Set 1:

This set yielded the only motif conserved across all the upstream
sequences. The motif identified is 'atggtg'. This was the best-
conserved motif across all genes, being perfectly conserved
across all 9 sequences. Increasing motif size caused number of
genes sharing motifs to drop sharply (down to no greater than 4
genes sharing any 1 motif) even when increasing the maximum
parsimony score allowed and decreasing motif loss cost.

To compare, the motif found by DIALIGN was 'AAATTA' (occurs in 6
of 9 sequences). Footprinter did not identify this but this could
be because of choice of parameters to limit the number of motifs
found of length 6.

Inputs:

Protein	PID	Orgnanism
1.isoleucyl-tRNA synthetase	15673845	Lactococcus lactis subsp. lactis
2.isoleucyl-tRNA synthetase	15901494	Streptococcus pneumoniae TIGR4
3.isoleucyl-tRNA synthetase	15903545	Streptococcus pneumoniae R6
4.isoleucyl-tRNA synthetase	25010601	Streptococcus agalactiae NEM316
5.isoleucyl-tRNA synthetase	22536664	Streptococcus agalactiae 2603V/R
6.isoleucyl-tRNA synthetase	16804058	Listeria monocytogenes
7.isoleucyl-tRNA synthetase	16801193	Listeria innocua
8.isoleucyl-tRNA synthetase	15615108	Bacillus halodurans
9.isoleucyl-tRNA synthetase	16078607	Bacillus subtilis

Footprinter results:
The results URL is:
http://abstract.cs.washington.edu/~blanchem/FootPrinterWeb/__webq
uery__.fasta0.9273578379189370.802847297381945.main.html

The motif match summary:
Significance score: 1.768750
Parsimony score: 0.500000
Span: 32.849159
LISTERIA-MONOCY   -211   atggtg
LISTERIA-INNOCU   -211   atggtg
LACTOCOCCUS-LAC   -225   atggtg
STREPTO-AGAL.NE   -245   atggtg
STREPTO-AGAL.26   -245   atggtg
STREPTO-PNEUM.T   -51   atggtg
STREPTO-PNEUM.R   -78   atggtg
BACILLUS-HALODU   -177   atggtg
BACILLUS-SUBTIL   -222   atggtg

Footprinter Parameters used:
Motif size: 6
Maximum number of mutations (max. parsimony score): 1
Max. Number of mutations per branch: 1
Sub region size: 50
Sub region change cost: 1
Allow regulatory element loss: yes

Spanned tree significance level: significant
Motif loss cost: 1

Set 2:

There was no motif found conserved across all 7 upstream
sequences. This was after varying motif size through the allowed
range (6-12), allowing a maximum parsimony score of 4, maximum
mutations per branch to 2 and increasing motif loss cost to the
maximum allowed (2). The motif 'taaaat' was the best result found
by Footprinter. It was conserved across 5 gene upstream
sequences. It may not be a regulatory sequence as it looks like a
random repeating pattern. Since the sequence sizes varied from 67
to 300, sub region size was not used (i.e., it was set to 500,
with change cost set to 0).

To compare, the motif found by DIALIGN was 'ATTATA' (occurring in
4 of 7 sequences). Footprinter did not identify this. In this
case the motif identified by DIALIGN is more likely to be a
regulatory element.

Inputs:

Protein	PID	Orgnanism
1.arginyl-tRNA synthetase	15674020	Lactococcus lactis subsp. lactis
2.Arginyl-tRNA synthetase (arginine--tRNA ligase)	15903931	Streptococcus pneumoniae R6
3.arginyl-tRNA synthetase	15901895	Streptococcus pneumoniae TIGR4
4.arginyl-tRNA synthetase	25012094	Streptococcus agalactiae NEM316
5.arginine-tRNA ligase	18310643	Clostridium perfringens
6.Arginyl-tRNA synthetase	15894328	Clostridium acetobutylicum
7.arginine-tRNA ligase	15613397	Bacillus halodurans

Footprinter results:
The results URL is:
http://abstract.cs.washington.edu/~blanchem/FootPrinterWeb/__webq
uery__.fasta0.3271038696521660.158268579965792.main.html

The motif match summary:
Significance score: 0.137500
Parsimony score: 0.000000
Span: 18.695169
STREPTOCOCCUS-P   -46   taaaat
STREPTOCOCCUS-P   -35   taaaat
STREPTOCOCCUS-P   -46   taaaat
STREPTOCOCCUS-P   -35   taaaat
STREPTOCOCCUS-A   -44   taaaat
STREPTOCOCCUS-A   -18   taaaat
CLOSTRIDIUM-ACE   -122   taaaat
BACILLUS-HALODU   -282   taaaat

Footprinter Parameters:
Motif size: 6
Maximum number of mutations (max. parsimony score): 1
Max. Number of mutations per branch: 1
Sub region size: 500
Sub region change cost: 0
Allow regulatory element loss: yes
Spanned tree significance level: significant
Motif loss cost: 2

Set 3:

There was no motif found that was conserved across all 8 upstream
sequences. After varying motif length from 6 to 12 and
experimenting with different values for the parameters (Parsimony
score, maximum mutations per branch, sub region size, sub region
cost, motif loss cost), the motif 'gactgcaagc' was the best. It
is conserved across 6 of the 8 sequences. This motif could be a
regulatory element (the motif does not look like it a randomly
occurring series of nucleotides).

To compare, the motifs identified by DIALIGN were AGCCGTGACT
(occurring in 6 of 8 sequences) and AATAGG (occurring in 4 of 8
sequences). Footprinter identified neither. Both the Motif
identified by Footprinter and the first motif identified by
DIALIGN could be regulatory elements.

Inputs:

Protein	PID	Orgnanism
1.phenylalanil-tRNA synthetase alpha chain	15673911	Lactococcus lactis subsp. lactis
2.Phenylalanyl-tRNA synthetase alpha chain	15902551	Streptococcus pneumoniae R6
3.Phenylalanyl-tRNA synthetase	15900489	Streptococcus pneumoniae TIGR4
4.phenylalanyl-tRNA synthetase,	22537032	Streptococcus agalactiae 2603V/R
5.Phenylalanyl-tRNA synthetase alpha chain	25010941	Streptococcus agalactiae NEM316
6.phenylalany-tRNA synthetase alpha subunit	16800253	Listeria innocua
7.phenylalanyl-tRNA synthetase alpha subunit	23099586	Oceanobacillus iheyensis
8.phenylalany-tRNA synthetase alpha subunit	16803261	Listeria monocytogenes

Footprinter results:
The results URL is:
http://abstract.cs.washington.edu/~blanchem/FootPrinterWeb/__webq
uery__.fasta0.784122412919789 0.938679405473508.main.html

The motif match summary:
Significance score: 0.037500
Parsimony score: 1.000000
Span: 14.189537
STREPTOCOCCUS-A   -178   gactgcaagc
STREPTOCOCCUS-A   -179   gactgcaagc
STREPTOCOCCUS-P   -168   gactgcaagc
STREPTOCOCCUS-P   -87   gactgcaagc
LISTERIA-MONOCY   -215   gactgaaagc

```
LISTERIA-INNOCU     -206  gactgaaagc

Footprinter Parameters:
Motif size: 10
Maximum number of mutations (max. parsimony score): 1
Max. Number of mutations per branch: 1
Sub region size: 200
Sub region change cost: 1
Allow regulatory element loss: yes
Spanned tree significance level: significant
Motif loss cost: 2
```