgaaaggggtgatttttttatattttttcctttatcatggttaaactcgttaat-----gagAT-----gaaaggggtgagatgATGGCATCGGCTGT 4.15
fig|866895.3.peg.864 aataagaggggtgatttttttaattttactttattatggttaaactcgcataat---aaataaaagatagacgtagagaggagaatggattATGGCCTGGGCAGT 4.72
fig|2045014.3.peg.1092 aaaatgaggggtgatttttttgaaatatcctttattatggttaaactcgcataatgtgataaaaagagttaacttgagaggagattatagtATGGCACTAGCTGT
***** ***** * * ***** ***** ***** ***** * * * **
```

fig|1036673.3.peg.173  
fig|717606.6.peg.3631

**Paenibacillus mucilaginosus KNP414**  
**Paenibacillus curdlanolyticus YK9**

YWYWTTAnYwwRnTAAAGnr

```
fig|1036673.3.peg.173 ctaacgattacattccttcaccttggtagtatggtaacatatttagcggactaaagcgttctgaaatattcatga 3.8
fig|717606.6.peg.3631 gtaataaggagtcgctttaacgcatgaagcatcaacctaacagatga-tacgaattgatgataaatgaaaga 4.33
*** * * *** * * * ** * * * ** * * * ** * * * **
```

```
fig|1036673.3.peg.173 ctaaacaagatagat--aggggtgaactccaTTGTCCAAACCGAAGGTGTTTGAGAAGCCTACAGGCGTGAAGGACT
fig|717606.6.peg.3631 acaaacacgattagaagaagggtgagtcccGT-GTCGAAACCAAAGTATTTGAAAACCGATCGGCGTTAAAGATT
**** ***** ** * * ** * ** * * * * * * * * * * * * * * * * * * *
```





fig|1414720.3.peg.2641

Clostridium sp. JCC

fig|931276.5.peg.1495

Clostridium saccharoperbutylaceticum N1-4 (HMT)

```

nRyTTTAnYRYRnTAAArYr
fig|1414720.3.peg.2641 aataattgacaaataaaacatata----aggtattattgataataactt--taa cactttaatctgctaaggg-ggcgatttgaattaaaaatattat 4.24
fig|931276.5.peg.1495 aattttttagaattattgacatgcaatagaaaatactatattataaatttaatta cagtttagcacggttaaagcgataaaatataaatggataatagtgg 4.82
* * * * * * * * * * * * * * * * * * * * * * *

```

fig|1414720.3.peg.2641 accagaagga**ATGCGAGATTTTACGG**

fig|931276.5.peg.1495 gg--tggaaa**ATGAGTAAGAAGAATA**

\*\*\*\* \* \* \*

fig|1871029.3.peg.2807

Intestinibacillus massiliensis strain type strain: Marseille-P3216

fig|2364792.3.peg.822

Intestinibacillus sp. Marseille-P6563

```

nRyTTTAnYRYRnTAAArYr
fig|1871029.3.peg.2807 cccgcagccgccttttttttg---caaaattttttat tgatthttcttgaataaagtgtgt-acaatattaaa-gactgttcattttaaacttttagga----- 4.07
fig|2364792.3.peg.822 tacaaaaattctttggatcgactcgattttgctgt tgatthttctgaactaaattgtctacaataataacacggttcttgaaaagggatagtagaaagcgg 4.50
* * * * * * * * * * * * * * * * * * * * * * *

```

fig|1871029.3.peg.2807 ---ggactatactt**ATGAAGATGAAGCGCATTGCAGGCCTGGTGCTGGCC**

fig|2364792.3.peg.822 aaaggaatgtactt**ATGAAAGCAAAGAGACTCTTCAGCCTGCTTCTGGCG**

\*\*\* \* \*\*\*\*\* \* \* \* \*\*\*\*\* \* \*\*\*\*\*

fig|1121323.3.peg.2212

[Clostridium] lactatifermentans DSM 14214

fig|991789.4.peg.2525

[Clostridium] propionicum DSM 1682

fig|1123282.4.peg.2032

Sporobacter termitidis DSM 10068

```

nRyTTTAnYRYRnTAAArYr
fig|1121323.3.peg.2212 aaatatctgaataaaaacttagcaatttgtatatacatacaaactttaataatctac ttgactaaagtgtctaaactggcttaatatgaagaaaagca 3.70
fig|991789.4.peg.2525 aaaaaatagaaagataaaaacgagtttcagcaaatgatataacttttagtggttctac ttgactaaagtgtgtaactgatataagattaaagtaaat 3.76
fig|1123282.4.peg.2032 aaaaaagacgcgagaacaacttcgatgccaatgaat-tatcata--ttaccgaattac ttgactaaatgctaaagtatcgtatactgatgtcaccaa 3.76
*** * * * * * * * * * * * * * * * * * * * * * *

```

fig|1121323.3.peg.2212 tatg---ggaggtgcctt**ATGACTGATTGGGATGGATTATTGAAAAAAGAAGAACAGACGGTGT**

fig|991789.4.peg.2525 gacgtgcggaggtgcaac**ATGACATTGCCGAAAATGCTTTTGAAAAAAGAAGAGGAAGTTATAT**

fig|1123282.4.peg.2032 cagaaagtgagaagagc**ATGACCGTTTCAGAGGATGTCTTGAAAAACGAGGAGAAGCGGATGA**

\* \*\*\* \*\*\*\*\* \* \* \*\*\*\*\* \*\* \*\* \* \*

**fig|476272.5.peg.243** **Blautia hydrogenotrophica DSM 10507**  
**fig|411459.7.peg.1542** **Ruminococcus obeum ATCC 29174**  
**fig|1095771.3.peg.1573** **Ruminococcus sp. JC304**

fig|476272.5.peg.243 gggattgactttttcagcattttttcttataatctaactaatttgttttatgtgctaa **tatagtagcgagataaagta** ctgaatt--aacg----- 4.17  
fig|1095771.3.peg.1573 g--attgacattttcagaataaaccagtataatttta---tctagtttatgcatcac **tgtagtaacatggttagcatg** ctaaagcaaatgg-----  
fig|411459.7.peg.1542 gggattgacattttctattttttctactataatttt----cttgttttacatagtaa **catggttagcacggttaaagtg** cttatctgtaatga----- 4.87  
\*\*\*\*\* \*\* \* \*\*\*\*\* \* \*\* \* \* \* \* \* \* \* \* \* \* \* \* \*

fig|476272.5.peg.243 -----ggagagaatgat**A-TGGAAGAATTTTTACACACCAGAGGGAGTCAGGGATATCTGCCAGGAA**  
fig|1095771.3.peg.1573 -----aagaca-----aggaggatcacgc**aATGAGCGTATTTTTACACACCAGAGGGTGAAGGGATATTTACAATGGT**  
fig|411459.7.peg.1542 -----ggaggaaacgact**A-TGCAGCGAATTTTTACACACTGAAGGTGTTCCGGATATTTATAATGGA**  
\* \*

.....  
**fig|1410649.3.peg.2004** **Blautia schinkii DSM 10518**  
**fig|657314.3.peg.2472** **Ruminococcus obeum A2-162**  
**fig|40520.1561.peg.1222** **Blautia obeum strain af\_0058**

fig|1410649.3.peg.2004 ggaatccaattagagaggaaaagtgggattgacttatgtaatgcac **catggttagtatggttagcaca** aataaagcattggttgtaagtaagacacacgag 4.26  
fig|657314.3.peg.2472 ggattgacaatttggtttttcctattat-aatttccttggttt-at **tatagtagcatggttagcacg** ttaaagcatc--tgcctaca-----tgag  
fig|40520.1561.peg.1222 ggattgacattttctattttttctactat-aatttccttggtttta **catagtagcatggttagcacg** ttaaagtgtt---tatctgtaa-----tgag  
\* \*

fig|1410649.3.peg.2004 gaggaatagac**ATGCAGCGAATTTTTACACACCAGGGGTGTACGGGA**  
fig|657314.3.peg.2472 gaggaaccgatt**ATGCAGCGAATTTTTACACACCAGGGGTGTACGGGA**  
fig|40520.1561.peg.1222 gaggaacgact**ATGCAGCGAATTTTTACACACCAGGGGTGTACGGGA**  
\*\*\*\*\* \*\*\*\*\* \* \* \* \* \* \* \* \* \* \* \* \* \*

.....  
**fig|657324.3.peg.1663** **Butyrivibrio fibrisolvens 16/4**  
**fig|1123011.4.peg.475** **Pseudobutyrvibrio ruminis DSM 9787**

fig|657324.3.peg.1663 gtttttttacgccctgaggccattggtgacttcggggctttattaaagtataataacttttag **ttctttatcagactaaagtgt** ---ttttataagagg 4.20  
fig|1123011.4.peg.475 gaaatagagtgccccgaggccat-gttgacttcggggctttattaaagtataataacttttag **ttctttatcacgctaagtg** tagttttgtacaagagg 4.78  
\* \*

fig|657324.3.peg.1663 acaaaa**ATGAAAAATCTTACTTTACACACTCCAGAGGGAGTAAGAGATAT**  
fig|1123011.4.peg.475 gcgaa**A-TGAAAAATTTAACTTTACACACTCCAGAAGGAGTTCGAGACAT**  
\* \*



**fig|887325.3.peg.962**  
**fig|467210.3.peg.308**  
**fig|575593.3.peg.2148**

**Eubacterium saburreum DSM 3986**  
**Lachnoanaerobaculum saburreum strain DNF00896**  
**Lachnospiraceae oral taxon 107 str. F0167**

nRyTTTAnYRYRnTAAArYr

```
fig|887325.3.peg.962  gagtcaaagagtgctacaagtctgaataaagcagaaatactttaatgtagtagagaaaacaaagtgttgacaagaacacttttagtatgtctaccataagaaagt 4.27
fig|467210.3.peg.308  gaatcaaagggatctacaaatcagaacaaagcagaatttacttttagttagtagagaaaataaagtgttgacagatacacttttagtgtgctaccataagaaagt
fig|575593.3.peg.2148  ataatggctctttttattaataatgatgaa-tataaaaaactttattttagtagagaaaataaagtgttgacaaaactacttttagcatgtctaccataagaaagt
** *          ** ** * ** * ***** * ***** * ***** * ***** * ***** * ***** * ***** * *****
```

```
fig|887325.3.peg.962  gttgaggcaaagcaa-----agaaagtaatgaggtgtattATGAAAAAAAAATTTATTTGTTGGAGTATTGGCGGCACTTA
fig|467210.3.peg.308  gtttaagacaaagcaacaagcaaagaaagtaatgaggtgtattATGAAGAAAAATTTATTTGTTGGAGTATTGGCGGCACTTA
fig|575593.3.peg.2148  acaaaggaacattaaaaaagcaagcgaaaaaatgaggtgtattATGAAAAAAAAATGATTTATTGGGAATTTGGCAGCATTTG
** * * **          ** ***** * ***** * ***** * ***** * ***** * ***** * *****
```

**fig|1291051.3.peg.2211**  
**fig|658655.3.peg.174**  
**fig|290055.3.peg.1890**  
**fig|411470.6.peg.2154**

**[Clostridium] glycyrrhizinilyticum JCM 13369**  
**Lachnospiraceae bacterium 1\_4\_56FAA**  
**[Eubacterium] fissicatena strain KCTC 15010**  
**Ruminococcus gnavus ATCC 29149**

nRyTTTAnYRYRnTAAArYr

```
fig|1291051.3.peg.2211  agaaagatctggatgagcgggaagaactcgggattgacatttagtcggaaaattcgtataatg-actaaagattttaacgtagtagcagagtaaaagtgcg 4.01
fig|290055.3.peg.1890  agaaggatttggat-----accaggattgacattttctgagaaaactcgtataatttacagggtgatttaacgaattagcatgataaaagtgaa 4.70
fig|411470.6.peg.2154  aaaaagatttggat-----gacaggattgacagaagaaatgaaagttcgtataatt-acagaatgatttaatgtattagcatagtaaaagtgaa 4.79
fig|658655.3.peg.174  agaaagatctggatgagcgggaagaactcgggattgacatttagtcggaaaattcgtataatg-actaaagattttaacgtagtagcagagtaaaagtgcg
* ** ** * *****          ***** * ***** * ***** * ***** * *****
```

```
fig|1291051.3.peg.2211  gaggtgttgag---ATGGAACAAAAATTACATACGCCGGAGGGTGTAAGAGATATTTACAGTACA
fig|290055.3.peg.1890  agg---caggaaatATGGAACAGAAATTACATACACCTGAAGGCGTCAGGGATATATACAATACA
fig|411470.6.peg.2154  aggggtgttatgaattATGGAACAAAAGTTACATACTCCCGAAGGAGTTCGAGATATTTACAGTAAT
fig|658655.3.peg.174  gaggtgttgag---ATGGAACAAAAATTACATACGCCGGAGGGTGTAAGAGATATTTACAGTACA
*          *          ***** * ***** * ***** * ***** * *****
```

**fig|693746.9.peg.4538**  
**fig|1007096.3.peg.459**

**Oscillibacter valericigenes Sjm18-20**  
**Oscillibacter ruminantium GH1**

nRyTTTAnYRYRnTAAArYr

```
fig|693746.9.peg.4538  aaaggggaaatttaccocgtagaaaaaatgtactttactttgatactcggatgaagtattataaaatcattaaactgtccgacatggcggaacacgga 3.66
fig|1007096.3.peg.459  aaaagaggaaattccgcatataaaaaaatgtactttactttgataacaagatgaagtattgataaatccattaaactgtccgatataggcggataacgga 4.07
*** * * ** * * * ** * ***** * ***** * ***** * ***** * ***** * *****
fig|693746.9.peg.4538  gggacatATGGAGTTGGATCTGGGAATATTGAAACCGCAGGAGC
fig|1007096.3.peg.459  gggacatATGGAGCTGGATCTGGGGATATTGAAACCGCAGGAGC
***** * ***** * *****
```





fig|485916.5.peg.755 *Desulfotomaculum acetoxidans* DSM 771  
 fig|1121428.3.peg.2623 *Desulfotomaculum hydrothermale* Lam5 = DSM 18033  
 fig|696281.4.peg.1365 *Desulfotomaculum ruminis* DSM 2154  
 fig|868595.4.peg.1894 *Desulfotomaculum carboxydivorans* CO-1-SRB  
 fig|696369.3.peg.2190 *Desulfotomaculum nigrificans* DSM 574  
 fig|349161.6.peg.2584 *Desulfotomaculum reducens* MI-1  
 fig|635013.3.peg.889 *Thermincola* sp. JR

nRyTTTAnYRYRnTAAArYr

fig 485916.5.peg.755	aataagca-ttgacagccccggaagagttgatttattattatagcgtattaa	tgctttaatatagtaaagag	gtgaagtga	4.15
fig 1121428.3.peg.2623	aattagtagcttgacaccaccttgcta-tcaatttaacataatattatattag	taatttaacaagtaaagag	gttgt-tta	4.09
fig 868595.4.peg.1894	aaaaagtcttgacaccatcataatg-ctatntttaccatatacattatgtag	taatttaataaaactaaaggg	ggccgggtt	4.26
fig 696369.3.peg.2190	aaaaagtcttgacaccatcataatg-ctatntttaccatatacattatgtag	taatttaataaaactaaaggg	ggccgggtt	4.26
fig 696281.4.peg.1365	aaaaaactattgacagcagcaagatg-ctatntttacatattattatattac	taatttaattaactaaaggag	gggaccgcc	4.10
fig 349161.6.peg.2584	aaattttacttgacacctttactacatgcat-ttaacatatacattatattag	taatttaataagtaaagag	gtgatacc	4.05
fig 635013.3.peg.889	taccttgattgacatttat-caaaatccatgctataatnttcacatgtag	cgctttaacgcgctaaagcg	aaaggggtc	4.82

\*           \*\*\*\*\*                   \*\* \*\* \* \* \*\*\*           \*\*\*\*\*           \*\*\*\*\*

fig 485916.5.peg.755	tTTGGATTCTTCCAGG-GGGACTTTGACCATAGCCATACCTAAAGGGACTTTA
fig 1121428.3.peg.2623	tTTGGTGGCGGCCAGG-GAGCTGTTAACCATTTGCTCTGCCCAAAGGGACTCTC
fig 868595.4.peg.1894	tTTGGTTGCTGATAGA-GATATTTTTAACCGTGGCCTTGCCCAAAGGGCACACTT
fig 696369.3.peg.2190	tTTGGTTGCTGATAGA-GATATTTTTAACCGTGGCCTTGCCCAAAGGGCACACTT
fig 696281.4.peg.1365	TT-GGTGCGGCCAGA-GATATTTTTAACAGTGGCTTTGCCCAAAGGGACCCTT
fig 349161.6.peg.2584	GT-GGTGCGGCCAGA-GAACAATTAACACTATTGCTTTGCCCAAAGGGAACCCTA
fig 635013.3.peg.889	ct-ggtgttttATGGATGATTATTTATCTATAGCTATTTCCCAAAGGGACACTA

\* \*\*\*                   \* \*                   \*\* \* \* \* \* \* \* \* \* \* \*

fig|457570.14.peg.2058 *Natranaerobius thermophilus* JW/NM-WN-LF  
 fig|555088.4.peg.1367 *Dethiobacter alkaliphilus* AHT 1  
 fig|574087.3.peg.1493 *Acetohalobium arabaticum* DSM 5501

nRyTTTAnYRYRnTAAArYr

fig 457570.14.peg.2058	tacagagaccattttattgacggaataaaaatagcgtgataagatattaccaat--gattaattaa	cactttaacgagttaacgaa	ataaaggaggttc	3.88
fig 574087.3.peg.1493	tgtaggatt--ttctatatttatattgacattttaaataagtagtatattt-taataagttaa	tactttactgtattaaaaata	ataaag--aggtg	5.08
fig 555088.4.peg.1367	ggtaaagggtt--gt-taactttgtactaaaattattgacacggcgcctgtgcatgttaaatttaatacgttat	cactttaacgtattaaagcg	gggta	5.02

\*           \* \*\*                   \* \* \* \* \*                   \*\*           \* \* \* \* \* \* \* \* \*                   \*\*\*\* \* \*\*\*\*\* \* \*

fig 457570.14.peg.2058	cttaaTTGGTCCCAAGAGAGATTTCAGCTGTTACAAAACCATCAGGAGTCAACGAT---
fig 574087.3.peg.1493	gtagaATGGA---AAGTGAATTGAATTTACTAAGATGCAGACCCCGATGGTACTAAG
fig 555088.4.peg.1367	tctacATGAA---TGATAAATGAAACGACTGCTGACGCCGAAGGTGTCCGCGATTG

\*\*                   \* \* \* \*                   \* \*                   \* \*                   \*

fig|589861.3.peg.975 **Thermoanaerobacter ethanolicus** CCSD1  
fig|1125975.3.peg.2106 **Thermoanaerobacter indiensis** BSB-33  
fig|880478.3.peg.1806 **Thermoanaerobacter siderophilus** SR4  
fig|399726.7.peg.1046 **Thermoanaerobacter** sp. X514

fig|589861.3.peg.975 gtctatgcttgacatacattgaaaattaggatacaataaaagaagtacagttttat **tactttatcacgttacaaag**acagagtaaattgtgggggtgtttta 4.11  
fig|1125975.3.peg.2106 gtctatgcttgacatacattgaaaattaggatacaataaaagaagtacagttttat **tactttatcacgttacaaag**acagagtaaattgtgggggtgtttta 4.11  
fig|880478.3.peg.1806 gtctatgcttgacatacattgaaaattaggatacaataaaagaagtacagttttat **tactttatcacgttacaaag**acagagtaaattgtgggggtgtttta 4.33  
fig|399726.7.peg.1046 gtctgtgcttgacatacattgaaaattaggatacaataaaagaataacagttttat **tactttatcacgttacaaag**acagagtaaattgtgggggtgtttta  
\*\*\*\* \*\*\*\*\*

fig|589861.3.peg.975 **ATGATATATTTGCCTGATGGTGTACAAGATTTTTTGCCGGAAGAATATAG**  
fig|880478.3.peg.1806 **ATGATATATTTGCCTGATGGTGTACAAGATTTTTTGCCGGAAGAATATAG**  
fig|1125975.3.peg.2106 **ATGATATATTTGCCTGATGGTGTACAAGATTTTTTGCCGGAAGAATATAG**  
fig|399726.7.peg.1046 **ATGATATATTTGCCTGATGGTGTACAAGATTTTTTGCCAGAGGAATATAA**  
\*\*\*\*\* \*\* \*\*\*\*\*

fig|1421016.3.peg.2602 **Thermoanaerobacterium aotearoense** SCUT27  
fig|1094508.3.peg.1187 **Thermoanaerobacterium saccharolyticum** JW/SL-YS485  
fig|858215.3.peg.1086 **Thermoanaerobacterium xylanolyticum** LX-11  
fig|698948.3.peg.535 **Thermoanaerobacterium thermosaccharolyticum** M0795  
fig|580331.4.peg.1866 **Thermoanaerobacter italicus** Ab9

fig|1421016.3.peg.2602 gaaaatgtacttgacacatagacaaaaataaattattattttaaataa--caagttattactttat **cactttacaaagataaaagta**a----agtaatagaa 4.08  
fig|1094508.3.peg.1187 gaaaatgtacttgacacatagacaaaaataaattattattttaaataa--caagttattactttat **cactttacaaagataaaagta**a----agtaatagaa 4.08  
fig|858215.3.peg.1086 gaaaatgtacttgacacatagacaaaaataaattattattttaaataa--caagtcattactttat **cactttacaaataaaagta**a----aatagtataa 5.02  
fig|698948.3.peg.535 attaaagtacttgacacatggacaaaaataaattattatgtaaataa--caagtcattactttat **cactctacaaagataaaagta**atntagatggtagc  
fig|580331.4.peg.1866 ttgtctgtgcttgacatacattgaaaattaggatacaataaaagaataacagttttattactttat **cacggtacaaagacagagta**a-----attgt----  
\*\* \*\*\*\*\* \* \*\*\*\* \* \*\* \* \* \* \* \* \*\*\*\*\* \* \* \* \* \* \* \*

fig|1421016.3.peg.2602 gaggtgctttg**ATGAAGAATCTACCTGATGGAGTACAGGATTTTTGCCTGATGAGTTGAGG**  
fig|1094508.3.peg.1187 gaggtgctttg**ATGAAGAATCTACCTGATGGAGTACAGGATTTTTGCCTGATGAGTTGAGG**  
fig|858215.3.peg.1086 gaggtgctttg**ATGAAGAATTTACCTGACGGAGTACAGGATTTTTGCCTGATGAGTTGAGA**  
fig|698948.3.peg.535 gaggtgctttg**ATGAAAAATTTACCTGATGGAGTACAGGATTTTTGCCTGAGGAATTAAGA**  
fig|580331.4.peg.1866 ggggtgtttta**ATGATATATTTGCCTGATGGTGTACAAGATTTTTGCCAGAGGAATATAAG**  
\* \*\*\*\* \*

**fig|520767.4.peg.2208**      **Thermovenabulum gondwanense strain R270**  
**fig|632348.3.peg.1306**      **Caldicellulosiruptor kronotskyensis 2002**  
**fig|521460.8.peg.1549**      **Anaerocellum thermophilum DSM 6725**  
**fig|31899.10.peg.1973**      **Caldicellulosiruptor bescii strain DSMZ 6725**

nRyTTTAnYRYRnTAAArYr

```

fig|520767.4.peg.2208  tttcatatTTTTTacttgccaaaattccacatataatttataatattatcaaattaactatTTTAAATATTAAGTAataaatttacataatggcggggt 4.06
fig|521460.8.peg.1549  ggctccactTTTTTgTtGctTTTTtagTtctTTTTgTaatataatataa-catgatatctattcatcacttcattatgataaa----gtaatgaaaggg- 4.69
fig|31899.10.peg.1973  ggctccactTTTTTgTtGctTTTTtagTtctTTTTgTaatataatataa-catgatatctattcatcacttcattatgataaa----gtaatgaaaggg- 4.69
fig|632348.3.peg.1306  ggctccactTTTTTgTtGctTTTTtagTtctTTTTgTaatataatataa-catgatatctattcatcacttcattatgataaa----gtaatgaaaggg- 4.69
                        *  *****  ***      *  *  *  *  *****  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *
fig|520767.4.peg.2208  cgcataaaATGGGTAAAAAATTCACGGATTACTGCCCGCCGGGCAGTATT
fig|521460.8.peg.1549  -agaaaaATGAGCAAAAAGTTTGTGGATTACTATCCTCCAGTGTCTGAT
fig|31899.10.peg.1973  -agaaaaATGAGCAAAAAGTTTGTGGATTACTATCCTCCAGTGTCTGAT
fig|632348.3.peg.1306  -agaaaaATGAGCAAAAAGTTTGTGGATTACTATCCTCCAGTGTCTGAT
                        *  *****  *  *****  **  *****  *  *  *  *  *  *  *  *  *  *  *

```

**fig|525903.6.peg.232**      **Thermanaerovibrio acidaminovorans DSM 6589**  
**fig|584708.3.peg.2236**      **Aminomonas paucivorans DSM 12260**  
**fig|926567.3.peg.1615**      **Thermanaerovibrio velox DSM 12556**

```

fig|525903.6.peg.232  aacacctccgTtgattttccgCGGCCCCatgatatcatgcaaaccacgTttcggctaccaattcatcagctaacaaggc-actggaggta----- 4.31
fig|584708.3.peg.2236  --cacctccgTtgctgtcaaacgTttggatgatatcatgCGGaccatgtcaaggatgaccctttatcccgatagcatgatgaacggaggggaccgcc 3.93
fig|926567.3.peg.1615  aacacctccgTtgcaTttccgCGGacacatgataacatgggatccacgTttgggatgataatttatcagctaatatgac-actggaggtaa----- 5.31
                        *****  *  ***  *****  *****  ***  *  **  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *
fig|525903.6.peg.232  -tgaccatgtca--cctaaccgccaactcaggtgggtcccc---cgggag-----gaggtgccccgtaATGC--GGGTCCCC
fig|584708.3.peg.2236  atgaccgtcgcac-cgcagcgccccgcctccccgtctcccgcgcgaaagcgcggaaccggaggtggggggccATGCTCCGAACCCC
fig|926567.3.peg.1615  -cagccatgagaaacgCGcacccaaatcgccgggacccccgtttgaagcg-----gaggtgcccggccATG---AGACCCCC
                        **  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

```

**Legend and Abbreviations:**

To validate candidate HisR-binding sites with scores below threshold (weak sites) we applied a phylogenetic footprinting approach to upstream regions of these genes using multiple sequence alignment of upstream gene regions for each group of closely-related orthologs as previously described in [1, 2]. Multiple alignments of DNA upstream regions were obtained by ClustalW2 integrated in the SEED platform. Genes/genomes with weak candidate HisR binding sites are highlighted in red in the header line, while other genes/genomes in black possess high-scored HisR sites. Predicted HisR-binding sites are highlighted by yellow boxes with a taxon-specific consensus shown on top. Gene coding regions are shown in capital red letters.

[1] Khoroshkin MS, Leyn SA, Van Sinderen D, Rodionov DA. 2016. Transcriptional regulation of carbohydrate utilization pathways in the *Bifidobacterium* genus. Front Microbiol 7:120.

[2] Ravcheev DA, Godzik A, Osterman AL, Rodionov DA. 2013. Polysaccharides utilization in human gut bacterium *Bacteroides thetaiotaomicron*: comparative genomics reconstruction of metabolic and regulatory networks. BMC Genomics 14:873.