

**Additional File 6. Phylogenetic footprinting of upstream regions of genes controlled by candidate Histidine-dependent translational attenuators (see legend below).**

**A. Novel candidate His attenuator upstream of the *hisZGDBHAFIEJ* operon in the *Bacillus cereus* group (Bacillales).**

**222523.1.peg.1517** *Bacillus cereus* ATCC 10987  
**527024.3.peg.2634** *Bacillus thuringiensis* serovar *tochigiensis* BGSC 4Y1  
**1245029.3.peg.3591** *Bacillus anthracis* str. *Carbosap*  
**315730.11.peg.1914** *Bacillus weihenstephanensis* KBAB4  
**526997.3.peg.1686** *Bacillus mycoides* DSM 2048  
**315749.8.peg.1198** *Bacillus cytotoxicus* NVH 391-98

```

-35 box Promoter  -10 box           Leader RBS       Leader start           His codons  Stop
315730.11.peg.1914 tatttacattcagaattaactgttatataaatatgttaaaaaatcagaggagaaaaggaggtaagtattATGAATATCGTGCAATCTATTAAATGCTTACAAAACTACATCATACATATGA
526997.3.peg.1686 tatttacattcggaattaactgttatataaatatgttaaaaaatcagaggagaaaaggaggtaagtattATGAATATCGTGCAATCTATTAAATGCTTACAAAATTACATCATACATATGA
527024.3.peg.2634 tatttacattcagaattaagtgttatataactatgttaaaaa-ttagaggagaaaaggaggtaagtattATGAAAACAATGCAATCTATTAAATGCTTACACAACTACATCATACATATAA
1245029.3.peg.3591 tatttacattcagaattatgtgttatataactatgttaaaaa-ttagaggagaaaaggaggtaagtattATGAAAACAATGCAATCTATTAAATGCTTACACAACTACATCATACAGTAA
222523.1.peg.1517 tatttacattcagaattaagtgttatataactatgttaaaaa-ttagaggagaaaaggaggtaagtattATGAAAACAATGCAATCTATTAAATGCTTACACAACTACATCATACATATAA
315749.8.peg.1198 tatttacttttctgaatagttgttatataaatgtgtgaaaa-taagcggagaaaaggaggtaaggattATGTTTGCAATGAAGTGATAAATATGCTTACAAAACTACACCATCATCCATAA
***** ** * ***** ** * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

Attenuator hairpin           terminator hairpin
315730.11.peg.1914 aacccttgaacctaacgtgaaa-acgtgtagggacaagggttattcccgtgttacctacactcactactaaa-gagccttgtaggta--ttttgtctacaaggctctttt
526997.3.peg.1686 aacccttgaacctaacgtgaaa-acgtgtagggacaagggttattcccgtgttacctacactcactactaaa-gagccttgtaggta--ttttgtctacaaggctctttt
527024.3.peg.2634 aacccttgaacctagcgtgaaa-acgtgtaggtgcaagggttattcccgtgttacctactcgctactaaa-gagccttgtaggta--ttttgtctacaaggctctttt
1245029.3.peg.3591 agcccttgaacctaacgtgaaa-acgtgtaggtgcaagggttattcccgtgttacctacacgcctacgaaa-gagccttgtaggta--ttttgtctacaaggctctttt
222523.1.peg.1517 agcccttgaacctagcgtgaaa-acgtgtaggtgcaagggttattcccgtgttacctacacgcctactaaa-gagccttgtaggta--ttttgtctacaaggctctttt
315749.8.peg.1198 aacccttgctcctagcgtgaaaacgtgtaggtacaagggttattcctgtgttacctacactctataaaaa-gagccttgtaggta--ttttgtctadaaggctcttt
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

hisZ gene start
315730.11.peg.1914 tttttcggtttgggtggatggaatgg-cataaattaaagaggggaagggtgtgtgaattcagATGACAAAGTGGAAACGGGCAAATCCGAATGGAACGAG
526997.3.peg.1686 tttttcggtttgggtggatggaatgggcataaattaaagaggggaagggtgtgtgaattcagATGACAAAGTGGAAACGGGCAAATCCGAATGGAACGAG
527024.3.peg.2634 tttttcggtttaggtgcatgaaataagcatacattagaaagggaaaggtgtgagaatttagATGACAAAATGGAAACGGGCAAACCCAAATGGAACGAG
1245029.3.peg.3591 tttttcggtttaggtgcatgaaataagcatacattagaaagggaaaggtgtgagaatttagATGACAAAATGGAAACGGGCAAACCCAAATGGAACGAG
222523.1.peg.1517 tttttcggtttaggtgcatgaaataagcatacattagaaagggaaaggtgtgagaatttagATGACAAAATGGAAACGGGCAAATCCAAATGGAACGAG
315749.8.peg.1198 ttctttgtttgtgattgtttaggtgagcatatacataaaaaagaaaggg---ggagttagcATGACGAACTGGAGACGTACAAAACCAAATGGAACGAG
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```





**C. Novel candidate His attenuator upstream of the *hisZGDCBHAF(IE)* operon in the Erysipelotrichales group.**

- 451640.5.peg.1377 **Catenibacterium mitsuokai** DSM 15897
- 1297794.4.peg.3389 **Coprobacillus cateniformis** JCM 10604
- 999415.3.peg.1341 **Eggerthia catenaformis** OT 569 = DSM 20559
- 445974.6.peg.1707 **Erysipelatoclostridium ramosum** DSM 1402
- 69824.3.peg.1410 **[Clostridium] cocleatum** strain DSM 1551
- 428126.7.peg.165 **[Clostridium] spiroforme** DSM 1552

	leader RBS	leader start	His codons	Stop	attenuator hairpin	
445974.6.peg.1707	tacaatgaga-ggtg	ttaattATG	AAAAACAACA	CAAAATTTACAATTA-	CATCACCATCAT	TGGATT--GCCTAG
69824.3.peg.1410	tacaatgaga-ggtg	tctaattATG	AAAAACAATCACA	AATTTACAACTA-	CATCACCATCAT	TGGATT--GCCTAG
428126.7.peg.165	tacaatgaga-ggtg	gataaattATG	AAAAACAATAATTA	AATTTACAATTA-	CATCACCATCAT	TGGATT--GCCTAG
1297794.4.peg.3389	aaaatgaggtga	tatttATG	AAAAACAATTTTAT	TATTTACAATTT-	CATCACCATCAT	TGGAAC--AACTAG
999415.3.peg.1341	aaaaatgaggtga	tatttATG	CAAACTATACAGTTA-	AACAATTT-	CATCATCACCAT	TGGCGT--ATTTAG
451640.5.peg.1377	aaaactgaggtga	tatttATG	AAAAATACAATTTT	TGTCAACAATTTAC	CACCATCACCAT	CGTGGTCTAGTTAG
	* * **** * *	**	* * * * * * * * * *	***** ****	***** **	
			Terminator hairpin			<i>hisZ</i> gene start
445974.6.peg.1707	atthttg-tg	ttaaatattatt	ttgggtgtatct	aatgac---aat	attacttttag-	cgagtcattagt
69824.3.peg.1410	atthttg-tatt	aaatagtttgg	gtgtatctaat	gac---aat	attacttttag-	cgagtcattagt
428126.7.peg.165	atthttg-tgt	ataaatttgg	ttgtatctaat	gac---aat	ttgttataaac	gagtcattagta-
1297794.4.peg.3389	atthttg-tgt	atataaattt	gggtgtatctg	atgaggtttat	atgg-cgagtc	attaa---gact
999415.3.peg.1341	gttttgtttt	acaattaag	caaggtgtca--	atctaat---	aacactaga	aacgaggttatt
451640.5.peg.1377	gttttgacata	caactcaaaaa	acggacacgaa	acaat---gac	ataagggaa	caggtcattg---
	**** * **	*	*	** **	***	* *

**Legend and Abbreviations:**

To identify candidate His attenuators we applied a phylogenetic footprinting approach to upstream regions of *his* genes using multiple sequence alignment of upstream gene regions for each group of closely-related orthologs. Multiple alignments of DNA upstream regions were obtained by ClustalW2 integrated in the SEED platform. Legend for attenuator elements:

LEADER PEPTIDE    His codons    Stop codon    attenuator hairpin    terminator hairpin    Potential antiterminator    Coding region of downstream gene