

(Circular) MAP of: Plh29.Seq check: 6918 from: 1 to: 11675

With 212 enzymes: *

MaxCuts: 5

October 22, 1999 16:00 ..

M
u
n
I

ATGAAGATTCTTGCTCAATTGTTATCAGCTATGCGCCGACCAGAACACCTTGCCGATCAG
1 -----+-----+-----+-----+-----+-----+ 60
TACTTCTAAGAACGAGTTAACAATAGTCGATACGCGGCTGGTCTTGTGGAACGGCTAGTC

E
a
r
I

R
l
e
A
I

CCAAACGTCTCTTCAGGCCACTGACTAGCCATAACTTTCCCACAAACGGAACAACCTCTCA
61 -----+-----+-----+-----+-----+-----+ 120
GGTTTGCAGAGAAGTCCGGTACTGATCGGTATTGAAAGGGGTGTTGCCTTGTTGAGAGT

R
l
e
A
I

TTGCATGGGATCATTGGGTACTGTGGGTTTAGTGGTTGTAAAAACACCTGACCGCTATCC
121 -----+-----+-----+-----+-----+-----+ 180
AACGTACCCTAGTAACCCATGACACCCAAATCACCAACATTTTTGTGGACTGGCGATAGG

CTGATCAGTTTCTTGAAGGTAAACTCATCCCCCAAGTCTGGCTATGCAGAAATCACCT
181 -----+-----+-----+-----+-----+-----+ 240
GACTAGTCAAAGAACTTCCATTTGAGTAGTGGGGTTTCAGACCGATACGTCTTTAGTGGA

H
i
n
d
I
I
I

GGCTCAACAGCCTGCTCAGGGTCAACGAGAATTAACATTCCGTCAGGAAAGCTTGGCTTG
241 -----+-----+-----+-----+-----+-----+ 300
CCGAGTTGTCGGACGAGTCCAGTTGCTCTTAATTGTAAGGCAGTCCTTTTGAACCGAAC

GAGCCTGTTGGTGCGGTCATGGAATTACCTTCAACCTCAAGCCAGAATGCAGAATCACTG
301 -----+-----+-----+-----+-----+-----+ 360
CTCGGACAACCACGCCAGTACCTTAATGGAAGTTGGAGTTCGGTCTTACGTCTTAGTGAC

H
i
n
d
I
I
I

361 GCTTTTTTGGTTGTGCTTACCCATCTCTCCGCATCACCTTTGGTAAAGGTTCTAAGCTTA 420
-----+-----+-----+-----+-----+-----+-----+
CGAAAAAACCAACACGAATGGGTAGAGAGGCGTAGTGAAACCATTTCCAAGATTCGAAT

421 GGTGAGAACATCCCTGCCTGAACATGAGAAAAACAGGGTACTCATACTCACTTCTAAGT 480
-----+-----+-----+-----+-----+-----+-----+
CCACTCTTGTAGGGACGGACTTGTACTCTTTTTTGTCCCATGAGTATGAGTGAAGATTCA

P
f
l
l
l
0
8
I

481 GACGGCTGCATACTAACCGCTTCATACATCTCGTAGATTTCTGTGGCGATTGAAGGGCTA 540
-----+-----+-----+-----+-----+-----+-----+
CTGCCGACGTATGATTGGCGAAGTATGTAGAGCATCTAAAGACACCGCTAACTTCCCGAT

A A
p p
a a
B B
I I

541 AATTCTTCAACGCTAACTTTGAGAATTTTTGCAAGCAATGCGGCGTTATAAGCATTTAAT 600
-----+-----+-----+-----+-----+-----+-----+
TTAAGAAGTTGCGATTGAAACTCTTAAAAACGTTTCGTTACGCCGCAATATTCGTAAATTA

N
s
i
I

601 GCATTGATGCCATTAATAAAGCACCAACGCCTGACTGCCCCATCCCCATCTTGTCTGCG 660
-----+-----+-----+-----+-----+-----+-----+
CGTAACTACGGTAATTTATTTTCGTGGTTGCGGACTGACGGGGTAGGGGTAGAACAGACGC

661 ACAGATTCCTGGGATAAGCCAAGTTCATTTTTCTTTTTTTCATAAATTGCTTTAAGGCGA 720
-----+-----+-----+-----+-----+-----+-----+
TGTCTAAGGACCCTATTCGGTTCAAGTAAAAAGAAAAAAGTATTTAACGAAATTCCGCT

P
s
h
A
I

721 CGTGCCTCCTCAAGCTGCTCTTGTGTTAATGGTTTCTTTTTTGTGCTCATACGTTAAATC 780
-----+-----+-----+-----+-----+-----+-----+
-----+-----+-----+-----+-----+-----+-----+

GCACGCAGGAGTTCGACGAGAACACAATTACCAAAGAAAAACACGAGTATGCAATTTAG

B
s
a
B
I
B
s
b
I
H
g
i
E
I
I

781 TATCACCGCAAGGGATAAATATCTAACACCGTGCGTGTTGACTATTTTACCTCTGGCGGT
-----+-----+-----+-----+-----+-----+ 840
ATAGTGGCGTTCCCTATTTATAGATTGTGGCACGCACAACCTGATAAAATGGAGACCGCCA

NT
sa
pt
II

841 GATAATGGTTGCATGTACTAAGGAGGTTGTATGCCACAATTTGATATATTATGTAAAACA
-----+-----+-----+-----+-----+-----+ 900
CTATTACCAACGTACATGATTCCTCCAACATACGGTGTTAAACTATATAATACATTTTGT

BH
sg
ui
3E
6I
II
/

901 CCACCTAAGGTGCTTGTTCGTCAGTTTGTGGAAAGGTTTGAAAGACCTTCAGGTGAGAAA
-----+-----+-----+-----+-----+-----+ 960
GGTGGATTCCACGAACAAGCAGTCAAACACCTTTCCAAACTTTCTGGAAGTCCACTCTTT

A
p
a
B
I

961 ATAGCATTATGTGCTGCTGAACTAACCTATTTATGTTGGATGATTACACATAACGGAACA
-----+-----+-----+-----+-----+-----+ 1020
TATCGTAATACACGACGACTTGATTGGATAAATACAACCTACTAATGTGTATTGCCTTGT

R
c
a
I

1021 GCAATCAAGAGAGCCACATTCATGAGCTATAATACTATCATAAGCAATTCGCTGAGTTTC
-----+-----+-----+-----+-----+-----+ 1080
CGTTAGTTCTCTCGGTGTAAGTACTCGATATTATGATAGTATTTCGTTAAGCGACTCAAAG

1081 GATATTGTCAATAAATCACTCCAGTTTAAATACAAGACGCAAAAAGCAACAATTCTGGAA
-----+-----+-----+-----+-----+-----+ 1140
CTATAACAGTTATTTAGTGAGGTCAAATTTATGTTCTGCGTTTTTTCGTTGTTAAGACCTT

1141 GCCTCATTAAAGAAATTGATTCCTGCTTGGGAATTTACAATTATTCCTTACTATGGACAA
-----+-----+-----+-----+-----+-----+ 1200
CGGAGTAATTTCTTTAACTAAGGACGAACCCTTAAATGTTAATAAGGAATGATACCTGTT

AAACATCAATCTGATGGAAGATCCCCCTGCCCGGTTATTATTATTTTTGACACCAGACC
1201 -----+-----+-----+-----+-----+-----+ 1260
TTTGTAGTTAGACTACCTTCTAGGGGGGACGGGCCAATAATAATAAAAACTGTGGTCTGG

B
P
u
1 E B
1 c s B
0 o p a
2 R 4 I
I I I I

AACTGGTAATGGTAGCGACCGGCGCTCAGCTGGAATTCGCCGATACTGACGGGCTCCAG
1261 -----+-----+-----+-----+-----+-----+ 1320
TTGACCATTACCATCGCTGGCCGCGAGTCGACCTTAAGGCGGCTATGACTGCCCGAGGTC

N
d
e
I

GAGTCGTCGCCACCAATCCCCATATGGAAACCGTCGATATTCAGCCATGTGCCTTCTTCC
1321 -----+-----+-----+-----+-----+-----+ 1380
CTCAGCAGCGGTGGTTAGGGGTATACCTTTGGCAGCTATAAGTCGGTACACGGAAGAAGG

S
f
c
I

GCGTGCAGCAGATGGCGATGGCTGGTTTTCCATCAGTTGCTGTTGACTGTAGCGGCTGATG
1381 -----+-----+-----+-----+-----+-----+ 1440
CGCACGTCGTCTACCGCTACCGACCAAAGGTAGTCAACGACAACCTGACATCGCCGACTAC

R B
l s
e t
A X
I I

TTGAACTGGAAGTCGCCGCGCCACTGGTGTGGGCCATAATTCAATTCGCGCGTCCCGCAG
1441 -----+-----+-----+-----+-----+-----+ 1500
AACTTGACCTTCAGCGGCGCGGTGACCACACCCGGTATTAAGTTAAGCGCGCAGGGCGTC

B B
s s
t p
1 L
1 U N
0 1 s
7 1 p
I I I

CGCAGACCGTTTTTCGCTCGGGAAGACGTACGGGGTATACATGTCTGACAATGGCAGATCC
1501 -----+-----+-----+-----+-----+-----+ 1560
GCGTCTGGCAAAAGCGAGCCCTTCTGCATGCCCCATATGTACAGACTGTTACCGTCTAGG

CAGCGGTCAAACAGGCGGCAGTAAGGCGGTCTGGGATAGTTTTCTTGGCGCCCTAATCCG

1561 -----+-----+-----+-----+-----+-----+ 1620
GTCGCCAGTTTTGTCCGCCGTCATTCCGCCAGCCCTATCAAAAGAACGCCGGGATTAGGC

B
s
p
M
I

AGCCAGTTTACCCGCTCTGCTACCTGCGCCAGCTGGCAGTTCAGGCCAATCCGCGCCGGA
1621 -----+-----+-----+-----+-----+-----+ 1680
TCGGTCAAATGGGCGAGACGATGGACGCGGTTCGACCGTCAAGTCCGGTTAGGCGCGGCCT

TGCGGTGTATCGCTCGCCACTTCAACATCAACGGTAATCGCCATTTGACCACTACCATCA
1681 -----+-----+-----+-----+-----+-----+ 1740
ACGCCACATAGCGAGCGGTGAAGTTGTAGTTGCCATTAGCGGTAAACTGGTGATGGTAGT

M
l
u
I

ATCCGGTAGGTTTTCCGGCTGATAAATAAGGTTTTCCCCTGATGCTGCCACGCGTGAGCG
1741 -----+-----+-----+-----+-----+-----+ 1800
TAGGCCATCCAAAAGGCCGACTATTTATTCCAAAAGGGGACTACGACGGTGCGCACTCGC

A
p
a
L
I

GTCGTAATCAGCACCGCATCAGCAAGTGTATCTGCCGTGCACTGCAACAACGCTGCTTCG
1801 -----+-----+-----+-----+-----+-----+ 1860
CAGCATTAGTCGTGGCGTAGTCGTTACATAGACGGCACGTGACGTTGTTGCGACGAAGC

GCCTGGTAATGGCCCGCCGCTTCCAGCGTTCGACCCAGGCGTTAGGGTCAATGCGGGTC
1861 -----+-----+-----+-----+-----+-----+ 1920
CGGACCATTACCGGGCGGCGGAAGGTCGCAAGCTGGGTCCGCAATCCCAGTTACGCCCG

A
p
a
L
I

GCTTCACTTACGCCAATGTCGTTATCCAGCGGTGCACGGGTGAACTGATCGCGCAGCGGC
1921 -----+-----+-----+-----+-----+-----+ 1980
CGAAGTGAATGCGGTTACAGCAATAGGTCGCCACGTGCCCACTTGACTAGCGCGTCGCCG

B
g
l
I

GTCAGCAGTTGTTTTTATCGCCAATCCACATCTGTGAAAGAAAGCCTGACTGGCGGTTA
1981 -----+-----+-----+-----+-----+-----+ 2040
CAGTCGTCAACAAAAAATAGCGGTTAGGTGTAGACACTTTCTTTCGGACTGACCCCAAT

B
S
t
X
I

2041 AATTGCCAACGCTTATTACCCAGCTCGATGCAAAAATCCATTTTCGCTGGTGGTCAGATGC 2100
-----+-----+-----+-----+-----+-----+
TTAACGGTTGCGAATAATGGGTTCGAGCTACGTTTTTAGGTAAAGCGACCACCAGTCTACG

2101 GGGATGGCGTGGGACGCGCGGGGAGCGTCACTGAGGTTTTCCGCCAGACGCCACTGC 2160
-----+-----+-----+-----+-----+-----+
CCCTACCGCACCTGCGCCGCCCTCGCAGTGTGACTCCAAAAGGCGGTCTGCGGTGACG

B D M
m r l
g d u
I I I

2161 TGCCAGGCGCTGATGTGCCCGGCTTCTGACCATGCGGTTCGCGTTTCGCTTGCCTACGCGT 2220
-----+-----+-----+-----+-----+-----+
ACGGTCCGCGACTACACGGGCCGAAGACTGGTACGCCAGCGCAAGCCAACGTGATGCGCA

2221 ACTGTGAGCCAGAGTTGCCCGGCGCTCTCCGGCTGCGGTAGTTTCAGGCAGTTCAATCAAC 2280
-----+-----+-----+-----+-----+-----+
TGACACTCGGTCTCAACGGGCCGCGAGAGGCCGACGCCATCAAGTCCGTCAAGTTAGTTG

2281 TGTTTACCTTGTGGAGCGACATCCAGAGGCACTTCACCGCTTGCCAGCGGCTTACCATCC 2340
-----+-----+-----+-----+-----+-----+
ACAAATGGAACACCTCGCTGTAGGTCTCCGTGAAGTGGCGAACGGTCGCCGAATGGTAGG

B
aS
ns
It
II
/

2341 AGCGCCACCATCCAGTGCAGGAGCTCGTTATCGCTATGACGGAACAGGTATTCGCTGGTC 2400
-----+-----+-----+-----+-----+-----+
TCGCGGTGGTAGGTACGTCCTCGAGCAATAGCGATACTGCCTTGTCCATAAGCGACCAG

2401 ACTTCGATGGTTTGCCCGGATAAACGGAACTGGAAAACTGCTGCTGGTGTGTTTTGCTTCC 2460
-----+-----+-----+-----+-----+-----+
TGAAGCTACCAAACGGGCCTATTTGCCTTGACCTTTTTGACGACGACCACAAAACGAAGG

E
c
o
4
7
I
I
I

GTCAGCGCTGGATGCGGCGTTCGGTTCGGCAAAGACCAGACCGTTCATACAGAAGTGGCGA

2461 -----+-----+-----+-----+-----+-----+ 2520
CAGTCGCGACCTACGCCGCACGCCAGCCGTTTCTGGTCTGGCAAGTATGTCTTGACCGCT

P
v
u
I

TCGTTTCGGCGTATCGCCAAAATCACCGCCGTAAGCCGACCACGGGTTGCCGTTTTTCATCA
2521 -----+-----+-----+-----+-----+-----+ 2580
AGCAAGCCGCATAGCGGTTTTAGTGGCGGCATTTCGGCTGGTGCCCAACGGCAAAAGTAGT

T
B a
f q
i I
I I

TATTTAATCAGCGACTGATCCACCCAGTCCCAGACGAAGCCGCCCTGTAAACGGGGATAC
2581 -----+-----+-----+-----+-----+-----+ 2640
ATAAATTAGTCGCTGACTAGGTGGGTTCAGGGTCTGCTTCGGCGGGACATTTGCCCCCTATG

TGACGAAACGCCTGCCAGTATTTAGCGAAACCGCCAAGACTGTTACCCATCGCGTGGGCG
2641 -----+-----+-----+-----+-----+-----+ 2700
ACTGCTTTGCGGACGGTCATAAATCGCTTTGGCGGTTCTGACAATGGGTAGCGCACCCGC

TATTCGCAAAGGATCAGCGGGCGCGTCTCTCCAGGTAGCGAAAGCCATTTTTTGTATGGAC
2701 -----+-----+-----+-----+-----+-----+ 2760
ATAAGCGTTTTCTAGTCGCCCCGCGCAGAGAGGTCCATCGCTTTTCGGTAAAAAACTACCTG

B
s
s
H
I
I

CATTTTCGGCACAGCCGGGAAGGGCTGGTCTTCATCCACGCGCGGTACATCGGGCAAATA
2761 -----+-----+-----+-----+-----+-----+ 2820
GTAAAGCCGTGTTCGGCCCTTCCCGACCAGAAGTAGGTGCGCGCGCATGTAGCCCGTTTTAT

ATATCGGTGGCCGTGGTGTTCGGCTCCGCCGCTTCATACTGCACCGGGCGGGAAGGATCG
2821 -----+-----+-----+-----+-----+-----+ 2880
TATAGCCACCGGCACCACAGCCGAGGCGGCGGAAGTATGACGTGGCCCGCCCTTCTTAGC

ACAGATTTGATCCAGCGATACAGCGCGTCGTGATTAGCGCCGTGGCCTGATTCATTCCCC
2881 -----+-----+-----+-----+-----+-----+ 2940
TGTCTAAACTAGGTTCGCTATGTCGCGCAGCACTAATCGCGGCACCGGACTAAGTAAGGGG

A B
v s P M
v a v l

```

                a           B   u           u
                I           I   I           I
2941  -----+-----+-----+-----+-----+-----+ 3000
      TCGCTGGTCTACTAGTGTGAGCCCACTAATGCTAGCGCGACGTGGTAAGCGCAATGCGCA

      TCGCTCATCGCCGGTAGCCAGCGCGGATCATCGGTGAGACGATTATTGGCACCATGCCG
3001  -----+-----+-----+-----+-----+-----+ 3060
      AGCGAGTAGCGGCCATCGGTGCGGCCTAGTAGCCAGTCTGCTAAGTAACCGTGGTACGGC

                S           D
                s           r
                p           a
                I           I
                I           I
3061  -----+-----+-----+-----+-----+-----+ 3120
      ACCCAAAGTTATAACCGAAGTAGGTGGTGTATGTCCGGCATCGCCAGCGTGTGCGACATG

                D
                r
                d
                I
                I
3121  -----+-----+-----+-----+-----+-----+ 3180
      CACAGCGGATGGTTCGGATAATGCGAACAGCGCACGGCGTTAAAGTTGTTCTGCTTCATC
      GTGTGCGCTACCAAGCCTATTACGCTTGTGCGGTGCCGCAATTTCAACAAGACGAAGTAG

                E           B
                c           s
                o           s
                R           S
                V           I
3181  -----+-----+-----+-----+-----+-----+ 3240
      AGCAGGATATCCTGCACCATCGTCTGCTCATCCATGACCTGACCATGCAGAGGATGATGC
      TCGTCCTATAGGACGTGGTAGCAGACGAGTAGTACTGGACTGGTACGTCTCCTACTACG

                H
                p
                a
                I
3241  -----+-----+-----+-----+-----+-----+ 3300
      TCGTGACGGTTAACGCCTCGAATCAGCAACGGCTTGCCGTTTCAGCAGCAGCAGACCATTT
      AGCACTGCCAATTGCGGAGCTTAGTCGTTGCCGAACGGCAAGTCGTTCGTCTGGTAAA

      TCAATCCGCACCTCGCGGAAACCGACATCGCAGGCTTCTGCTTCAATCAGCGTGCCGTCG
3301  -----+-----+-----+-----+-----+-----+ 3360
      AGTTAGGCGTGGAGCGCCTTTGGCTGTAGCGTCCGAAGACGAAGTTAGTCGCACGGCAGC

      GCGGTGTGCAGTTCAACCACCGCACGATAGAGATTTCGGGATTTTCGGCGCTCCACAGTTTC
3361  -----+-----+-----+-----+-----+-----+ 3420
      CGCCACACGTCAAGTTGGTGGCGTGCTATCTCTAAGCCCTAAAGCCGCGAGGTGTCAAAG

```


v l
u a
I I
GGGTTTTTCGACGTTTCAGACGTAGTGTGACGCGATCGGCATAACCACCACGCTCATCGATA
3421 -----+-----+-----+-----+-----+-----+ 3480
CCCAAAGCTGCAAGTCTGCATCACACTGCGCTAGCCGTATTGGTGGTGCGAGTAGCTAT

B
S
P
M
I
ATTTACCCGCCGAAAGGCGCGGTGCCGCTGGCGACCTGCGTTTTACCCTGCCATAAAGAA
3481 -----+-----+-----+-----+-----+-----+ 3540
TAAAGTGGCGGCTTTCCGCGCCACGGCGACCGCTGGACGCAAAGTGGGACGGTATTTCTT

T
a
t
I
ACTGTTACCCGTAGGTAGTCACGCAACTCGCCGCACATCTGAACTTCAGCCTCCAGTACA
3541 -----+-----+-----+-----+-----+-----+ 3600
TGACAATGGGCATCCATCAGTGC GTT GAGCGGCGTGTAGACTTGAAGTCGGAGGTCATGT

GCGCGGCTGAAATCATCATTAAAGCGAGTGGCAACATGGAAATCGCTGATTTGTGTAGTC
3601 -----+-----+-----+-----+-----+-----+ 3660
CGCGCCGACTTTAGTAGTAATTTGCTCACCGTTGTACCTTTAGCGACTAAACACATCAG

A
a
t
I
I
GGTTTATGCAGCAACGAGACGTACGGAAAATGCCGCTCATCCGCCACATATCCTGATCT
3661 -----+-----+-----+-----+-----+-----+ 3720
CCAAATACGTCGTTGCTCTGCAGTGCCTTTTACGGCGAGTAGGCGGTGTATAGGACTAGA

TCCAGATAACTGCCGTCCTCCAACGCAGCACCATCACCGCGAGGCGGTTTTCTCCGGCG
3721 -----+-----+-----+-----+-----+-----+ 3780
AGGTCTATTGACGGCAGTGAGGTTGCGTCGTGGTAGTGGCGCTCCGCCAAAAGAGGCCGC

CGTAAAAATGCGCTCAGGTCAAATTCAGACGGCAAACGACTGTCCTGGCCGTAACCGACC
3781 -----+-----+-----+-----+-----+-----+ 3840
GCATTTTTACGCGAGTCCAGTTTAAGTCTGCCGTTTGTGCTGACAGGACCGGCATTGGCTGG

H
P
a
I
CAGCGCCCGTTGCACCACAGATGAAACGCCGAGTTAACGCCATCAAAAATAATTGCGGTC
3841 -----+-----+-----+-----+-----+-----+ 3900
GTCGCGGGCAACGTGGTGTCTACTTTGCGGCTCAATTGCGGTAGTTTTTTATTAAGCGCAG

S
f

c
I

3901 TGGCCTTCCTGTAGCCAGCTTTTCATCAACATTAATGTGAGCGAGTAACAACCCGTCGGA
-----+-----+-----+-----+-----+ 3960
ACCGGAAGGACATCGGTTCGAAAGTAGTTGTAATTTACACTCGCTCATTGTTGGGCAGCCT

3961 TTCTCCGTGGGAACAAACGGCGGATTGACCGTAATGGGATAGGTTACGTTGGTGTAGATG
-----+-----+-----+-----+-----+ 4020
AAGAGGCACCCTTGTTTGCCGCCTAACTGGCATTACCCTATCCAATGCAACCACATCTAC

B
s
B u
a 3
e 6
I I

4021 GGCGCATCGTAACCGTGCATCTGCCAGTTTGAGGGGACGACGACAGTATCGGCCTCAGGA
-----+-----+-----+-----+-----+ 4080
CCGCGTAGCATTGGCACGTAGACGGTCAAACCTCCCCTGCTGCTGTCATAGCCGGAGTCTT

H
g
i
E
I
I

4081 AGATCGCACTCCAGCCAGCTTTCCGGCACCCTTCTGGTGCCGAAACCAGGCAGGCGC
-----+-----+-----+-----+-----+ 4140
TCTAGCGTGAGGTTCGGTCGAAAGGCCGTGGCGAAGACCACGGCCTTTGGTCCGTTTTCGCG

B F P E
g s v a
l p u r
I I I I

4141 CATTGCCATTTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTA
-----+-----+-----+-----+-----+ 4200
GTAAGCGGTAAGTCCGACGCGTTGACAACCCTTCCCCTAGCCACGCCCGGAGAAGCGAT

4201 TTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGG
-----+-----+-----+-----+-----+ 4260
AATGCGGTTCGACCGCTTTCCCCTACACGACGTTCCGCTAATTCAACCCATTGCGGTCCC

B B
f a
i m
I H
I I

4261 TTTTCCCAGTACGACGTTGTAAAACGACGGGATCCGCGCTAGAGAATAGGAACTTCGGA
-----+-----+-----+-----+-----+ 4320
AAAAGGGTCAGTGCTGCAACATTTTGCTGCCCTAGGCGCGATCTCTTATCCTTGAAGCCT

E
c

o
4
7
I
I
I

F
s
P
I

4321 ATAGGAACTTCAAAGCGTTTCCGAAAACGAGCGCTTCCGAAAATGCAACGCGAGCTGCGC 4380
-----+-----+-----+-----+-----+-----+
TATCCTTGAAGTTTCGCAAAGGCTTTTGCTCGCGAAGGCTTTTACGTTGCGCTCGACGCG

B
s
b
I

4381 ACATACAGCTCACTGTTTCACGTGCGCACCTATATCTGCGTGTTCCTGTATATATATATAC 4440
-----+-----+-----+-----+-----+-----+
TGTATGTCGAGTGACAAGTGCAGCGTGGATATAGACGCACAACGGACATATATATATATG

4441 ATGAGAAGAACGGCATAGTGCCTGTTTATGCTTAAATGCGTACTTATATGCGTCTATTTA 4500
-----+-----+-----+-----+-----+-----+
TACTCTTCTTGCCGTATCACGCACAAATACGAATTTACGCATGAATATACGCAGATAAAT

4501 TGTAGGATGAAAGGTAGTCTAGTACCTCCTGTGATATTATCCCATTCCATGCGGGGTATC 4560
-----+-----+-----+-----+-----+-----+
ACATCCTACTTTCCATCAGATCATGGAGGACACTATAATAGGGTAAGGTACGCCCCATAG

B
s
e
R
I

M
u
n
I

4561 GTATGCTTCCTTCAGCACTACCCTTTAGCTGTTCTATATGCTGCCACTCCTCAATTGGAT 4620
-----+-----+-----+-----+-----+-----+
CATACGAAGGAAGTCGTGATGGGAAATCGACAAGATATACGACGGTGAGGAGTTAACCTA

N
d
e
I

N
s
i
I

4621 TAGTCTCATCCTTCAATGCTATCATTTTCTTTGATATTGGATCATATGCATAGTACCGAG 4680
-----+-----+-----+-----+-----+-----+
ATCAGAGTAGGAAGTTACGATAGTAAAGGAAACTATAACCTAGTATACGTATCATGGCTC

S
p
e
I

B
s
t
l
1
0
7
I

M
s
c
I

4681 AAAC TAGTGCGAAGTAGTGATCAGGTATTGCTGTTATCTGATGAGTATACGTTGTCCTGG 4740
-----+-----+-----+-----+-----+-----+
TTTGATCAGCTTCATCACTAGTCCATAACGACAATAGACTACTCATATGCAACAGGACC

E
c
o
R O
l l
e 0
A 9
I I

4741 -----+-----+-----+-----+-----+-----+ 4800
 CCACGGCAGAAGCACGCTTATCGCTCCAATTTCCACAACATTAGTCAACTCCGTTAGGC
 GGTGCCGTCTTCGTGCGAATAGCGAGGTTAAAGGGTGTGTAATCAGTTGAGGCAATCCG

4801 -----+-----+-----+-----+-----+-----+ 4860
 CCTTCATTGAAAGAAATGAGGTCATCAAATGTCTTCCAATGTGAGATTTTGGGCCATTTT
 GGAAGTAACTTTCTTTACTCCAGTAGTTTACAGAAGGTTACTACTCTAAAACCCGGTAAAA

H
i
n
d
I
I
I

4861 -----+-----+-----+-----+-----+-----+ 4920
 TTATAGCAAAGATTGAATAAGGCGCATTTTTCTTCAAAGCTTTATTGTACGATCTGACTA
 AATATCGTTTTCTAACTTATTCCGCGTAAAAAGAAGTTTCGAAATAACATGCTAGACTGAT

4921 -----+-----+-----+-----+-----+-----+ 4980
 AGTTATCTTTTAATAATTGGTATTCTGTTTATTGCTTGAAGAATTGCCGGTCTATTTA
 TCAATAGAAAATTATTAACCATAAGGACAAATAACGAACTTCTTAACGGCCAGGATAAAT

D E
r c C
d o l
I R a
I I I

4981 -----+-----+-----+-----+-----+-----+ 5040
 CTCGTTTTAGGACTGGTTCAGAATTCCTCAAAAATTCATCCAAATATACAAGTGGATCGA
 GAGCAAAATCCTGACCAAGTCTTAAGGAGTTTTTAAGTAGGTTTATATGTTTACCTAGCT

5041 -----+-----+-----+-----+-----+-----+ 5100
 TCCTACCCCTTGCCTAAAGAAGTATATGTGCCTACTAACGCTTGTCTTTGTCTCTGTCA
 AGGATGGGGAACGCGATTTCTTCATATACACGGATGATTGCGAACAGAAACAGAGACAGT

S
w
a
I

5101 -----+-----+-----+-----+-----+-----+ 5160
 CTAAACACTGGATTATTACTCCAGATACTTATTTTTGGACTAATTTAAATGATTTCCGGAT
 GATTTGTGACCTAATAATGAGGGTCTATGAATAAAACCTGATTAAATTTACTAAAGCCTA

P
 s
 p
 l
 4 M E
 0 u a
 6 n r
 I I I
 CAACGTTCTTAATATCGCTGAATCTTCCACAATTGATGAAAGTAGCTAGGAAGAGGAATT
 5161 -----+-----+-----+-----+-----+-----+ 5220
 GTTGCAAGAATTATAGCGACTTAGAAGGTGTTAACTACTTTCATCGATCCTTCTCCTTAA

B
 s
 t
 l
 l
 0
 7
 I
 GGTATAAAGTTTTTGTGTTTTGTAAATCTCGAAGTATACTCAAACGAATTTAGTATTTTTCT
 5221 -----+-----+-----+-----+-----+ 5280
 CCATATTTCAAAAACAAAAACATTTAGAGCTTCATATGAGTTTGCTTAAATCATAAAAGA
 CAGTGATCTCCCAGATGCTTTCACCCTCACTTAGAAGTGCTTTAAGCATTTTTTTACTGT
 5281 -----+-----+-----+-----+-----+ 5340
 GTCCTAGAGGGTCTACGAAAGTGGGAGTGAATCTTCACGAAATTCGTAAAAAATGACA

N
 s
 p
 V
 GGCTATTTCCCTTATCTGCTTCTTCCGATGATTCGAACTGTAATTGCAAACACTTACAA
 5341 -----+-----+-----+-----+-----+ 5400
 CCGATAAAGGGAATAGACGAAGAAGGCTACTAAGCTTGACATTAACGTTTGATGAATGTT

E
 c
 o
 R
 V
 TATCAGTGATATCAGATTGATGTTTTTGTCCATAGTAAGGAATAATTGTAAATTCCCAAG
 5401 -----+-----+-----+-----+-----+ 5460
 ATAGTCACTATAGTCTAACTACAAAAACAGGTATCATTCCCTTATTAACATTTAAGGGTTC
 CAGGAATCAATTTCTTTAATGAGGCTTCCAGAATTGTTGCTTTTTGCGTCTTGTATTTAA
 5461 -----+-----+-----+-----+-----+ 5520
 GTCCTTAGTTAAAGAAATTACTCCGAAGGTCTTAACAACGAAAAACGCAGAACATAAATT
 ACTGGAGTGATTTATTGACAATATCGAACTCAGCGAATTGCTTATGATAGTATTATAGC
 5521 -----+-----+-----+-----+-----+ 5580
 TGACCTCACTAAATAACTGTTATAGCTTTGAGTCGCTTAACGAATACTATCATAATATCG

R
c
a
I

5581 TCATGAATGTGGCTCTCTTGATTGCTGTTCCGTTATGTGTAATCATCCAACATAAATAGG 5640
-----+-----+-----+-----+-----+-----+-----+
AGTACTTACACCGAGAGAACTAACGACAAGGCAATACACATTAGTAGGTTGTATTTATCC

A
p
a
B
I

5641 TTAGTTTCAGCAGCACATAATGCTATTTTTCTCACCTGAAGGTCTTTCAAACCTTTCCACAA 5700
-----+-----+-----+-----+-----+-----+-----+
AATCAAGTCGTCGTGTATTACGATAAAAAGAGTGGACTTCCAGAAAGTTTGGAAAGGTGTT

B
s
u
3
6
I

E
c
o
R
V

NS
sp
ph
II
/

5701 ACTGACGAACAAGCACCTtaggcgggggttttacacaggatatcgaactgcggCatGCTTA 5760
-----+-----+-----+-----+-----+-----+-----+
TGACTGCTTGTTCGTGGAatccgccccaaaatgtgtcctatagcttgacgccGtaCGAAT

B
a
m
H
I

A S
v m
a a
I I

K
p
n
I

S
t
u
I

B E
aSc
nso
ItR
III
/

5761 ATCCTCCAATGGATCCCCGGGTACCCGTTTTAAACCAGGCCTCAGTACCGAGCTCGAATT 5820
-----+-----+-----+-----+-----+-----+-----+
TAGGAGGTTACCTAGGGGCCCATGGGCAAATTTGGTCCGGAGTCATGGCTCGAGCTTAA

5821 CGTAATCATGTCATAGCTGTTTTCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAA 5880
-----+-----+-----+-----+-----+-----+-----+
GCATTAGTACAGTATCGACAAAGGACACACTTTAACAATAGGCGAGTGTTAAGGTGTGTT

5881 CATAAGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCAC 5940
-----+-----+-----+-----+-----+-----+-----+
GTATGCTCGGCCTTCGTATTTACATTTTCGACCCACGGATTACTCACTCGATTGAGTG

V
s
p
I

E
c
o
R
I

5941 ATTAATTGCGTTGCGCTCACTGCCCCGTTTTCCAGTCGGGAAACCTGTCGTGCCAGGGGAA 6000
-----+-----+-----+-----+-----+-----+-----+
TAATTAACGCAACGCGAGTGACGGGCGAAAGGTCAGCCCTTTGGACAGCACGGTCCCCTT

			E		
			c		
			o		
			OP		
	A		lp	M	F
	v		0u	s	s
	a		9M	c	p
	I		II	I	I
			/		

TTCCCCGGATCTTCCGATCTTCTCGGGCAGCGTTGGGTCTTGCCACGGGTGCGCATGAT
 6001 -----+-----+-----+-----+-----+-----+ 6060
 AAGGGGCCTAGAAGGCTAGAAGAGCCCGTCGCAACCCAGGACCGGTGCCACGCGTACTA

		E		
		c		
		o		
		OP		
		lp		
		0u		
		9M		
		II		
		/		

CGTGCTCCTGTCGTTGTTGAGGACCCGGCTAGGCTGGCGGGGTTGCCTTACTGGTTAGCA
 6061 -----+-----+-----+-----+-----+ 6120
 GCACGAGGACAGCAACAACCTCTGGGCCGATCCGACCGCCCAACGGAATGACCAATCGT

 GAATGAATCACCGATACGCGAGCGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCGAC
 6121 -----+-----+-----+-----+-----+ 6180
 CTTACTTAGTGGCTATGCGCTCGCTTGCCTTGCCTGACGACGACGTTTTGACGACGCTG

 CTGAGCAACAACATGAATGGTCTTCGGTTTTCCGTGTTTTCGTAAAGTCTGGAAACGCGGAA
 6181 -----+-----+-----+-----+-----+ 6240
 GACTCGTTGTTGTACTTACCAGAAGCCAAAGGCACAAAGCATTTCAGACCTTTGCGCCTT

	B		D
	s		r
	a		d
	A		I
	I		I

GTCCCCTACGTGCTGCTGAAGTTGCCCGCAACAGAGAGTGGAACCAACCGGTGATACCAC
 6241 -----+-----+-----+-----+-----+ 6300
 CAGGGGATGCACGACGACTTCAACGGGCGTTGTCTCTCACCTTGGTTGGCCACTATGGTG

	A
	h
	d
	I

GATACTATGACTGAGAGTCAACGCCATGAGCGGCCTCATTCTTATTCTGAGTTACAACA
 6301 -----+-----+-----+-----+-----+ 6360
 CTATGATACTGACTCTCAGTTGCGGTACTCGCCGGAGTAAAGAATAAGACTCAATGTTGT

P
s
h
A
I

6361 GTCCGCACCGCTGTCCGGTAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGC
-----+-----+-----+-----+-----+-----+ 6420
CAGGCGTGGCGACAGGCCATCGAGGAAGGCCACCCGCGCCCCGTACTGATAGCAGCGGCG

P
f
l N
l g
l o
0 A
8 I
I V

6421 ACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCCCAACAG
-----+-----+-----+-----+-----+ 6480
TGAATACTGACAGAAGAAATAGTACGTTGAGCATCCTGTCCACGGCCGTGCGGGTTGTC

E
c
o
O
l
0
9
I

6481 TCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCCCTGCACCATTAT
-----+-----+-----+-----+-----+ 6540
AGGGGGCCGGTGCCCCGGACGGTGGTATGGGTGCGGCTTTGTTTCGCGGGACGTGGTAATA

B B
s s
p a
E B
I I

6541 GTTCCGGATCTGCATCGCAGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATT
-----+-----+-----+-----+-----+ 6600
CAAGGCCTAGACGTAGCGTCCTACGACGACCGATGGGACACCTTGTGGATGTAGACATAA

E
c
o
4
7
I
I
I

6601 AACGAAGCGCTAACCGTTTTTATCAGGCTCTGGGAGGCAGAATAAATGATCATATCGTCA
-----+-----+-----+-----+-----+ 6660
TTGCTTCGCGATTGGCAAAAATAGTCCGAGACCCTCCGTCTTATTTACTAGTATAGCAGT

B
s
u
3
6
I

ATTATTACCTCCACGGGGAGAGCCTGAGCAAACCTGGCCTCAGGCATTTGAGAAGCACACG
6661 -----+-----+-----+-----+-----+-----+ 6720
TAATAATGGAGGTGCCCCTCTCGGACTCGTTTGACCGGAGTCCGTAAACTCTTCGTGTGC

GTCACACTGCTTCCGGTAGTCAATAAACCGGTAAACCAGCAATAGACATAAGCGGCTATT
6721 -----+-----+-----+-----+-----+-----+ 6780
CAGTGTGACGAAGGCCATCAGTTATTTGGCCATTTGGTCGTTATCTGTATTGCGCGATAA

T
t
h
Dl N
r1 s
d1 p
II V
/

TAACGACCCTGCCCTGAACCGACGACCGGGTCAATTTGCTTTTCAATTTCTGCCATTCA
6781 -----+-----+-----+-----+-----+-----+ 6840
ATTGCTGGGACGGGACTTGGCTGCTGGCCCAGCTTAAACGAAAGCTTAAAGACGGTAAGT

B
m
g
I

TCCGCTTATTATCACTTATTAGGGCGTAGCACCAGGCGTTTAAAGGCACCAATAACTGCC
6841 -----+-----+-----+-----+-----+-----+ 6900
AGGCGAATAATAGTGAATAAGTCCGCATCGTGGTCCGCAAATTCCTGGTTATTGACGG

ST
ca
at
II
/

TTAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCAT
6901 -----+-----+-----+-----+-----+-----+ 6960
AATTTTTTTAATGCGGGGCGGGACGGTGAGTAGCGTCATGACAACATTAAGTAATTCGTA

TCTGCCGACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAG
6961 -----+-----+-----+-----+-----+-----+ 7020
AGACGGCTGTACCTTCGGTAGTGTCTGCCGTAATACTTGGACTTAGCGGTCGCCGTAGTC

S NS
s ct
p oY
I II
/

CACCTTGTCGCCTTGCCTATAATATTTGCCCATGGTGAAAACGGGGGCGAAGAAGTTGTC
7021 -----+-----+-----+-----+-----+-----+ 7080
GTGGAACAGCGGAACGCATATTATAAACGGGTACCACTTTTGCCCCGCTTCTTCAACAG

M T
s a
c q
I I

7081 CATATTGGCCACGTTTAAATCAAACTGGTGAAACTCACCCAGGGATTGGCTGAGACGAA 7140
-----+-----+-----+-----+-----+-----+-----+
GTATAACCGGTGCAAATTTAGTTTTGACCACTTTGAGTGGGTCCCTAACCGACTCTGCTT

7141 AAACATATTCTCAATAAACCCCTTTAGGGAAATAGGCCAGGTTTTACCGTAACACGCCAC 7200
-----+-----+-----+-----+-----+-----+-----+
TTTGTATAAGAGTTATTTGGGAAATCCCTTTATCCGGTCCAAAAGTGGCATTGTGCGGTG

7201 ATCTTGCGAATATATGTGTAGAACTGCCGGAATCGTCGTGGTATTCCTCCAGAGCGA 7260
-----+-----+-----+-----+-----+-----+-----+
TAGAACGCTTATATACACATCTTTGACGGCCTTTAGCAGCACCATAAGTGAGGTCTCGCT

P
s
p
l
4
0
6
I

7261 TGAAAACGTTTTCAGTTTGCTCATGGAAAACGGTGTAACAAGGGTGAACACTATCCCATAT 7320
-----+-----+-----+-----+-----+-----+-----+
ACTTTTGCAAAGTCAAACGAGTACCTTTTGCCACATTGTTCCCACTTGTGATAGGGTATA

E B
c s
o p
R E
I I

7321 CACCAGCTCACCGTCTTTTCATTGCCATACGAATTCGGATGAGCATTTCATCAGGCGGGCA 7380
-----+-----+-----+-----+-----+-----+-----+
GTGGTCGAGTGGCAGAAAGTAACGGTATGCTTAAGGCCTACTCGTAAGTAGTCCGCCCGT

7381 AGAATGTGAATAAAGGCCGGATAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAG 7440
-----+-----+-----+-----+-----+-----+-----+
TCTTACACTTATTTCCGGCCTATTTTGAACACGAATAAAAAGAAATGCCAGAAATTTTTC

7441 GCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCC 7500
-----+-----+-----+-----+-----+-----+-----+
CGGCATTATAGGTGCGACTTGCCAGACCAATATCCATGTAACCTCGTTGACTGACTTTACGG

7501 TCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTATATCCAGTGATTTTT 7560
-----+-----+-----+-----+-----+-----+-----+
AGTTTTTACAAGAAATGCTACGGTAACCCTATATAGTTGCCACCATATAGGTCACTAAAAA

7561 TTCTCCATTTTAGCTTCCTTAGCTCCTGAAAATCTCGATAACTCAAAAAATACGCCCGGT 7620
-----+-----+-----+-----+-----+-----+-----+

AAGAGGTAAAATCGAAGGAATCGAGGACTTTTAGAGCTATTGAGTTTTTTATGCGGGCCA

B
s
a
A
I

AGTGATCTTATTTTATTATGGTCAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCAT
 7621 -----+-----+-----+-----+-----+-----+ 7680
 TCACTAGAATAAAGTAATACCACTTTCAACCTTGGAGAATGCACGGCTAGTTGCAGAGTA
 TTTCGCCAAAAGTTGGCCCAGGGCTTCCCGGTATCAACAGGGACACCAGGATTTATTTAT
 7681 -----+-----+-----+-----+-----+-----+ 7740
 AAAGCGGTTTTCAACCGGGTCCCGAAGGGCCATAGTTGTCCCTGTGGTCCTAAATAAATA
 TCTGCGAAGTGATCTTCCGTACAGGTATTTATTTCGGCGCAAAGTGCGTCGGGTGATGCT
 7741 -----+-----+-----+-----+-----+-----+ 7800
 AGACGCTTCACTAGAAGGCAGTGTCCATAAATAAGCCGCGTTTTACGCAGCCCACTACGA
 GCCAACTTACTGATTTAGTGTATGATGGTGTTTTTGAGGTGCTCCAGTGGCTTCTGTTTC
 7801 -----+-----+-----+-----+-----+-----+ 7860
 CGGTTGAATGACTAAATCACATACTACCACAAAACTCCACGAGGTACCCGAAGACAAAG
 TATCAGCTGTCCCTCCTGTTTCACTACTGACGGGGTGGTGCGTAACGGCAAAGCACCGC
 7861 -----+-----+-----+-----+-----+-----+ 7920
 ATAGTCGACAGGGAGGACAAGTCGATGACTGCCCCACCACGCATTGCCGTTTTTCGTGGCG

E
c
o
4
7N
Ih
Ie
II

S
a
l
I

E
a
g
I

CGGACATCAGCGCTAGCCGGTCGACCGCGAACGCCAGCAAGACGTAGCCAGCGCGTCGGC
 7921 -----+-----+-----+-----+-----+-----+ 7980
 GCCTGTAGTCGCGATCGGCCAGCTGGCGCTTGCGGTCGTTCTGCATCGGTCGCGCAGCCG

P
s
p
l
4
0
6
I

N
g
oB
Ag
Il
VI

CGCCATGCCGGCGATAATGGCTTGCTTCTCGCCGAAACGTTTTGGTGGCGGGACCAGTGAC
 7981 -----+-----+-----+-----+-----+-----+ 8040
 GCGGTACGGCCGCTATTACCGAACGAAGAGCGGCTTTGCAAACCACCGCCCTGGTCACTG
 GAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGT
 8041 -----+-----+-----+-----+-----+-----+ 8100
 CTTCCGAACTCGCTCCCGCACGTTCTAAGGCTTATGGCGTTTCGCTGTCCGGCTAGTAGCA

E
c
o N
4 g
7 o
I A
I I
I V

CGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCC
8101 -----+-----+-----+-----+-----+-----+ 8160
GCGCGAGGTCGCTTTTCGCCAGGAGCGGCTTTTACTGGGTCTCGCGACGGCCGTGGACAGG

P
f
l
l P
l s
0 h
8 A
I I
TACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGC
8161 -----+-----+-----+-----+-----+ 8220
ATGCTCAACGTACTATTTCTTCTGTCTGTCAGTATTCACGCCGCTGCTATCAGTACGGGGCGCG

B
f
i
I
CCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCCCCCAGCGTCTGC
8221 -----+-----+-----+-----+-----+ 8280
GGTGGCCTTCCTCGACTGACCCAACTTCCGAGAGTTCCCGTAGCCAGGGGGTTCGCAGACG

P
s
h
A
I
CAGAATTACGCGACGAAGACAGAAGTCCACAAAATAGCATTACAGGGCTGTCAGAGTCAGG
8281 -----+-----+-----+-----+-----+ 8340
GTCTTAATGCGCTGCTTCTGTCTTCAGGTGTTTTATCGTAAGTCCCGACAGTCTCAGTCC
TGTGGATTGTCATGCGCCAGAATGCCAGGGAATTATTGAATTACAGGGCGACAGCGATG
8341 -----+-----+-----+-----+-----+ 8400
ACACCTAACAGTACGCGGTCTTACGGGTCCCTTAATAACTTAATGTCCCGCTGTCGCTAC

P
f
l
l
l
0

8

I

8401 -----+-----+-----+-----+-----+-----+-----+ 8460
 CGGCGATTGTGAAAGGGCTTATTGCGGTCGTCTTTATTCTCTACGATCAGATGACGCCGC
 GCCGCTAACACTTTCCCGAATAACGCCAGCAGAAATAAGAGATGCTAGTCTACTGCGGGC

8461 -----+-----+-----+-----+-----+-----+-----+ 8520
 AGGATATTGTCAATTTTCGATGTGCGTCCGTGGTTTGAAAAAATGGCGCTCACCCAACATC
 TCCTATAACAGTTAAAGCTACACGCAGGCACCAAACCTTTTTTACC GCGAGTGGGTTGTAG

T
 a
 q
 I
 I

8521 -----+-----+-----+-----+-----+-----+-----+ 8580
 TCACCCCATCTCGTTTACAAGGTCTGGAAGCGATGATTCGCGCAATTCGCGCCAAAGCCG
 AGTGGGGTAGAGCAAGTGTTCCAGACCTTCGCTACTAAGCGCGTTAAGCGCGGTTTCGGC

8581 -----+-----+-----+-----+-----+-----+-----+ 8640
 CTGCACTTAGCTAAACTAAAAGGACAGCTTTCATCCGGGTTGGCTCACATTGCCGGATGG
 GACGTGAATCGATTTGATTTTCTGTGCGAAAGTAGGCCCAACCGAGTGTAACGGCCTACC

B
 S
 P
 M
 I

8641 -----+-----+-----+-----+-----+-----+-----+ 8700
 TCAATGTTTACACTTGACCTGCACAGGAGTGTTTTGGTTTTAGGTGAACATAAGCGCCAT
 AGTTACAAATGTGAACTGGACGTGTCCTCACAAAACCAAAGTCCACTTGTATTTCGCGGTA

8701 -----+-----+-----+-----+-----+-----+-----+ 8760
 CCGACGTTTTCAGCGTGAGTCTCCGGCAAATACAGGAGGTTTACAATGAAACGCGGTCTT
 GGCTGCAAAGTCGCACTCAGAGGCCGTTTATGTCCTCAAATGTTACTTTGCGCGCAGAA

B
 g
 l
 I

8761 -----+-----+-----+-----+-----+-----+-----+ 8820
 TGCTTACACTCACGCTTATCGGCGCTTTTAGCGCCATCCAGGCTGCCTGGGCGGTTGATT
 ACGAATGTGAGTGCGAATAGCCGCGAAAATCGCGGTAGGTCCGACGGACCCGCCAACTAA

B
 S
 t
 l
 l
 0
 7
 I

8821 -----+-----+-----+-----+-----+-----+-----+ 8880
 ATCCGCTACCACCAACCGGAAGCCGACTGGTTGGGCAAATCAAACGTATACGGTGCAAG
 TAGGCGATGGTGGTTGGCCTTCGGCTGACCAACCCGTTTTAGTTTGCATATGCCACGTTT

8881 AAGGGGATAAAAACCTTCAGGCTATTGCCCGACGTTTTTGATACTGCGGCAATGTTGATCC 8940
-----+-----+-----+-----+-----+-----+
TTCCCCTATTTTTGGAAAGTCCGATAACGGGCTGCAAAACTATGACGCCGTTACAACCTAGG

S
e
x
A
I

8941 TTGAAGCCAATAAACAATATCGCCCCGGTGCCAAAACCTGGTACGACGATAACTATTTCCTT 9000
-----+-----+-----+-----+-----+-----+
AACTTCGGTTATTGTGATAGCGGGCCACGGTTTTGGACCATGCTGCTATTGATAAGGAA

9001 CACAACCTGTTATTACCTGATGCACCGCGTCAGGGGATTATCGTTAGCATAAAGTGTAAG 9060
-----+-----+-----+-----+-----+-----+
GTGTTGACAATAATGGACTACGTGGCGCAGTCCCCTAATAGCAATCGTATTTTCACATTTT

V
s
p
I

9061 CCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTT 9120
-----+-----+-----+-----+-----+-----+
GGACCCACGGATTACTCACTCGATTGAGTGTAATTAACGCAACGCGAGTGACGGGCGAA

V
s
p
I
B
s
a
X
I

9121 TCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGGGGAGAG 9180
-----+-----+-----+-----+-----+-----+
AGGTCAGCCCTTTGGACAGCACGGTGCAGTAATTAAGTCCGTTGCGCGCCCCTCTC

N
a
r
I

9181 GCGGTTTTCGCTATTGGGCGCCAGGGTGGTTTTTCTTTTACCAGTGAGACGGGCAACAGC 9240
-----+-----+-----+-----+-----+-----+
CGCCAAACGCATAACCCGCGGTCCCACCAAAAAGAAAAGTGGTCACTCTGCCCGTTGTCG

9241 TGATTGCCCTTACCCGCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGC 9300
-----+-----+-----+-----+-----+-----+
ACTAACGGGAAGTGGCGGACCGGGACTCTCTCAACGTCGTTCCGCGAGGTGCGACCAAACG

9301 CCCAGCAGGCGAAAATCCTGTTTGTGGTGGTTGACGGCGGGATATAACATGAGCTGTCT 9360
-----+-----+-----+-----+-----+-----+
GGGTCGTCCGCTTTTAGGACAAACTACCACCAACTGCCGCCCTATATTGTAICTGACAGA

E
c

O
R
V

TCGGTATCGTTCGTATCCCAGTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTA
9361 -----+-----+-----+-----+-----+-----+ 9420
AGCCATAGCAGCATAGGGTGATGGCTCTATAGGCGTGGTTGCGCGTCGGGCCTGAGCCAT

B
s
s
H
I
I

ATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACG
9421 -----+-----+-----+-----+-----+-----+ 9480
TACCGCGCGTAACGCGGGTCGCGGTAGACTAGCAACCGTTGGTTCGTAGCGTCACCCTTGC

ATGCCCTCATTTCAGCATTTGCATGGTTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCT
9481 -----+-----+-----+-----+-----+-----+ 9540
TACGGGAGTAAGTCGTAAACGTACCAACAACCTTTTGGCCTGTACCGTGAGGTCAGCGGA

TCCC GTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGA
9541 -----+-----+-----+-----+-----+-----+ 9600
AGGGCAAGGCGATAGCCGACTTAAACTAACGCTCACTCTATAAATACGGTCGGTCGGTCT

B
B Aa
m pn
g aI
I II
/
B
s
t
E
I
I

CGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCC
9601 -----+-----+-----+-----+-----+-----+ 9660
GCGTCTGCGCGGCTCTGTCTTGAATTACCCGGGCGATTGTGCGCTAAACGACCACTGGG

B
s
p
2
4
I
B
f
i
I
T
a
q
I
I

AATGCGACCAGATGCTCCACGCCAGTCGCGTACCGTCTTCATGGGAGAAAATAACTG
9661 -----+-----+-----+-----+-----+-----+ 9720
TTACGCTGGTCTACGAGGTGCGGGTCAGCGCATGGCAGAAGTACCCTCTTTTATTATGAC

TTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCT
9721 -----+-----+-----+-----+-----+-----+ 9780
AACTACCCACAGACCAGTCTCTGTAGTTCTTTATTGCGGCCTTGTAATCACGTCCGTCGA

B
s
t
M
l

s
p
l
4
0
6
I

A
p
a
B
I

TAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTACACCACC
 10141 -----+-----+-----+-----+-----+-----+-----+ 10200
 ATTCTCTGTGGCCGTATGAGACGCTGTAGCATATTGCAATGACCAAAGTGTAAAGTGGTGG

H
g
i
E
I
I

E
a
r
I

CTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTTCG
 10201 -----+-----+-----+-----+-----+-----+-----+ 10260
 GACTTAACTGAGAGAAGGCCCGCGATAGTACGGTATGGCGCTTTCAAAACGCGGTAAGC

B
s S
p s
M p
I I

ATGGTGTCAACCTTGCAGAGCTGCGCCTTTATTATTATCCGCCGGGAGAAAATATTGTGC
 10261 -----+-----+-----+-----+-----+-----+-----+ 10320
 TACCACAGTTGGAACGTCTCGACGCGGAAATAATAATAGGCGGCCCTCTTTTATAACACG

S
g M B
r l s
A u b
I I I

AGGTTTATCCAATAGGTATTGGATTGCAGGGGCTGGAAACGCCGGTGATGGAAACGCGTG
 10321 -----+-----+-----+-----+-----+-----+-----+ 10380
 TCCAAATAGGTTATCCATAACCTAACGTCCCCGACCTTTGCGGCCACTACCTTTGCGCAC

B E
s c
p o
G N
I I

TTGGGCAGAAAATCCCTAACCCAACCTGGACGCCTACGGCAGGCATTTCGTGACGTTTCGC
 10381 -----+-----+-----+-----+-----+-----+-----+ 10440
 AACCCGTCTTTTAGGGATTGGGTTGGACCTGCGGATGCCGTCCGTAAGCAGTCGCAAGCG

TGGAGCGTGGCATTAAATTACCGCCAGTCGTTCTGCGGACCAAATAACCCGCTAGGAC
 10441 -----+-----+-----+-----+-----+-----+-----+ 10500
 ACCTCGCACCGTAATTTAATGGCGGTGAGCAAGGACGGCCTGGTTTATTGGGCGATCCTG

K
p
n
I

GTTACGCACTGCGCCTCGCGCATGGTAATGGCGAATACCTCATTTCATGGTACCAGTGCGC

10501 -----+-----+-----+-----+-----+-----+ 10560
CAATGCGTGACGCGGAGCGCGTACCATTACCGCTTATGGAGTAAGTACCATGGTCACGCG

T
t
h
l
l
l
I

CGGACAGCGTCGGTTTTGCGCGTCAGTTCAGGGTGTATTTCGCATGAATGCACCGGATATTA
10561 -----+-----+-----+-----+-----+-----+ 10620
GCCTGTGCGCAGCCAAACGCGCAGTCAAGTCCCACATAAGCGTACTTACGTGGCCTATAAT

S
g
r
A
I

AAGCCTTGTTCTCCAGCGTGCGGACGGGAACGCCGGTGAAAGTGATCAACGAACCGGTGA
10621 -----+-----+-----+-----+-----+-----+ 10680
TTCGGAACAAGAGGTGCGCACGCCTGCCCTTGCGGCCACTTTCACTAGTTGCTTGGCCACT

S
s
p
I

S N
a h
l e
I I

AATATTCCGTGGAGCCTAACGGGATGCGTTATGTTGAANTGAAAACCGGTTCGACCGGCTA
10681 -----+-----+-----+-----+-----+-----+ 10740
TTATAAGGCACCTCGGATTGCCCTACGCAATACAACCTTNACTTTTGGCCAGCTGGCCGAT

B
s
t
l
l
0
7
I

X
m
n
I

GCGGAGTGTATACTGGCTTACTATGTTGGCACTGATGAGGGTGTGAGTGAAGTGCTTCAT
10741 -----+-----+-----+-----+-----+-----+ 10800
CGCCTCACATATGACCGAATGATACAACCGTGACTACTCCCACAGTCACTTACGAAGTA

S
g
r
A
I

GTGGCAGGAGAAAAAAGGCTGCACCGGTGCGTCAGCAGAATATGTGATACAGGATATATT
10801 -----+-----+-----+-----+-----+-----+ 10860
CACCGTCCTCTTTTTTCCGACGTGGCCACGCAGTCGTCTTATACTATGTCCTATATAA

CCGCTTCTCGCTCACTGACTCGCTACGCTCGGTTCGACTGCGGCGAGCGGAAATGG
10861 -----+-----+-----+-----+-----+-----+ 10920

GGCGAAGGAGCGAGTGACTGAGCGATGCGAGCCAGCAAGCTGACGCCGCTCGCCTTTACC
CTTACGAACGGGGCGGAGATTTCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGA
10921 -----+-----+-----+-----+-----+-----+ 10980
GAATGCTTGCCCCGCCTCTAAAGGACCTTCTACGGTCCTTCTATGAATTGTCCCTTCACT

S
a
c
I
I

GAGGGCCGCGGCAAAGCCGTTTTTCCATAGGCTCCGCCCCCTGACAAGCATCACGAAAT
10981 -----+-----+-----+-----+-----+-----+ 11040
CTCCCGGCGCCGTTTTCGGCAAAAGGTATCCGAGGCGGGGGGACTGTTTCGTAGTGCTTTA

CTGACGCTCAAATCAGTGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGGCGTTTTCC
11041 -----+-----+-----+-----+-----+-----+ 11100
GACTGCGAGTTTGTAGTCACCACCGCTTTGGGCTGTCCTGATATTTCTATGGTCCGCAAAGG

B
s
s
S
I

CCCTGGCGGCTCCCTCGTGCCTCTCCTGTTCTGCCTTTTCGGTTTACCGGTGTCATTCC
11101 -----+-----+-----+-----+-----+-----+ 11160
GGGACCGCCGAGGGAGCACGCGAGAGGACAAGGACGGAAGCCAAATGGCCACAGTAAGG

GCTGTTATGGCCGCGTTTTGTCTCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGTTC
11161 -----+-----+-----+-----+-----+-----+ 11220
CGACAATACCGGCGCAAACAGAGTAAGGTGCGGACTGTGAGTCAAGGCCCATCCGTCAAG

B
s
p
G
I

GCTCCAAGCTGGACTGTATGCACGAACCCCCGTTTCAGTCCGACCGCTGCGCCTTATCCG
11221 -----+-----+-----+-----+-----+-----+ 11280
CGAGGTTTCGACCTGACATACGTGCTTGGGGGGCAAGTCAGGCTGGCGACGCGGAATAGGC

N
s
p
I

GTA ACTATCGTCTTGAGTCCAACCCGAAAGACATGCAAAGCACCCTGGCAGCAGCCA
11281 -----+-----+-----+-----+-----+-----+ 11340
CATTGATAGCAGA ACTCAGGTTGGGCCTTTCTGTACGTTTTTCGTGGTGACCGTCGTCGGT

B
s
e
R
I

CTGGTAATTGATTTAGAGGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAA

11341 -----+-----+-----+-----+-----+-----+ 11400
 GACCATTAACATAAATCTCCTCAATCAGAACTTCAGTACGCGGCCAATTCCGATTTGACTT

B
 s
 p
 2
 4
 I
 I

AGGACAAGTTTTGGTGACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTA

11401 -----+-----+-----+-----+-----+-----+ 11460
 TCCTGTTCAAACCCTGACGCGAGGAGTTTCGGTCAATGGAGCCAAGTTTCTCAACCAT

N
 s
 p
 V

GCTCAGAGAACCTTCGAAAAACCGCCCTGCAAGGCGGTTTTTTTCGTTTTTCAGAGCAAGAG

11461 -----+-----+-----+-----+-----+-----+ 11520
 CGAGTCTCTTGGAAGCTTTTTGGCGGGACGTTCCGCCAAAAAAGCAAAGTCTCGTTCTC

V
 s
 p
 I
 S
 s
 p
 I
 X
 b
 a
 I

ATTACGCGCAGACCAAAACGATCTCAAGAAGATCATCTTATTAATCAGATAAAAATATTTTC

11521 -----+-----+-----+-----+-----+-----+ 11580
 TAATGCGCGTCTGGTTTTGCTAGAGTTCTTCTAGTAGAATAATTAGTCTATTTTATAAAG

E
 a
 r
 I

TAGATTTTCAGTGCAATTTATCTCTTCAAATGTAGCACCTGAAGTCAGCCCCATACGATAT

11581 -----+-----+-----+-----+-----+-----+ 11640
 ATCTAAAGTCACGTTAAATAGAGAAGTTTACATCGTGGACTTCAGTCGGGGTATGCTATA

AAGTTGTAATTCTCATGTTTGACAGCTTATCAGCG

11641 -----+-----+-----+-----+-----+-----+ 11675
 TTCAACATTAAGAGTACAACTGTCTGAATAGTCGC

Enzymes that do cut and were not excluded:

AatII	AhdI	ApaI	ApaBI	ApaLI	AvaI	BaeI	BamHI
BanII	BfiI	BglI	BmgI	Bpu1102I	BsaAI	BsaBI	BsaXI
BsbI	BseRI	Bsp24I	BspEI	BspGI	BspLU11I	BspMI	BssHII
BssSI	Bst1107I	BstEII	BstXI	Bsu36I	ClaI	DraIII	DrdI
DrdII	EagI	EarI	Eco47III	EcoNI	EcoO109I	EcoRI	EcoRV
FspI	HgiEII	HindIII	HpaI	KpnI	MluI	MscI	MunI
NarI	NcoI	NdeI	NgoAIV	NheI	NsiI	NspI	NspV
Pfl1108I	PpuMI	PshAI	Psp1406I	PvuI	RcaI	RleAI	SacII
SalI	ScaI	SexAI	SfcI	SgrAI	SmaI	SpeI	SphI
SspI	SstI	StuI	StyI	SunI	SwaI	TaqII	TatI
Tth111I	VspI	XbaI	XmnI				

Enzymes that do not cut:

AflII	AscI	AvrII	BglII	BplI	BsaI	BsrGI	FseI
NotI	NruI	PacI	PmeI	PmlI	PstI	RsrII	SanDI
SapI	SfiI	SgfI	SnaBI	SrfI	Sse8387I	Sse8647I	UbaDI
XhoI							

Enzymes excluded; MinCuts: 1 MaxCuts: 5

AccI	AceIII	AciI	AflIII	AluI	AlwI	Alw26I	AlwNI
ApoI	AvaII	BanI	BbsI	BbvI	BccI	Bce83I	BcefI
BcgI	BcgI	BclI	BfaI	BpmI	Bpu10I	BsaHI	BsaJI
BsaWI	BscGI	BsgI	BsiEI	BsiHKAI	BslI	BsmI	BsmBI
BsmFI	Bsp24I	Bsp1286I	BsrI	BsrBI	BsrDI	BsrFI	BstYI
Cac8I	CjeI	CjeI	CjePI	CjePI	CviJI	CviRI	DdeI
DpnI	DraI	DsaI	EaeI	EciI	Eco57I	EcoRII	FauI
Fnu4HI	FokI	GdiII	HaeI	HaeII	HaeIII	HgaI	HhaI
Hin4I	HincII	HinfI	HphI	MaeIII	MboII	MmeI	MnlI
MseI	MslI	MspI	MspAlI	MwoI	NciI	NlaIII	NlaIV
PflMI	PinAI	PleI	PvuII	RsaI	Sau96I	Sau3AI	ScrFI
SfaNI	SimI	TaiI	TaqI	TaqII	TauI	TfiI	ThaI
TseI	Tsp45I	Tsp4CI	Tsp509I	TspRI	Tth111II	XcmI	