Sequence of pTYB12 from Web -June1, 2005

Plasmid pTYB12

Update 3/26/04

Features:

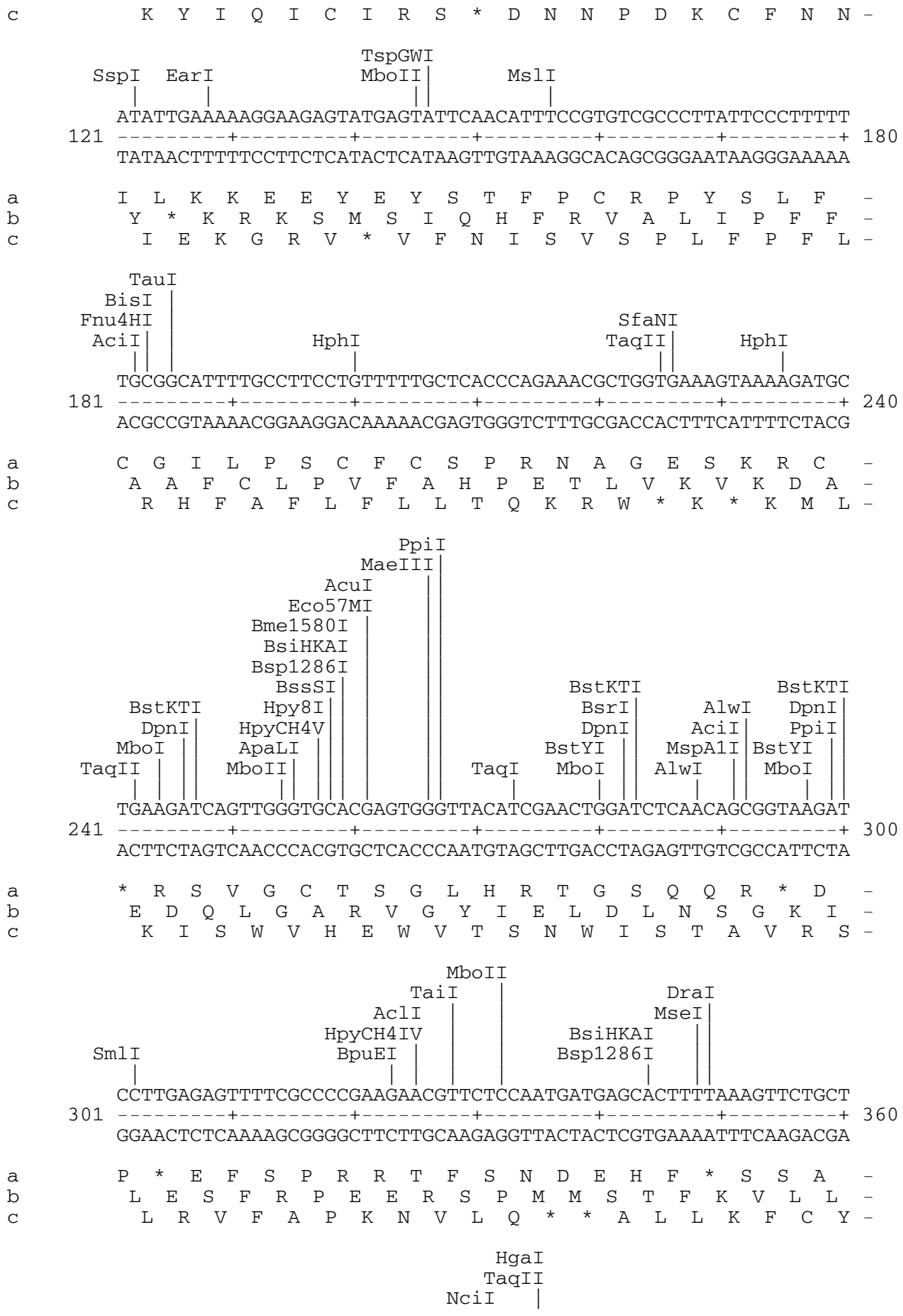
140-1000	beta-lactamase (bla; amp-r) CDS (start 140)
140- 208	beta-lactamase signal peptide CDS (start 140)
1042-1555	M13 origin of replication (- +)
1666-2254	pMB1 origin of replication (clockwise) (RNAII -35 to RNA/DNA switch point)
2814-2623	rop CDS (start 2814, complementary strand)
4453-3371	lacI CDS (start 4453, complementary strand)
4840-4857	T7 promoter (transcript start 4840 clockwise)
4840-4859	T7 universal primer sequence (NEB #S1248S)
4859-4883	lac operator
4913-4919	RBS for expression CDS
4928-6562	expression CDS (start 4928)
4928-4972	malE fragment (translational start)
4973-6502	Sce VMA intein with embedded CBD
4973-5791	Sce VMA intein 5' region
5798-5950	chitin binding domain (CBD)
5963-6502	Sce VMA intein 3' region
6503-6559	multiple cloning site (NdeI-SmaI)

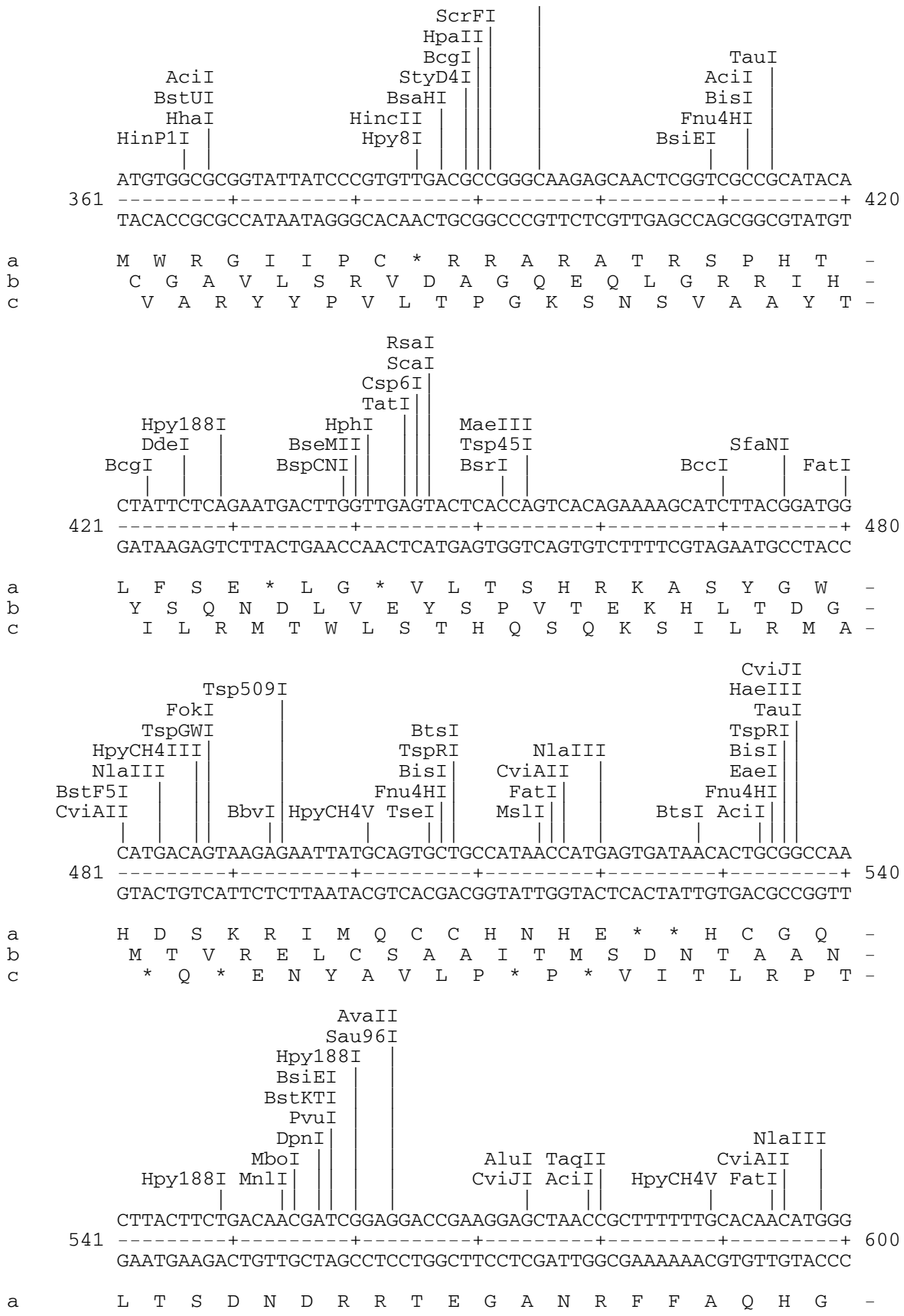
→ The target gene was inserted using Nde I and Eco RI sites

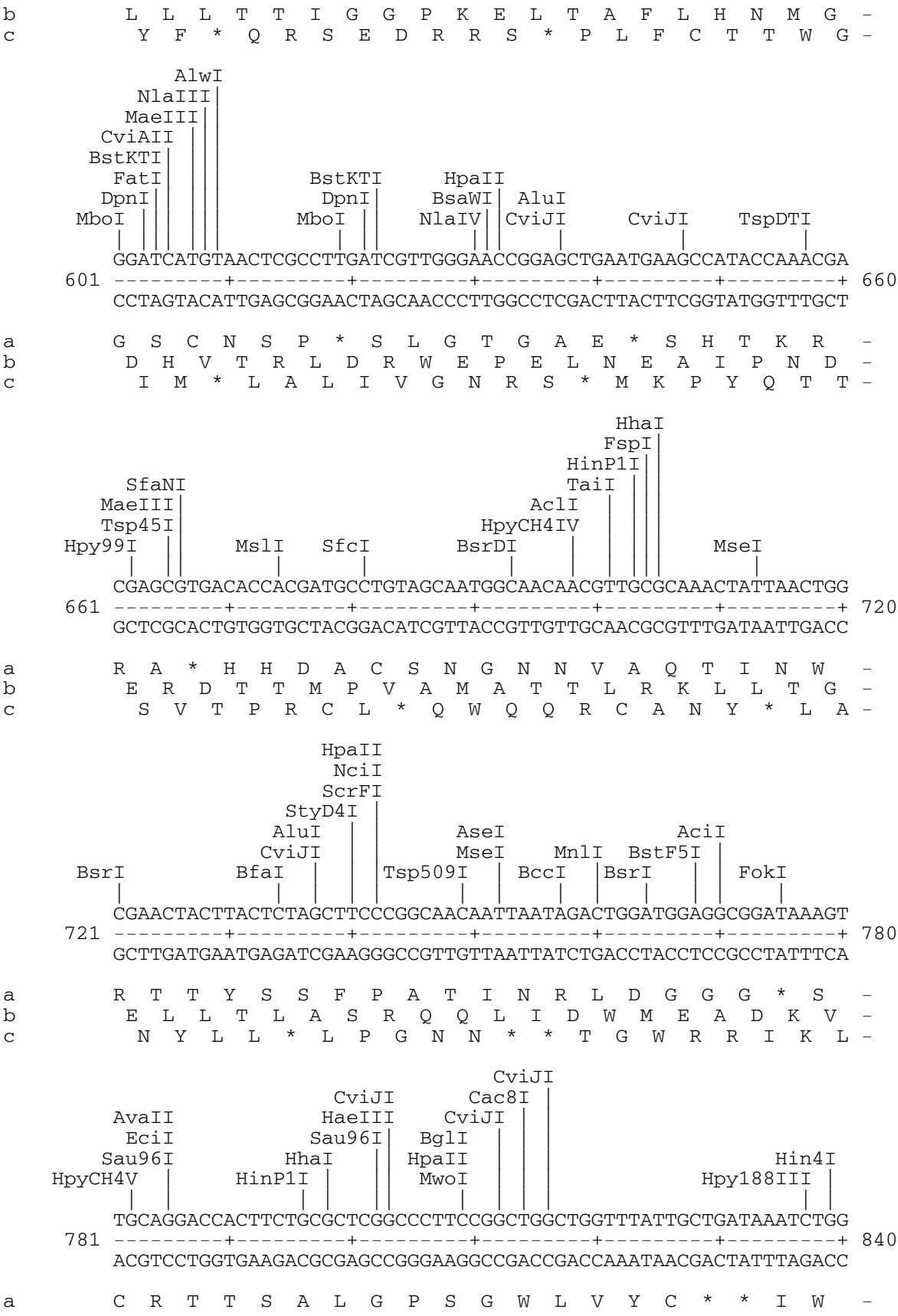
6362-6385	Intein forward primer sequence (NEB #S1263S)
6582-6704	T7 Tphi transcription terminator With 252 enzymes: *

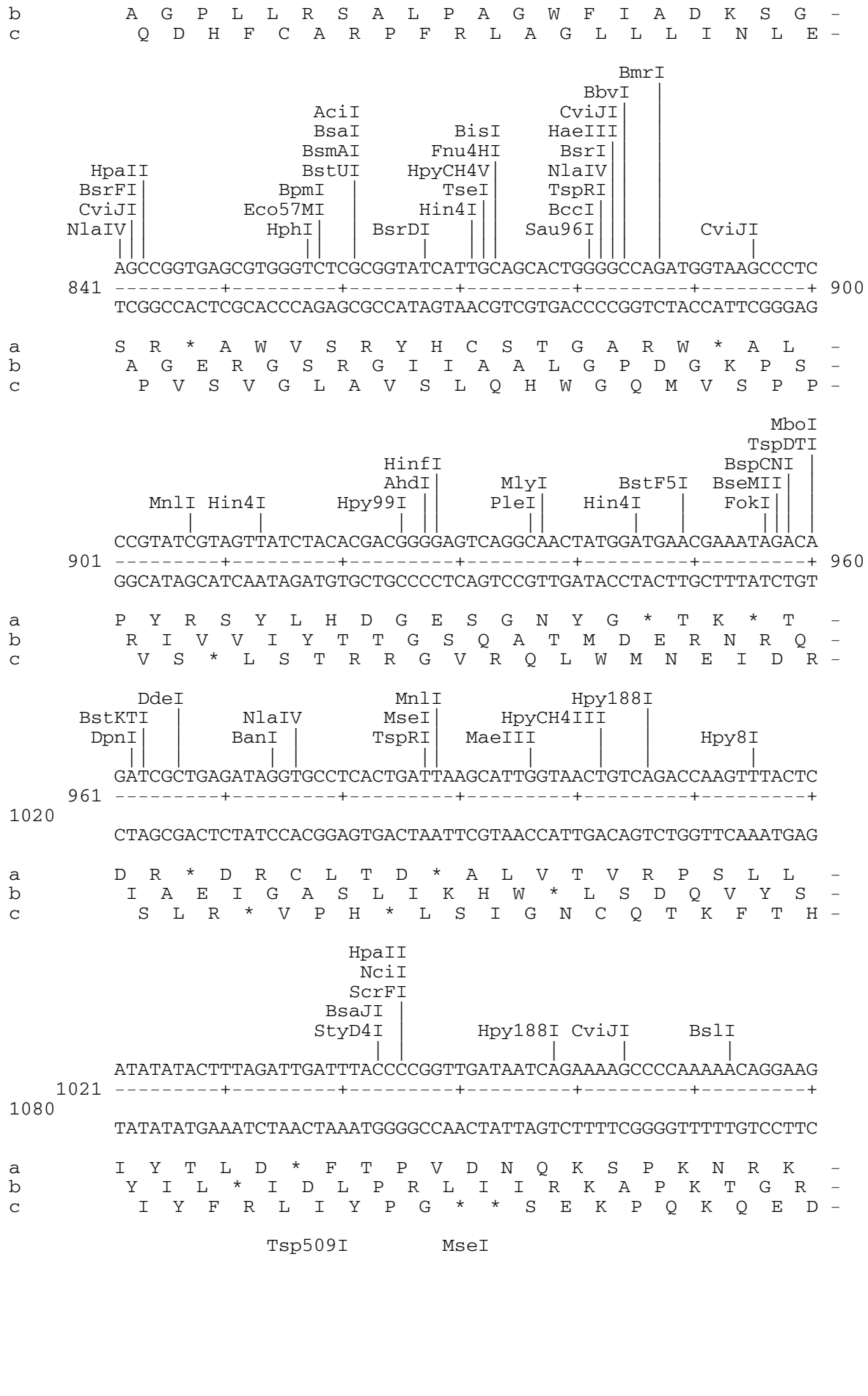
June 1, 2005 15:40 ..

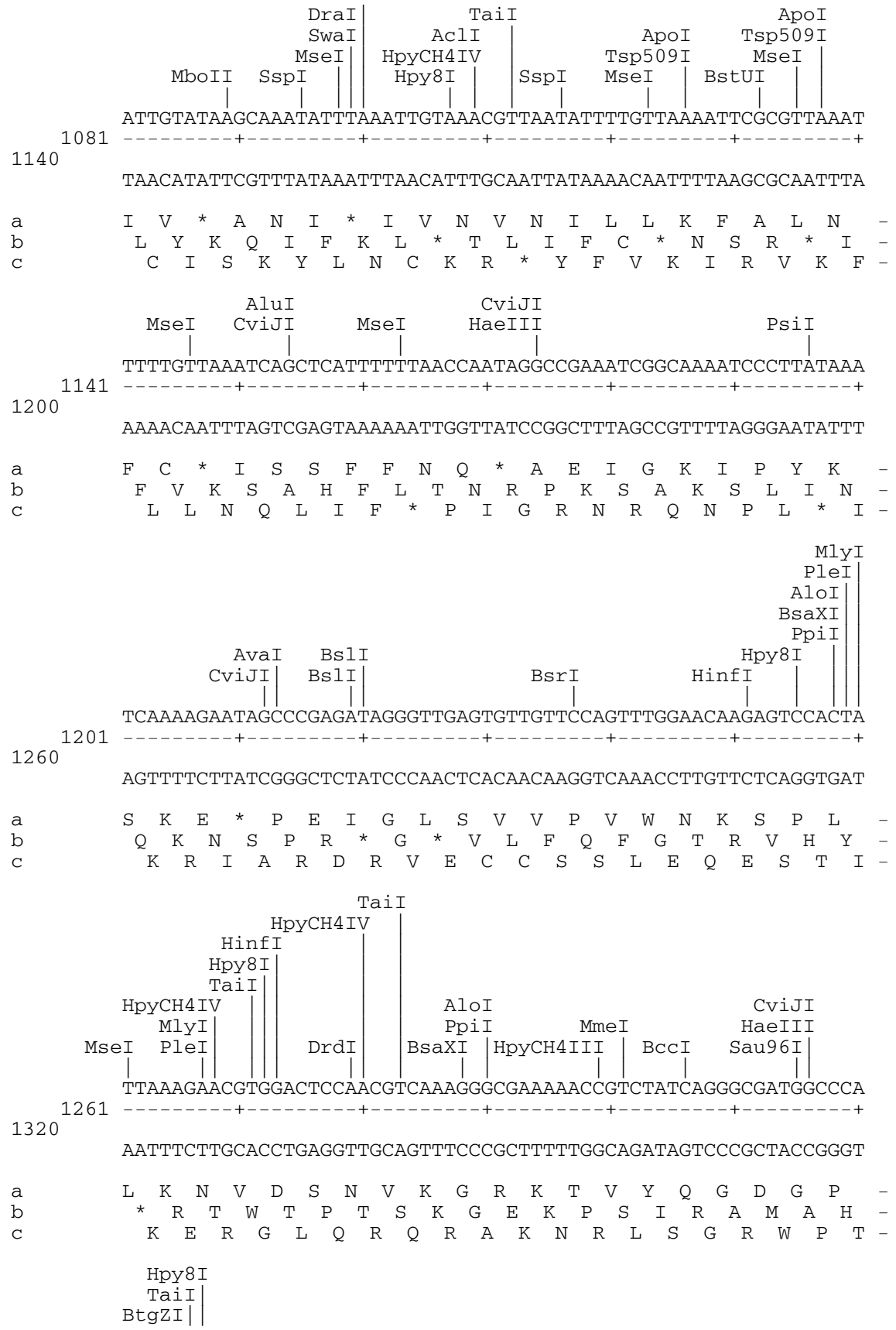
```
                  BccI
               BccI  |
          Hpy188III|  |
         BstF5I   ||  |  BstF5I
         CviJI|   ||  |  CviJI|   FokI
        HaeIII|   ||  | HaeIII| TspGWI
             ||   ||  |      ||      |
        AGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTTGTTTATTTTT
      1 ---------+---------+---------+---------+---------+---------+ 60
        TCTTCCGGTAGGACTGCCTACCGGAAAAACGCAAAGATGTTTGAGAAAAACAAATAAAAA

a       R  R  P  S  *  R  M  A  F  L  R  F  Y  K  L  F  L  F  I  F   -
b        E  G  H  P  D  G  W  P  F  C  V  S  T  N  S  F  C  L  F  F  -
c         K  A  I  L  T  D  G  L  F  A  F  L  Q  T  L  F  V  Y  F  S -

                                 NlaIII
                                 BciVI|
                              CviAII ||
                           Hpy188III ||
                              BspHI| ||
                               FatI| ||
                            BsmAI || ||
                            BsrBI || ||
                           AciI | || ||
                              | | || ||
        CTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATA
     61 ---------+---------+---------+---------+---------+---------+ 120
        GATTTATGTAAGTTTATACATAGGCGAGTACTCTGTTATTGGGACTATTTACGAAGTTAT

a       L  N  T  F  K  Y  V  S  A  H  E  T  I  T  L  I  N  A  S  I   -
b        *  I  H  S  N  M  Y  P  L  M  R  Q  *  P  *  *  M  L  Q  *  -
```

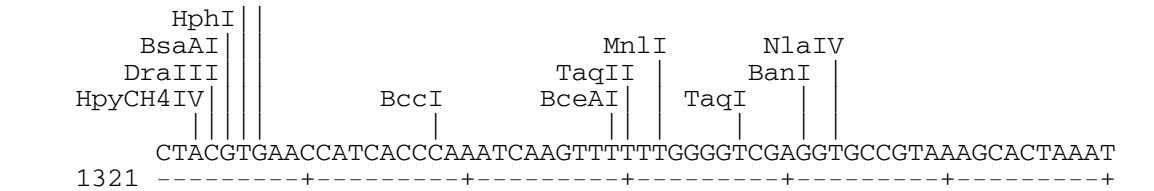

c K Y I Q I C I R S * D N N P D K C F N N -
TspGWI
SspI EarI MboII MslI
ATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTT
121 180
TATAACTTTTTCCTTCTCATACTCATAAGTTGTAAAGGCACAGCGGGAATAAGGGAAAAA
a I L K K E E Y E Y S T F P C R P Y S L F -
b Y * K R K S M S I Q H F R V A L I P F F -
c I E K G R V * V F N I S V S P L F P F L -
TauI
BisI
Fnu4HI
AciI HphI SfaNI TaqII HphI
TGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGC
181 240
ACGCCGTAAAACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTCATTTTCTACG
a C G I L P S C F C S P R N A G E S K R C -
b A A F C L P V F A H P E T L V K V K D A -
c R H F A F L F L L T Q K R W * K * K M L -
PpiI
MaeIII
AcuI
Eco57MI
Bme1580I
BsiHKAI
Bsp1286I
BssSI BstKTI BstKTI
BstKTI Hpy8I BsrI AlwI DpnI
DpnI HpyCH4V DpnI AciI PpiI
MboI ApaLI BstYI MspA1I BstYI
TaqII MboII TaqI MboI AlwI MboI
TGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGAT
241 300
ACTTCTAGTCAACCCACGTGCTCACCCAATGTAGCTTGACCTAGAGTTGTCGCCATTCTA
a * R S V G C T S G L H R T G S Q Q R * D -
b E D Q L G A R V G Y I E L D L N S G K I -
c K I S W V H E W V T S N W I S T A V R S -
MboII
TaiI DraI
AclI MseI
HpyCH4IV BsiHKAI
SmlI BpuEI Bsp1286I
CCTTGAGAGTTTTCGCCCCGAAGAACGTTCTCCAATGATGAGCACTTTTAAAGTTCTGCT
301 360
GGAACTCTCAAAAGCGGGGCTTCTTGCAAGAGGTTACTACTCGTGAAAATTTCAAGACGA
a P * E F S P R R T F S N D E H F * S S A -
b L E S F R P E E R S P M M S T F K V L L -
c L R V F A P K N V L Q * * A L L K F C Y -
HgaI
TaqII
NciI


ScrFI
HpaII
BcgI
TauI
AciI
StyD4I
AciI
BstUI
BsaHI
BisI
HhaI
HincII
Fnu4HI
HinP1I
Hpy8I
BsiEI
ATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACA
361
TACACCGCGCCATAATAGGGCACAACTGCGGCCCGTTCTCGTTGAGCCAGCGGCGTATGT
420
a M W R G I I P C * R R A R A T R S P H T -
b C G A V L S R V D A G Q E Q L G R R I H -
c V A R Y Y P V L T P G K S N S V A A Y T -
RsaI
ScaI
Csp6I
TatI
Hpy188I
HphI
MaeIII
DdeI
BseMII
Tsp45I
SfaNI
BcgI
BspCNI
BsrI
BccI
FatI
CTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGG
421
GATAAGAGTCTTACTGAACCAACTCATGAGTGGTCAGTGTCTTTTCGTAGAATGCCTACC
480
a L F S E * L G * V L T S H R K A S Y G W -
b Y S Q N D L V E Y S P V T E K H L T D G -
c I L R M T W L S T H Q S Q K S I L R M A -
CviJI
HaeIII
Tsp509I
TauI
FokI
TspRI
TspGWI
BtsI
TspRI
HpyCH4III
NlaIII
BisI
NlaIII
CviAII
EaeI
BstF5I
Fnu4HI
FatI
Fnu4HI
CviAII
BbvI
HpyCH4V
TseI
MslI
BtsI
AciI
CATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAA
481
GTACTGTCATTCTCTTAATACGTCACGACGGTATTGGTACTCACTATTGTGACGCCGGTT
540
a H D S K R I M Q C C H N H E * * H C G Q -
b M T V R E L C S A A I T M S D N T A A N -
c * Q * E N Y A V L P * P * V I T L R P T -
AvaII
Sau96I
Hpy188I
BsiEI
BstKTI
PvuI
DpnI
NlaIII
MboI
AluI
TaqII
CviAII
Hpy188I
MnlI
CviJI
AciI
HpyCH4V
FatI
CTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGG
541
GAATGAAGACTGTTGCTAGCCTCCTGGCTTCCTCGATTGGCGAAAAAACGTGTTGTACCC
600
a L T S D N D R R T E G A N R F F A Q H G -

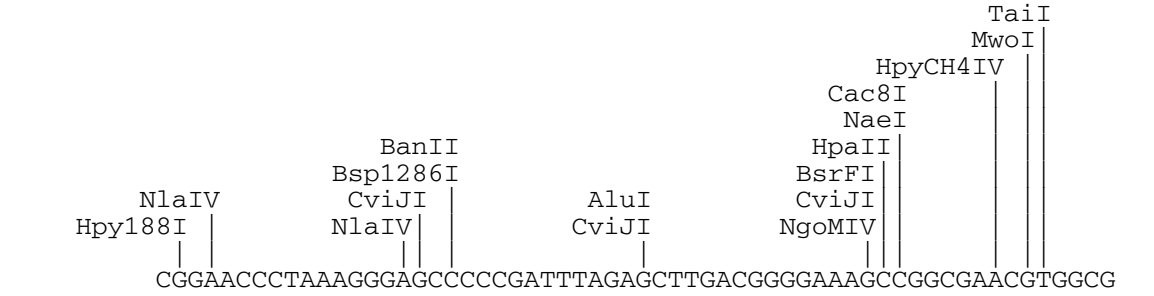

b L L L T T I G G P K E L T A F L H N M G -
c Y F * Q R S E D R R S * P L F C T T W G -
AlwI
NlaIII
MaeIII
CviAII
BstKTI
FatI
DpnI
MboI
BstKTI
DpnI
MboI
HpaII
BsaWI
AluI
NlaIV
CviJI
CviJI
TspDTI
GGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGA
601 ---------+---------+---------+---------+---------+---------+ 660
CCTAGTACATTGAGCGGAACTAGCAACCCTTGGCCTCGACTTACTTCGGTATGGTTTGCT
a G S C N S P * S L G T G A E * S H T K R -
b D H V T R L D R W E P E L N E A I P N D -
c I M * L A L I V G N R S * M K P Y Q T T -
HhaI
FspI
HinP1I
TaiI
SfaNI
MaeIII
Tsp45I
Hpy99I
AclI
HpyCH4IV
MslI
SfcI
BsrDI
MseI
CGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGG
661 ---------+---------+---------+---------+---------+---------+ 720
GCTCGCACTGTGGTGCTACGGACATCGTTACCGTTGTTGCAACGCGTTTGATAATTGACC
a R A * H H D A C S N G N N V A Q T I N W -
b E R D T T M P V A M A T T L R K L L T G -
c S V T P R C L * Q W Q Q R C A N Y * L A -
HpaII
NciI
ScrFI
StyD4I
AluI
CviJI
AseI
MseI
AciI
MnlI
BstF5I
BsrI
BfaI
Tsp509I
BccI
BsrI
FokI
CGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGT
721 ---------+---------+---------+---------+---------+---------+ 780
GCTTGATGAATGAGATCGAAGGGCCGTTGTTAATTATCTGACCTACCTCCGCCTATTTCA
a R T T Y S S F P A T I N R L D G G G * S -
b E L L T L A S R Q Q L I D W M E A D K V -
c N Y L L * L P G N N * * T G W R R I K L -
CviJI
CviJI
Cac8I
AvaII
HaeIII
CviJI
EciI
Sau96I
BglI
Sau96I
HhaI
HpaII
Hin4I
HpyCH4V
HinP1I
MwoI
Hpy188III
TGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGG
781 ---------+---------+---------+---------+---------+---------+ 840
ACGTCCTGGTGAAGACGCGAGCCGGGAAGGCCGACCGACCAAATAACGACTATTTAGACC
a C R T T S A L G P S G W L V Y C * * I W -


b A G P L L R S A L P A G W F I A D K S G -
c Q D H F C A R P F R L A G L L L I N L E -
BmrI
BbvI
AciI
CviJI
BsaI
BisI
HaeIII
BsmAI
Fnu4HI
BsrI
HpaII
BstUI
HpyCH4V
NlaIV
BsrFI
BpmI
TseI
TspRI
CviJI
Eco57MI
Hin4I
BccI
NlaIV
HphI
BsrDI
Sau96I
CviJI
AGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTC
841 ---------+---------+---------+---------+---------+---------+ 900
TCGGCCACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAG
a S R * A W V S R Y H C S T G A R W * A L -
b A G E R G S R G I I A A L G P D G K P S -
c P V S V G L A V S L Q H W G Q M V S P P -
MboI
TspDTI
HinfI
BspCNI
AhdI
MlyI
BstF5I
BseMII
MnlI Hin4I
Hpy99I
PleI
Hin4I
FokI
CCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACA
901 ---------+---------+---------+---------+---------+---------+ 960
GGCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGT
a P Y R S Y L H D G E S G N Y G * T K * T -
b R I V V I Y T T G S Q A T M D E R N R Q -
c V S * L S T R R G V R Q L W M N E I D R -
DdeI
MnlI
Hpy188I
BstKTI
NlaIV
MseI
HpyCH4IIII
DpnI
BanI
TspRI
MaeIII
Hpy8I
GATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTC
961 ---------+---------+---------+---------+---------+---------+
1020
CTAGCGACTCTATCCACGGAGTGACTAATTCGTAACCATTGACAGTCTGGTTCAAATGAG
a D R * D R C L T D * A L V T V R P S L L -
b I A E I G A S L I K H W * L S D Q V Y S -
c S L R * V P H * L S I G N C Q T K F T H -
HpaII
NciI
ScrFI
BsaJI
StyD4I
Hpy188I CviJI
BslI
ATATATACTTTAGATTGATTTACCCCGGTTGATAATCAGAAAAGCCCCAAAAACAGGAAG
1021 ---------+---------+---------+---------+---------+---------+
1080
TATATATGAAATCTAACTAAATGGGGCCAACTATTAGTCTTTTCGGGGTTTTTGTCCTTC
a I Y T L D * F T P V D N Q K S P K N R K -
b Y I L * I D L P R L I I R K A P K T G R -
c I Y F R L I Y P G * * S E K P Q K Q E D -
Tsp509I
MseI

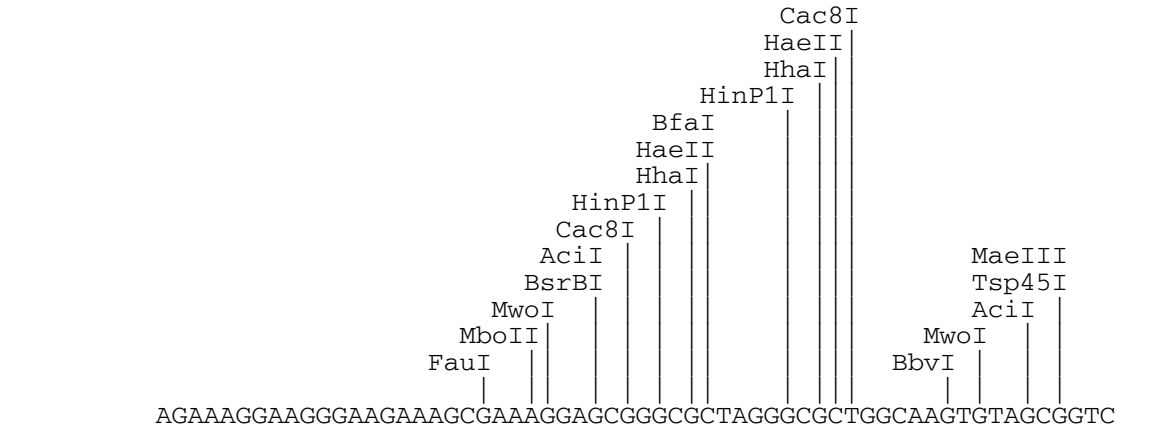

DraI
SwaI
MseI
TaiI
AclI
HpyCH4IV
ApoI
Tsp509I
MseI
MboII
SspI
Hpy8I
SspI
MseI
BstUI
ATTGTATAAGCAAATATTTAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAAT
1081
1140
TAACATATTCGTTTATAAATTTAACATTTGCAATTATAAAACAATTTTAAGCGCAATTTA
a I V * A N I * I V N V N I L L K F A L N -
b L Y K Q I F K L * T L I F C * N S R * I -
c C I S K Y L N C K R * Y F V K I R V K F -
AluI
CviJI
MseI
CviJI
MseI
HaeIII
PsiI
TTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAA
1141
1200
AAAACAATTTAGTCGAGTAAAAAATTGGTTATCCGGCTTTAGCCGTTTTAGGGAATATTT
a F C * I S S F F N Q * A E I G K I P Y K -
b F V K S A H F L T N R P K S A K S L I N -
c L L N Q L I F * P I G R N R Q N P L * I -
MlyI
PleI
AloI
BsaXI
PpiI
AvaI
CviJI
BslI
BslI
BsrI
Hpy8I
HinfI
TCAAAAGAATAGCCCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTA
1201
1260
AGTTTTCTTATCGGGCTCTATCCCAACTCACAACAAGGTCAAACCTTGTTCTCAGGTGAT
a S K E * P E I G L S V V P V W N K S P L -
b Q K N S P R * G * V L F Q F G T R V H Y -
c K R I A R D R V E C C S S L E Q E S T I -
TaiI
HpyCH4IV
HinfI
Hpy8I
TaiI
HpyCH4IV
MlyI
MseI
PleI
DrdI
AloI
PpiI
BsaXI
MmeI
HpyCH4III
BccI
CviJI
HaeIII
Sau96I
TTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCA
1261
1320
AATTTCTTGCACCTGAGGTTGCAGTTTCCCGCTTTTTGGCAGATAGTCCCGCTACCGGGT
a L K N V D S N V K G R K T V Y Q G D G P -
b * R T W T P T S K G E K P S I R A M A H -
c K E R G L Q R Q R A K N R L S G R W P T -
Hpy8I
TaiI
BtgZI



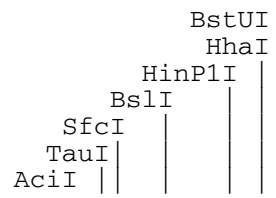
a L R E P S P K S S F L G S R C R K A L N -
b Y V N H H P N Q V F W G R G A V K H * I -
c T * T I T Q I K F F G V E V P * S T K S -

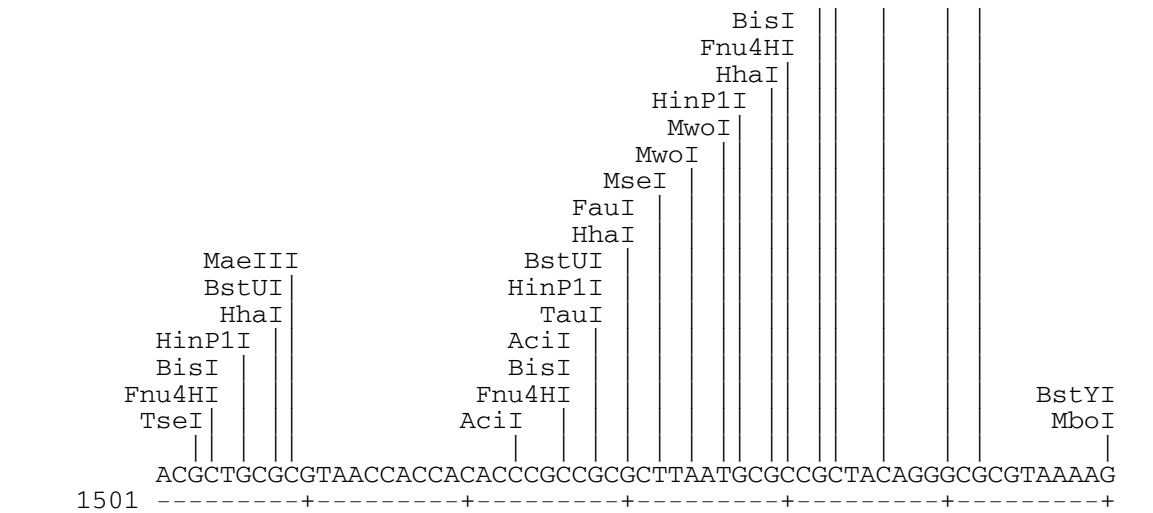


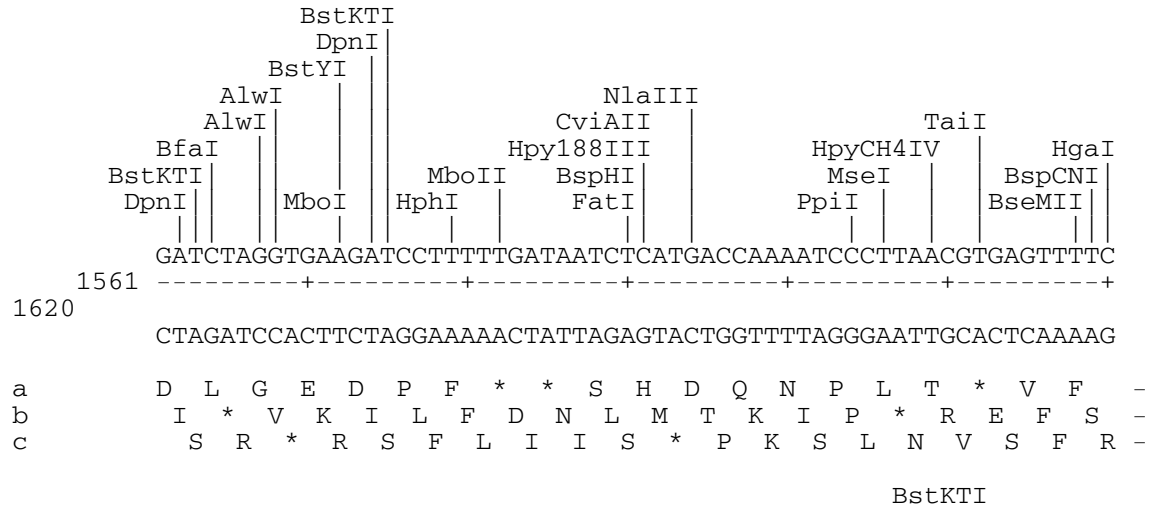
a R N P K G S P R F R A * R G K P A N V A -
b G T L K G A P D L E L D G E S R R T W R -
c E P * R E P P I * S L T G K A G E R G E -

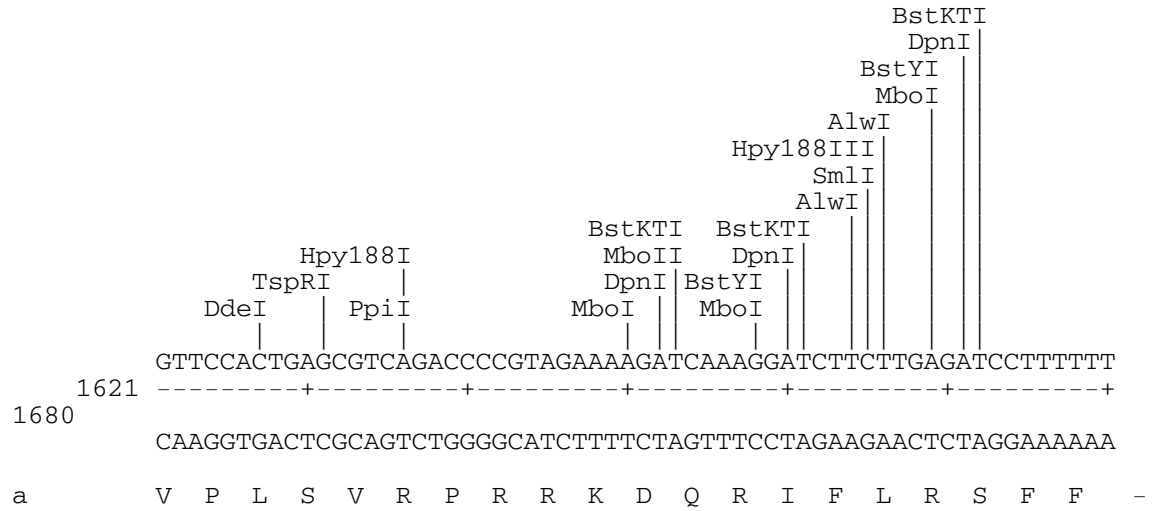


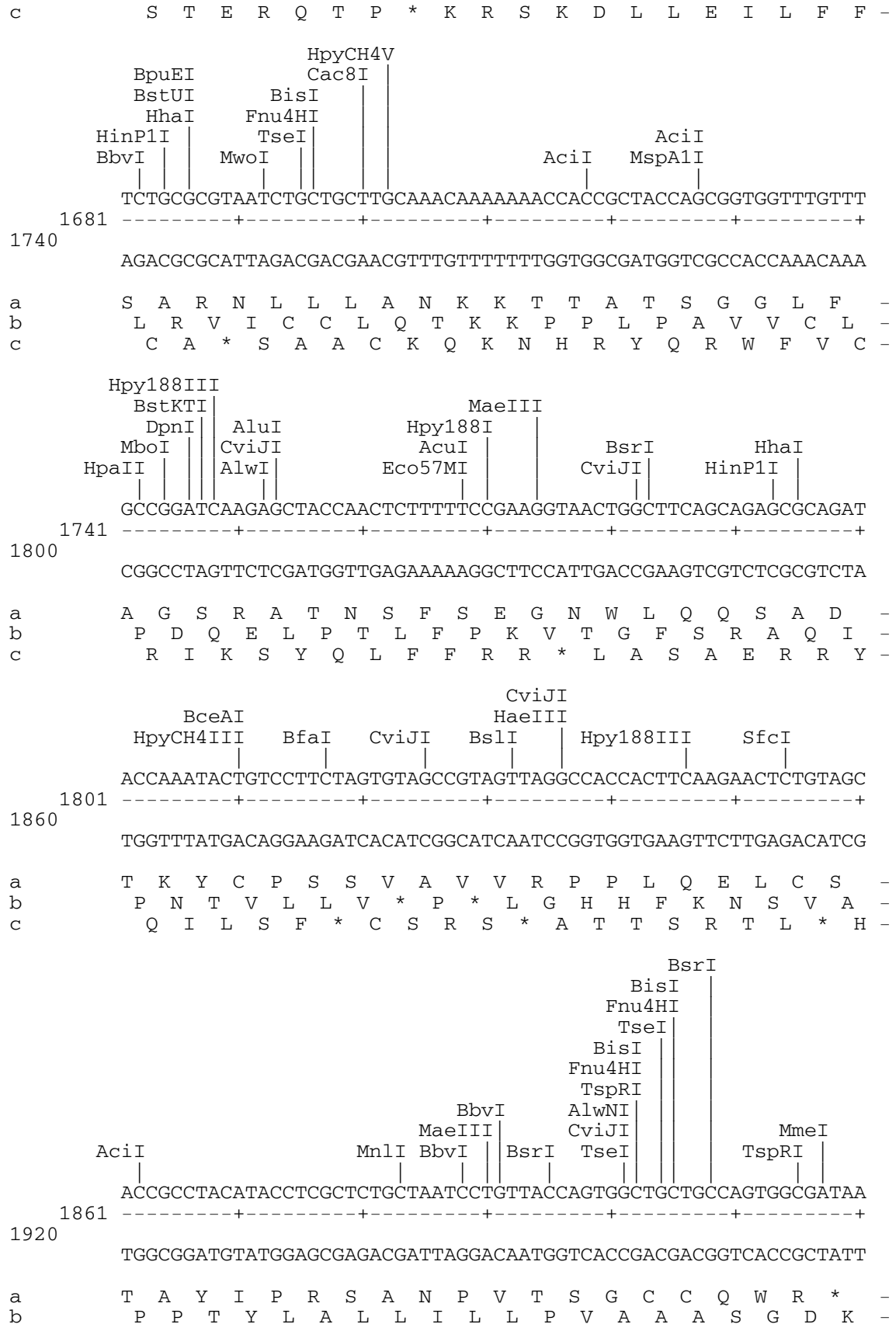
a R K E G K K A K G A G A R A L A S V A V -
b E R K G R K R K E R A L G R W Q V * R S -
c K G R E E S E R S G R * G A G K C S G H -

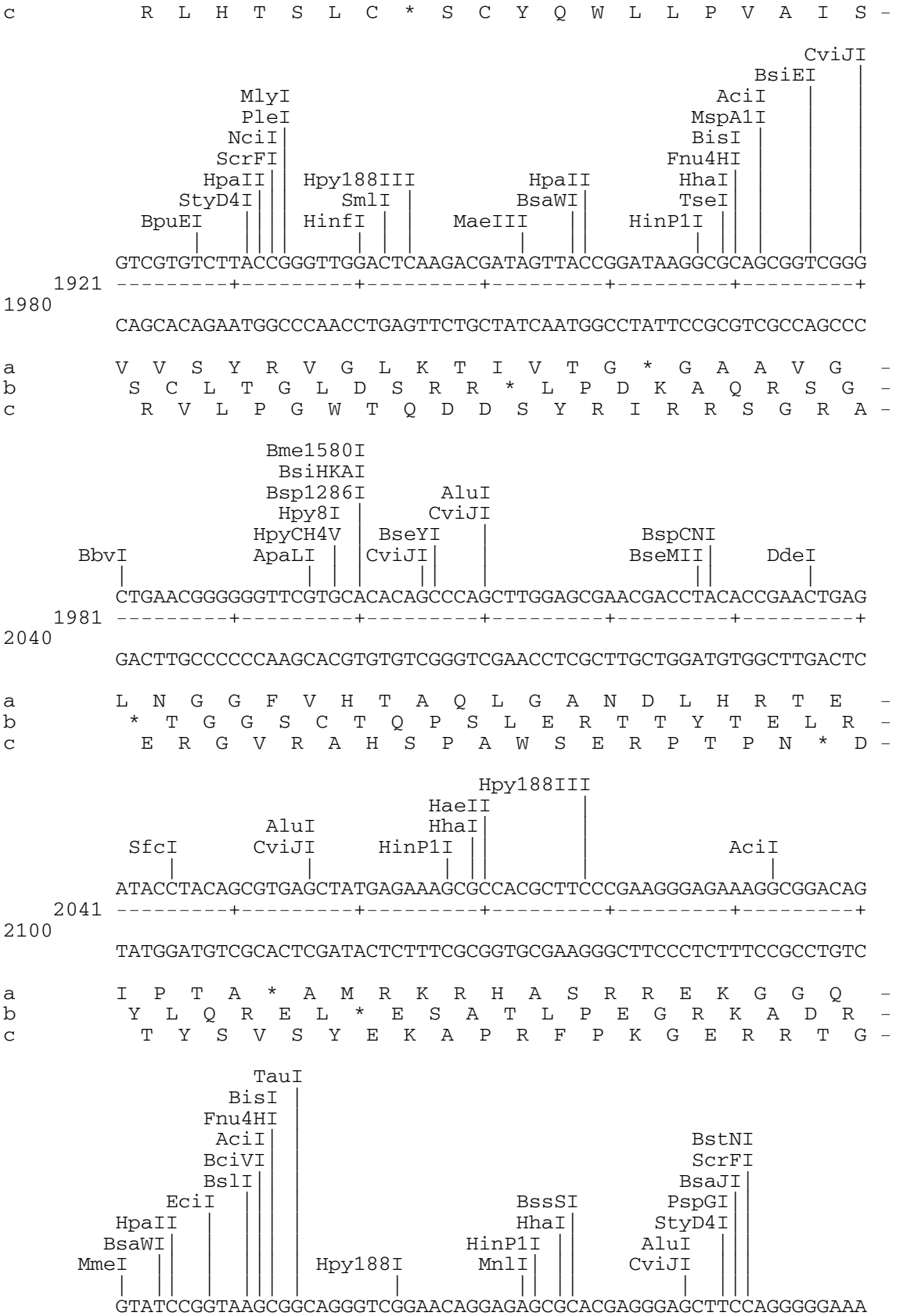


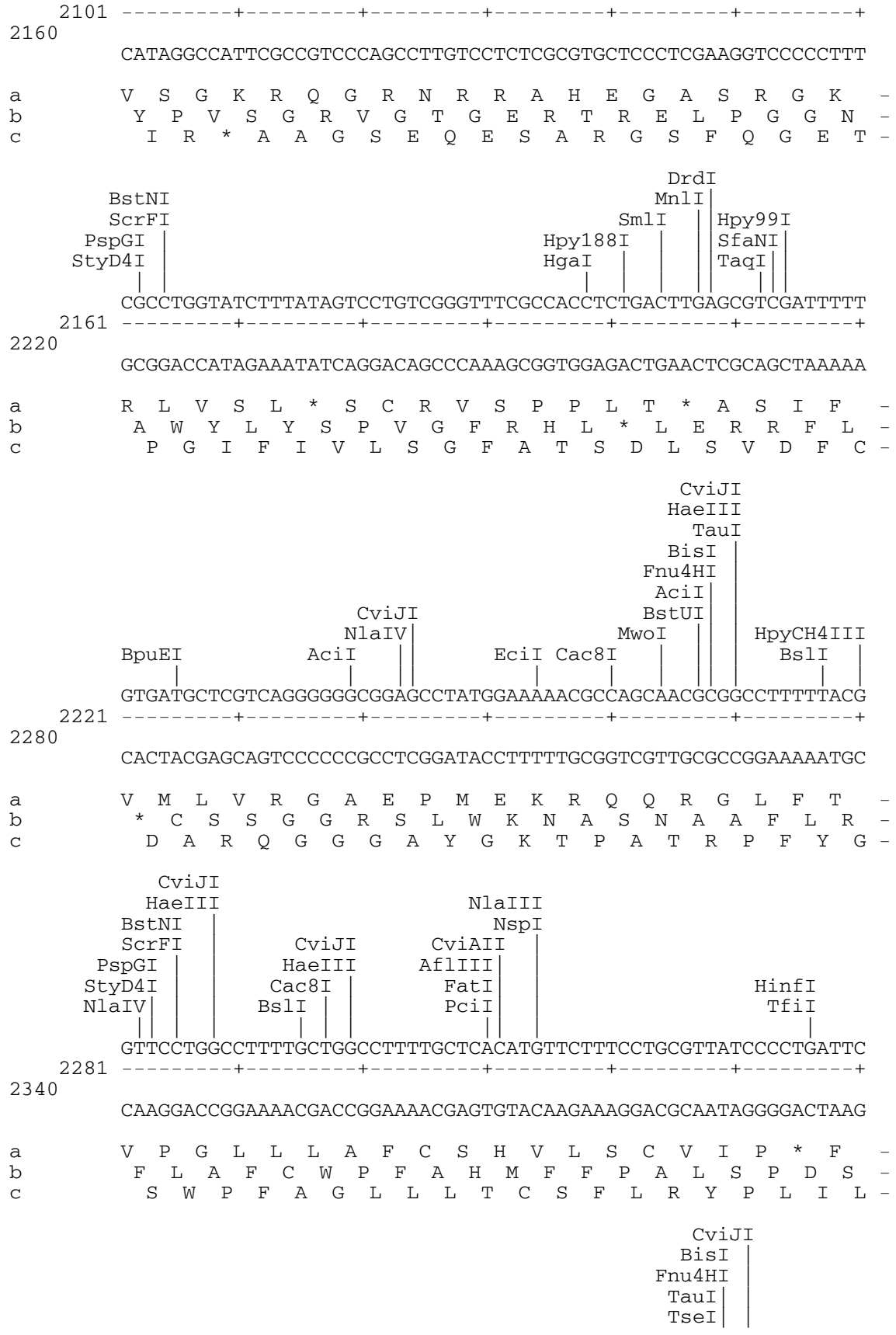

BisI
Fnu4HI
HhaI
HinP1I
MwoI
MwoI
MseI
FauI
HhaI
MaeIII
BstUI
HhaI
HinP1I
BisI
Fnu4HI
TseI
BstUI
HinP1I
TauI
AciI
BisI
Fnu4HI
AciI
BstYI
MboI
ACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTAAAAG
1501 ---------+---------+---------+---------+---------+---------+
1560
TGCGACGCGCATTGGTGGTGTGGGCGGCGCGAATTACGCGGCGATGTCCCGCGCATTTTC
a T L R V T T T P A A L N A P L Q G A * K -
b R C A * P P H P P R L M R R Y R A R K R -
c A A R N H H T R R A * C A A T G R V K G -

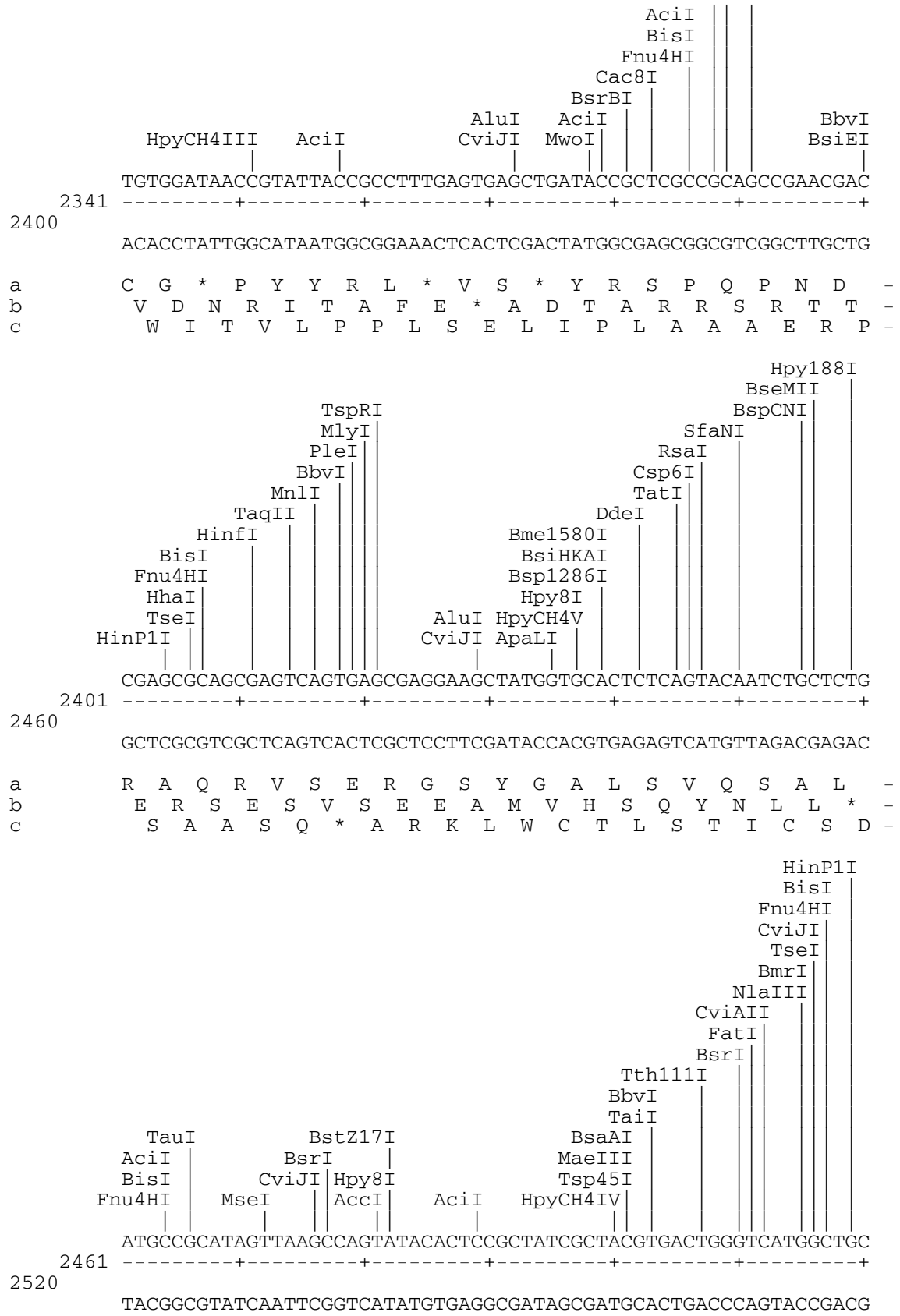

BstKTI
DpnI
BstYI
AlwI
NlaIII
AlwI
CviAII
TaiI
BfaI
Hpy188III
HpyCH4IV
HgaI
BstKTI
MboII
BspHI
MseI
BspCNI
DpnI
MboI
HphI
FatI
PpiI
BseMII
GATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTC
1561 ---------+---------+---------+---------+---------+---------+
1620
CTAGATCCACTTCTAGGAAAAACTATTAGAGTACTGGTTTTAGGGAATTGCACTCAAAAG
a D L G E D P F * * S H D Q N P L T * V F -
b I * V K I L F D N L M T K I P * R E F S -
c S R * R S F L I I S * P K S L N V S F R -

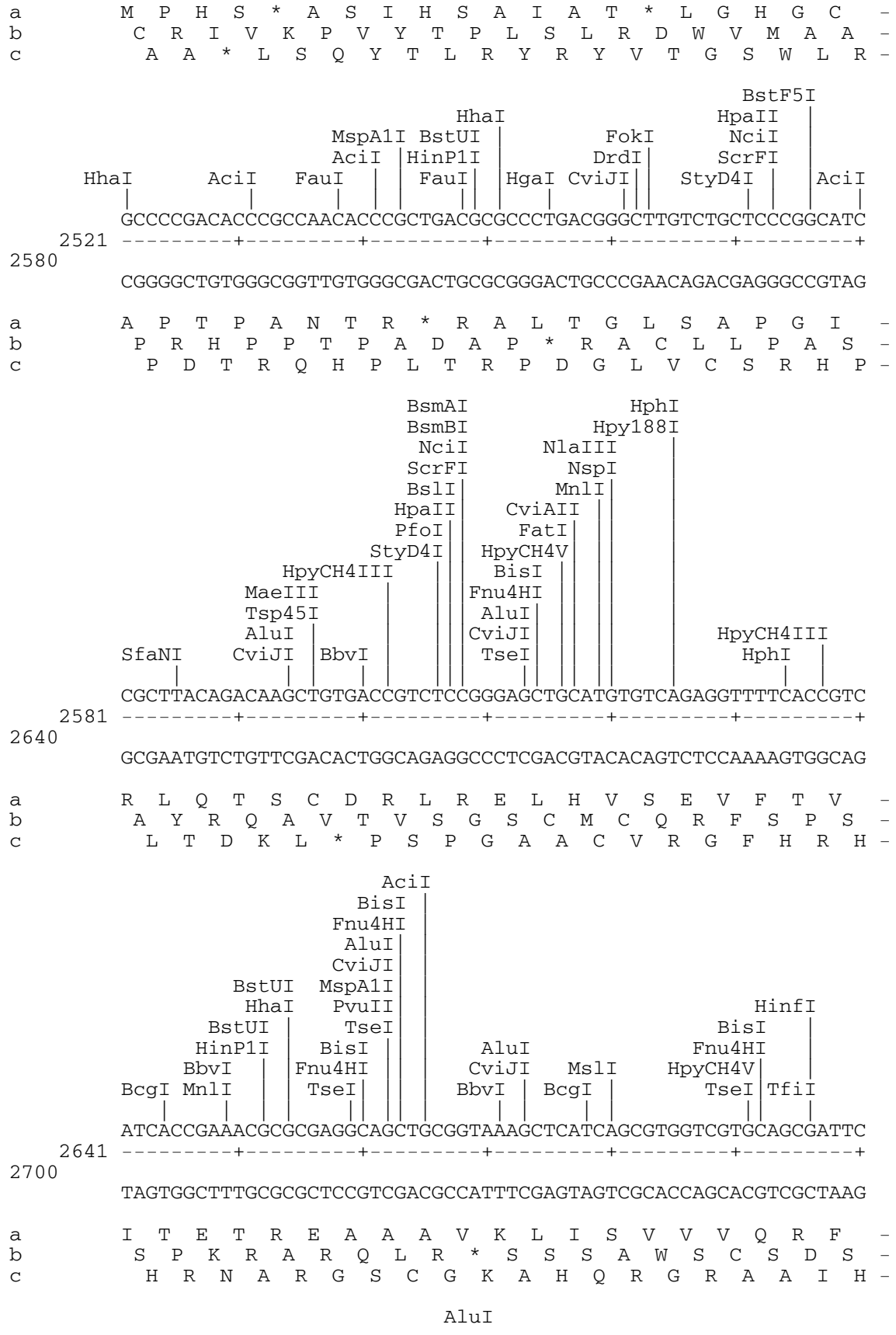

BstKTI
DpnI
BstYI
MboI
AlwI
Hpy188III
SmlI
AlwI
BstKTI
BstKTI
Hpy188I
MboII
DpnI
TspRI
DpnI
BstYI
DdeI
PpiI
MboI
MboI
GTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTT
1621 ---------+---------+---------+---------+---------+---------+
1680
CAAGGTGACTCGCAGTCTGGGGCATCTTTTCTAGTTTCCTAGAAGAACTCTAGGAAAAAA
a V P L S V R P R R K D Q R I F L R S F F -
b F H * A S D P V E K I K G S S * D P F F -

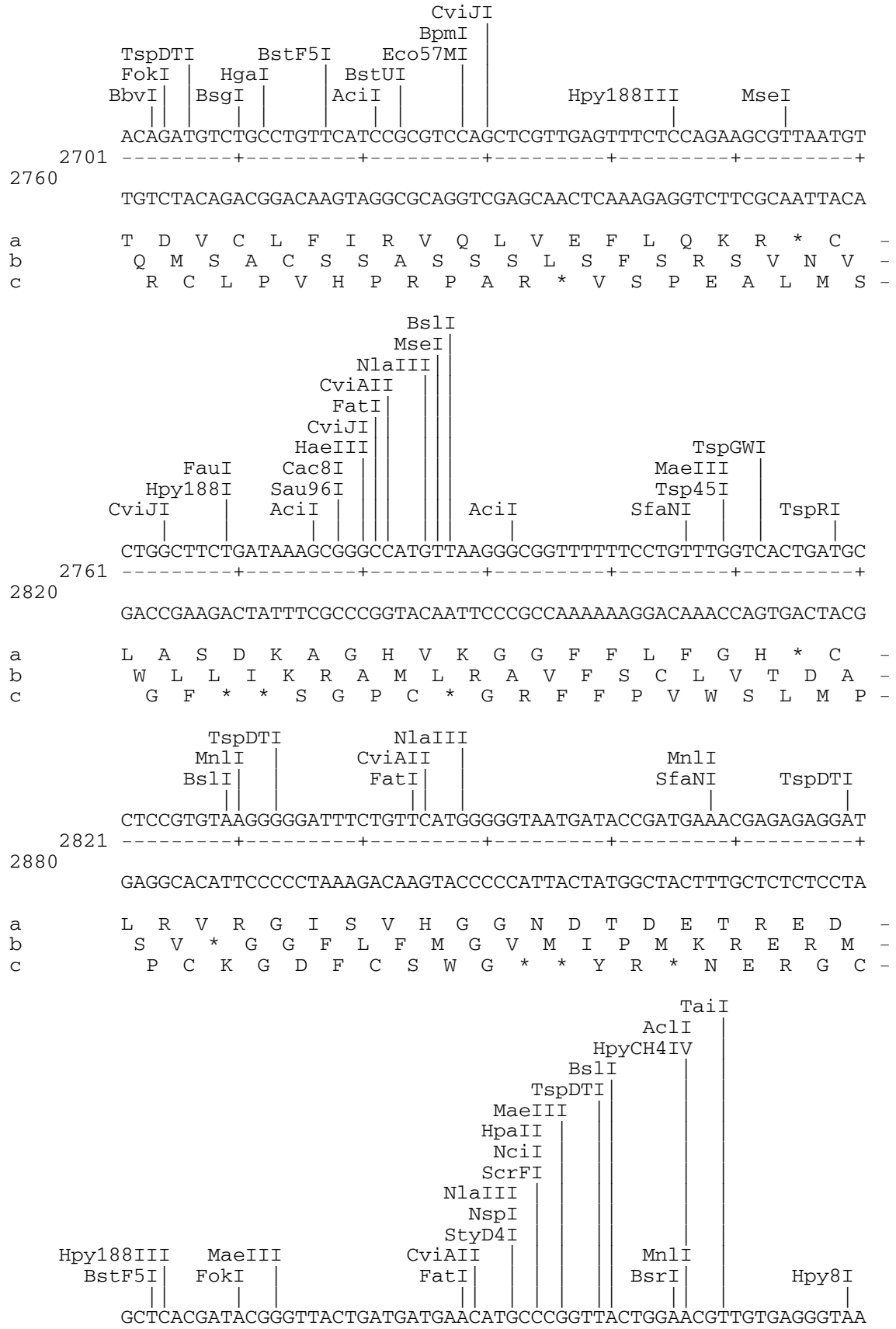

c S T E R Q T P * K R S K D L L E I L F F -
HpyCH4V
Cac8I
BpuEI
BstUI
BisI
HhaI
Fnu4HI
HinP1I
TseI
BbvI
MwoI
AciI
AciI
MspA1I
TCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTT
1681
1740
AGACGCGCATTAGACGACGAACGTTTGTTTTTTTGGTGGCGATGGTCGCCACCAAACAAA
a S A R N L L L A N K K T T A T S G G L F -
b L R V I C C L Q T K K P P L P A V V C L -
c C A * S A A C K Q K N H R Y Q R W F V C -
Hpy188III
BstKTI
MaeIII
DpnI
AluI
Hpy188I
MboI
CviJI
AcuI
BsrI
HhaI
HpaII
AlwI
Eco57MI
CviJI
HinP1I
GCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGAT
1741
1800
CGGCCTAGTTCTCGATGGTTGAGAAAAAGGCTTCCATTGACCGAAGTCGTCTCGCGTCTA
a A G S R A T N S F S E G N W L Q Q S A D -
b P D Q E L P T L F P K V T G F S R A Q I -
c R I K S Y Q L F F R R * L A S A E R R Y -
CviJI
HaeIII
BceAI
HpyCH4III
BfaI
CviJI
BslI
Hpy188III
SfcI
ACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGC
1801
1860
TGGTTTATGACAGGAAGATCACATCGGCATCAATCCGGTGGTGAAGTTCTTGAGACATCG
a T K Y C P S S V A V V R P P L Q E L C S -
b P N T V L L V * P * L G H H F K N S V A -
c Q I L S F * C S R S * A T T S R T L * H -
BsrI
BisI
Fnu4HI
TseI
BisI
Fnu4HI
TspRI
BbvI
AlwNI
MaeIII
CviJI
MmeI
AciI
MnlI
BbvI
BsrI
TseI
TspRI
ACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAA
1861
1920
TGGCGGATGTATGGAGCGAGACGATTAGGACAATGGTCACCGACGACGGTCACCGCTATT
a T A Y I P R S A N P V T S G C C Q W R * -
b P P T Y L A L L I L L P V A A A S G D K -

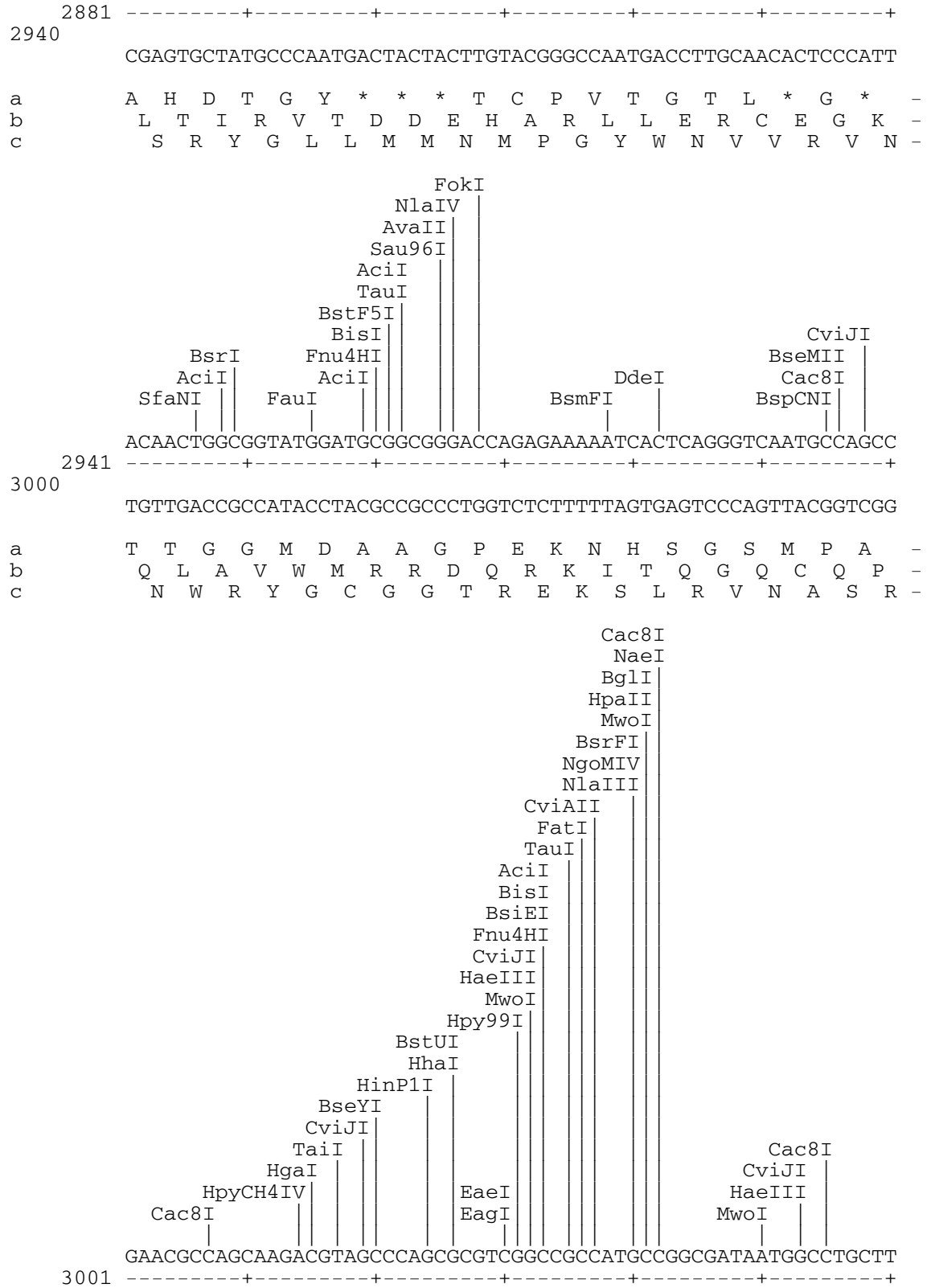

c R L H T S L C * S C Y Q W L L P V A I S -
CviJI
BsiEI
MlyI
AciI
PleI
MspA1I
NciI
BisI
ScrFI
Fnu4HI
HpaII
Hpy188III
HpaII
HhaI
StyD4I
SmlI
BsaWI
TseI
BpuEI
HinfI
MaeIII
HinP1I
GTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGG
1921
1980
CAGCACAGAATGGCCCAACCTGAGTTCTGCTATCAATGGCCTATTCCGCGTCGCCAGCCC
a V V S Y R V G L K T I V T G * G A A V G -
b S C L T G L D S R R * L P D K A Q R S G -
c R V L P G W T Q D D S Y R I R R S G R A -
Bme1580I
BsiHKAI
Bsp1286I
AluI
Hpy8I
CviJI
HpyCH4V
BseYI
BspCNI
BbvI
ApaLI
CviJI
BseMII
DdeI
CTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAG
1981
2040
GACTTGCCCCCCAAGCACGTGTGTCGGGTCGAACCTCGCTTGCTGGATGTGGCTTGACTC
a L N G G F V H T A Q L G A N D L H R T E -
b * T G G S C T Q P S L E R T T Y T E L R -
c E R G V R A H S P A W S E R P T P N * D -
Hpy188III
HaeII
AluI
HhaI
SfcI
CviJI
HinP1I
AciI
ATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAG
2041
2100
TATGGATGTCGCACTCGATACTCTTTCGCGGTGCGAAGGGCTTCCCTCTTTCCGCCTGTC
a I P T A * A M R K R H A S R R E K G G Q -
b Y L Q R E L * E S A T L P E G R K A D R -
c T Y S V S Y E K A P R F P K G E R R T G -
TauI
BisI
Fnu4HI
AciI
BstNI
BciVI
ScrFI
BslI
BsaJI
EciI
BssSI
PspGI
HpaII
HhaI
StyD4I
BsaWI
HinP1I
AluI
MmeI
Hpy188I
MnlI
CviJI
GTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAA

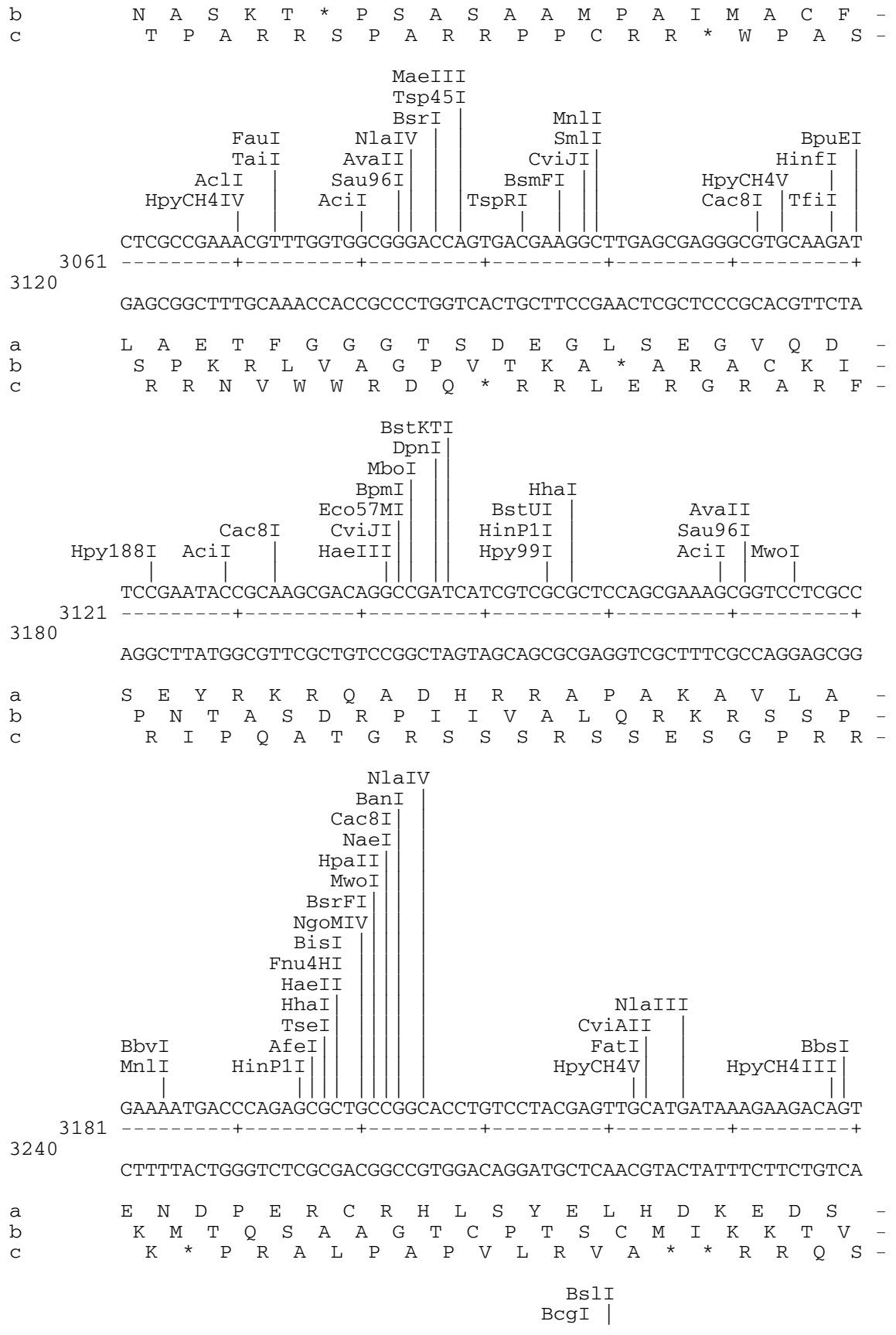

2101
2160
CATAGGCCATTCGCCGTCCCAGCCTTGTCCTCTCGCGTGCTCCCTCGAAGGTCCCCCTTT
a V S G K R Q G R N R R A H E G A S R G K -
b Y P V S G R V G T G E R T R E L P G G N -
c I R * A A G S E Q E S A R G S F Q G E T -
DrdI
BstNI MnlI
ScrFI SmlI Hpy99I
PspGI Hpy188I SfaNI
StyD4I HgaI TaqI
CGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTT
2161
2220
GCGGACCATAGAAATATCAGGACAGCCCAAAGCGGTGGAGACTGAACTCGCAGCTAAAAA
a R L V S L * S C R V S P P L T * A S I F -
b A W Y L Y S P V G F R H L * L E R R F L -
c P G I F I V L S G F A T S D L S V D F C -
CviJI
HaeIII
TauI
BisI
Fnu4HI
AciI
CviJI BstUI
NlaIV MwoI HpyCH4III
BpuEI AciI EciI Cac8I BslI
GTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACG
2221
2280
CACTACGAGCAGTCCCCCCGCCTCGGATACCTTTTTGCGGTCGTTGCGCCGGAAAAATGC
a V M L V R G A E P M E K R Q Q R G L F T -
b * C S S G G R S L W K N A S N A A F L R -
c D A R Q G G G A Y G K T P A T R P F Y G -
CviJI
HaeIII NlaIII
BstNI NspI
ScrFI CviJI CviAII
PspGI HaeIII AflIII
StyD4I Cac8I FatI HinfI
NlaIV BslI PciI TfiI
GTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTC
2281
2340
CAAGGACCGGAAAACGACCGGAAAACGAGTGTACAAGAAAGGACGCAATAGGGGACTAAG
a V P G L L L A F C S H V L S C V I P * F -
b F L A F C W P F A H M F F P A L S P D S -
c S W P F A G L L L T C S F L R Y P L I L -
CviJI
BisI
Fnu4HI
TauI
TseI

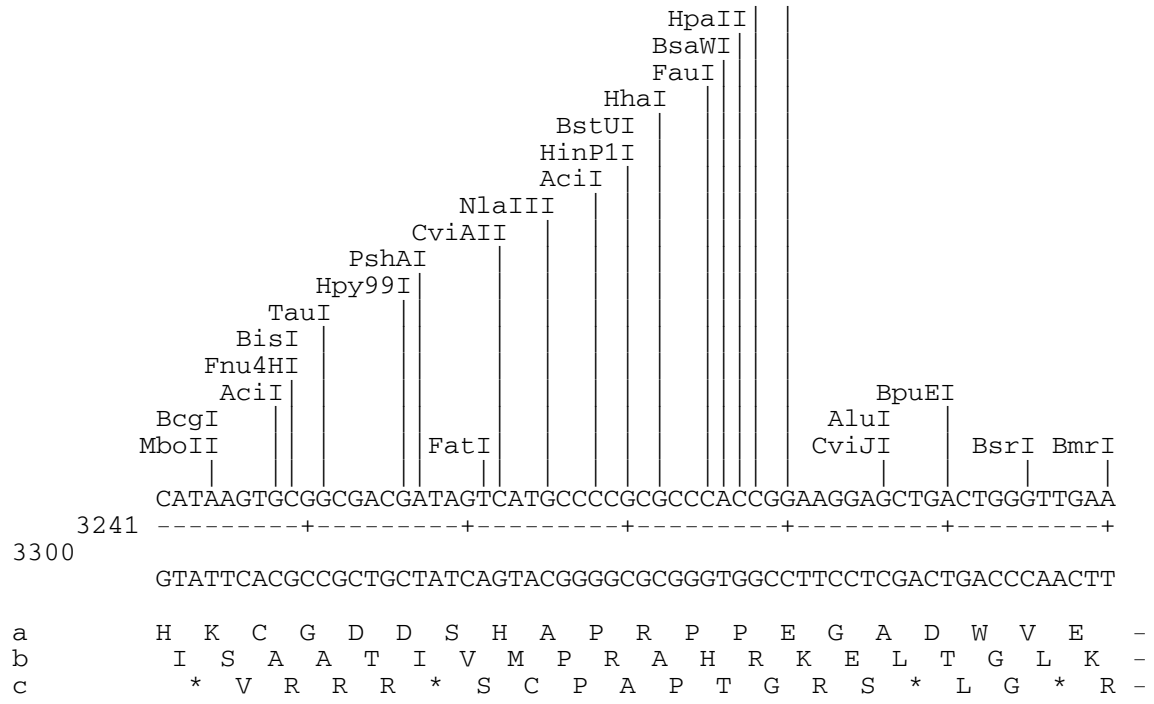

AciI
BisI
Fnu4HI
Cac8I
BsrBI
AluI
AciI
BbvI
HpyCH4III
AciI
CviJI
MwoI
BsiEI
TGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGAC
2341
2400
ACACCTATTGGCATAATGGCGGAAACTCACTCGACTATGGCGAGCGGCGTCGGCTTGCTG
a C G * P Y Y R L * V S * Y R S P Q P N D -
b V D N R I T A F E * A D T A R R S R T T -
c W I T V L P P L S E L I P L A A A E R P -
Hpy188I
BseMII
TspRI
BspCNI
MlyI
SfaNI
PleI
RsaI
BbvI
Csp6I
MnlI
TatI
TaqII
DdeI
HinfI
Bme1580I
BisI
BsiHKAI
Fnu4HI
Bsp1286I
HhaI
Hpy8I
TseI
AluI HpyCH4V
HinP1I
CviJI ApaLI
CGAGCGCAGCGAGTCAGTGAGCGAGGAAGCTATGGTGCACTCTCAGTACAATCTGCTCTG
2401
2460
GCTCGCGTCGCTCAGTCACTCGCTCCTTCGATACCACGTGAGAGTCATGTTAGACGAGAC
a R A Q R V S E R G S Y G A L S V Q S A L -
b E R S E S V S E E A M V H S Q Y N L L * -
c S A A S Q * A R K L W C T L S T I C S D -
HinP1I
BisI
Fnu4HI
CviJI
TseI
BmrI
NlaIII
CviAII
FatI
BsrI
Tth111I
BbvI
TaiI
TauI
BstZ17I
BsaAI
AciI
BsrI
MaeIII
BisI
CviJI
Hpy8I
Tsp45I
Fnu4HI
MseI
AccI
AciI
HpyCH4IV
ATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGC
2461
2520
TACGGCGTATCAATTCGGTCATATGTGAGGCGATAGCGATGCACTGACCCAGTACCGACG

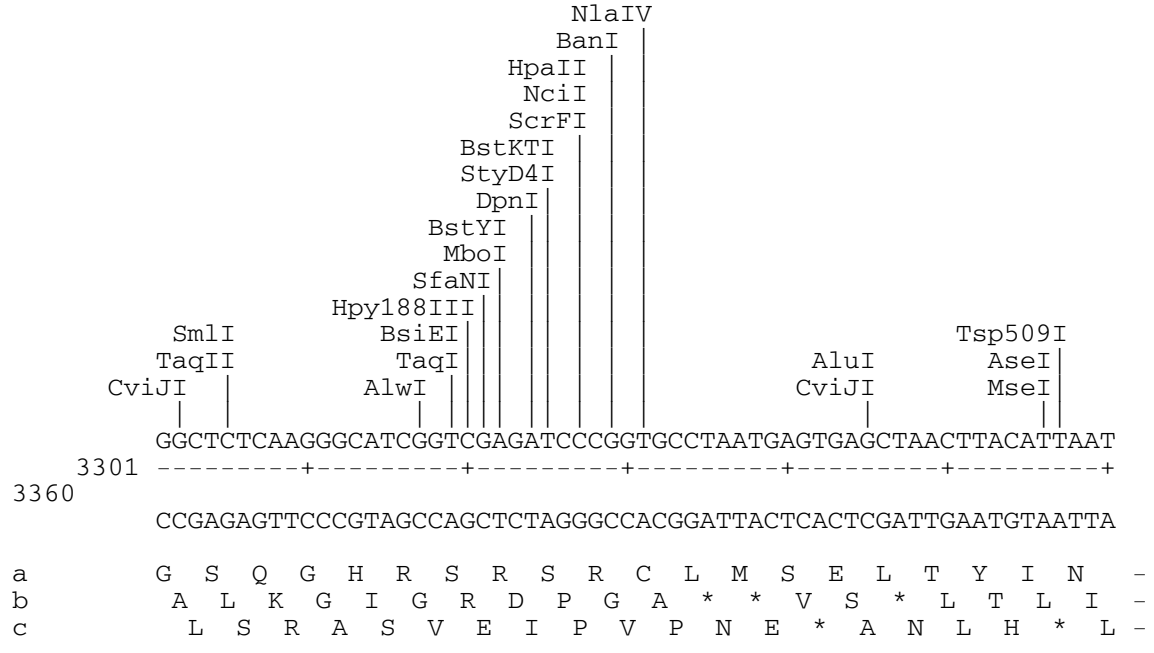

a M P H S * A S I H S A I A T * L G H G C -
b C R I V K P V Y T P L S L R D W V M A A -
c A A * L S Q Y T L R Y R Y V T G S W L R -
BstF5I
HhaI
HpaII
MspA1I BstUI
FokI
NciI
AciI HinP1I
DrdI
ScrFI
HhaI AciI FauI FauI HgaI CviJI StyD4I AciI
GCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATC
2521
2580
CGGGGCTGTGGGCGGTTGTGGGCGACTGCGCGGGACTGCCCGAACAGACGAGGGCCGTAG
a A P T P A N T R * R A L T G L S A P G I -
b P R H P P T P A D A P * R A C L L P A S -
c P D T R Q H P L T R P D G L V C S R H P -
BsmAI
HphI
BsmBI
Hpy188I
NciI
NlaIII
ScrFI
NspI
BslI
MnlI
HpaII
CviAII
PfoI
FatI
StyD4I
HpyCH4V
HpyCH4III
BisI
MaeIII
Fnu4HI
Tsp45I
AluI
AluI
CviJI
HpyCH4III
SfaNI CviJI BbvI
TseI
HphI
CGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTC
2581
2640
GCGAATGTCTGTTCGACACTGGCAGAGGCCCTCGACGTACACAGTCTCCAAAAGTGGCAG
a R L Q T S C D R L R E L H V S E V F T V -
b A Y R Q A V T V S G S C M C Q R F S P S -
c L T D K L * P S P G A A C V R G F H R H -
AciI
BisI
Fnu4HI
AluI
CviJI
BstUI MspA1I
HhaI PvuII
HinfI
BstUI TseI
BisI
HinP1I BisI
AluI
Fnu4HI
BbvI Fnu4HI
CviJI MslI
HpyCH4V
BcgI MnlI TseI
BbvI BcgI
TseI TfiI
ATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGCAGCGATTC
2641
2700
TAGTGGCTTTGCGCGCTCCGTCGACGCCATTTCGAGTAGTCGCACCAGCACGTCGCTAAG
a I T E T R E A A A V K L I S V V V Q R F -
b S P K R A R Q L R * S S S A W S C S D S -
c H R N A R G S C G K A H Q R G R A A I H -
AluI

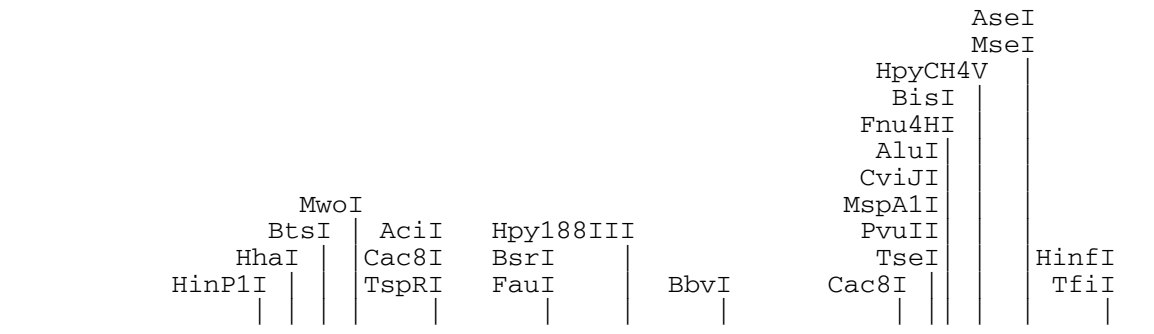

CviJI
BpmI
TspDTI
BstF5I
Eco57MI
FokI
HgaI
BstUI
BbvI
BsgI
AciI
Hpy188III
MseI
ACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGT
2701
2760
TGTCTACAGACGGACAAGTAGGCGCAGGTCGAGCAACTCAAAGAGGTCTTCGCAATTACA
a T D V C L F I R V Q L V E F L Q K R * C -
b Q M S A C S S A S S S L S F S R S V N V -
c R C L P V H P R P A R * V S P E A L M S -
BslI
MseI
NlaIII
CviAII
FatI
CviJI
HaeIII
TspGWI
FauI
Cac8I
MaeIII
Hpy188I
Sau96I
Tsp45I
CviJI
AciI
AciI
SfaNI
TspRI
CTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGC
2761
2820
GACCGAAGACTATTTCGCCCGGTACAATTCCCGCCAAAAAAGGACAAACCAGTGACTACG
a L A S D K A G H V K G G F F L F G H * C -
b W L L I K R A M L R A V F S C L V T D A -
c G F * * S G P C * G R F F P V W S L M P -
TspDTI
NlaIII
MnlI
CviAII
MnlI
BslI
FatI
SfaNI
TspDTI
CTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGAT
2821
2880
GAGGCACATTCCCCCTAAAGACAAGTACCCCCATTACTATGGCTACTTTGCTCTCTCCTA
a L R V R G I S V H G G N D T D E T R E D -
b S V * G G F L F M G V M I P M K R E R M -
c P C K G D F C S W G * * Y R * N E R G C -
TaiI
AclI
HpyCH4IV
BslI
TspDTI
MaeIII
HpaII
NciI
ScrFI
NlaIII
NspI
StyD4I
Hpy188III
MaeIII
CviAII
MnlI
BstF5I
FokI
FatI
BsrI
Hpy8I
GCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAA

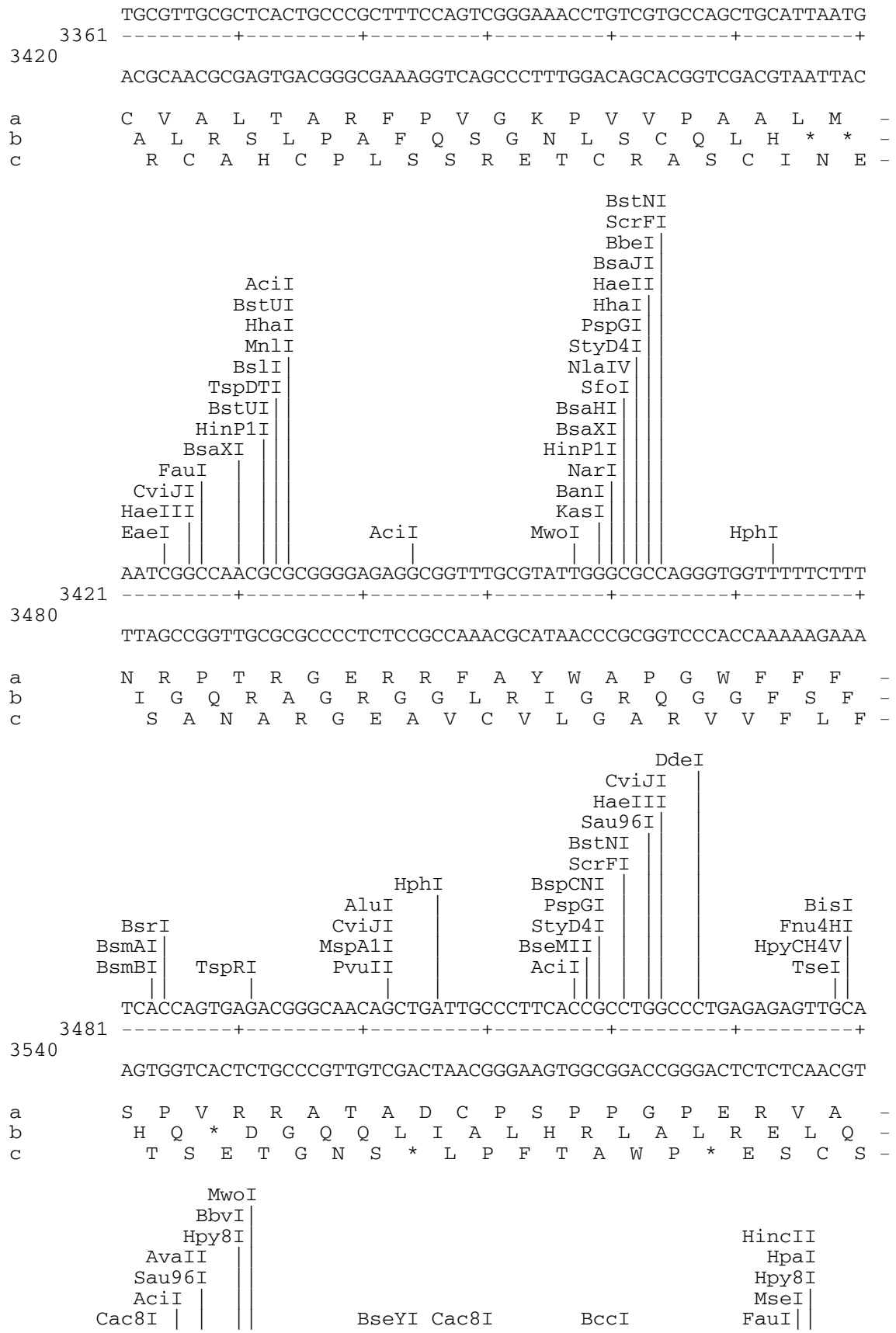

2881
2940
CGAGTGCTATGCCCAATGACTACTACTTGTACGGGCCAATGACCTTGCAACACTCCCATT
a A H D T G Y * * * T C P V T G T L * G *
b L T I R V T D D E H A R L L E R C E G K
c S R Y G L L M M N M P G Y W N V V R V N
FokI
NlaIV
AvaII
Sau96I
AciI
TauI
BstF5I
BisI
CviJI
BseMII
Cac8I
BspCNI
BsrI
AciI
Fnu4HI
AciI
DdeI
BsmFI
SfaNI
FauI
ACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCC
2941
3000
TGTTGACCGCCATACCTACGCCGCCCTGGTCTCTTTTTAGTGAGTCCCAGTTACGGTCGG
a T T G G M D A A G P E K N H S G S M P A
b Q L A V W M R R D Q R K I T Q G Q C Q P
c N W R Y G C G G T R E K S L R V N A S R
Cac8I
NaeI
BglI
HpaII
MwoI
BsrFI
NgoMIV
NlaIII
CviAII
FatI
TauI
AciI
BisI
BsiEI
Fnu4HI
CviJI
HaeIII
MwoI
Hpy99I
BstUI
HhaI
HinP1I
BseYI
CviJI
TaiI
HgaI
HpyCH4IV
Cac8I
EaeI
EagI
Cac8I
CviJI
HaeIII
MwoI
GAACGCCAGCAAGACGTAGCCCAGCGCGTCGGCCGCCATGCCGGCGATAATGGCCTGCTT
3001
3060
CTTGCGGTCGTTCTGCATCGGGTCGCGCAGCCGGCGGTACGGCCGCTATTACCGGACGAA
a E R Q Q D V A Q R V G R H A G D N G L L

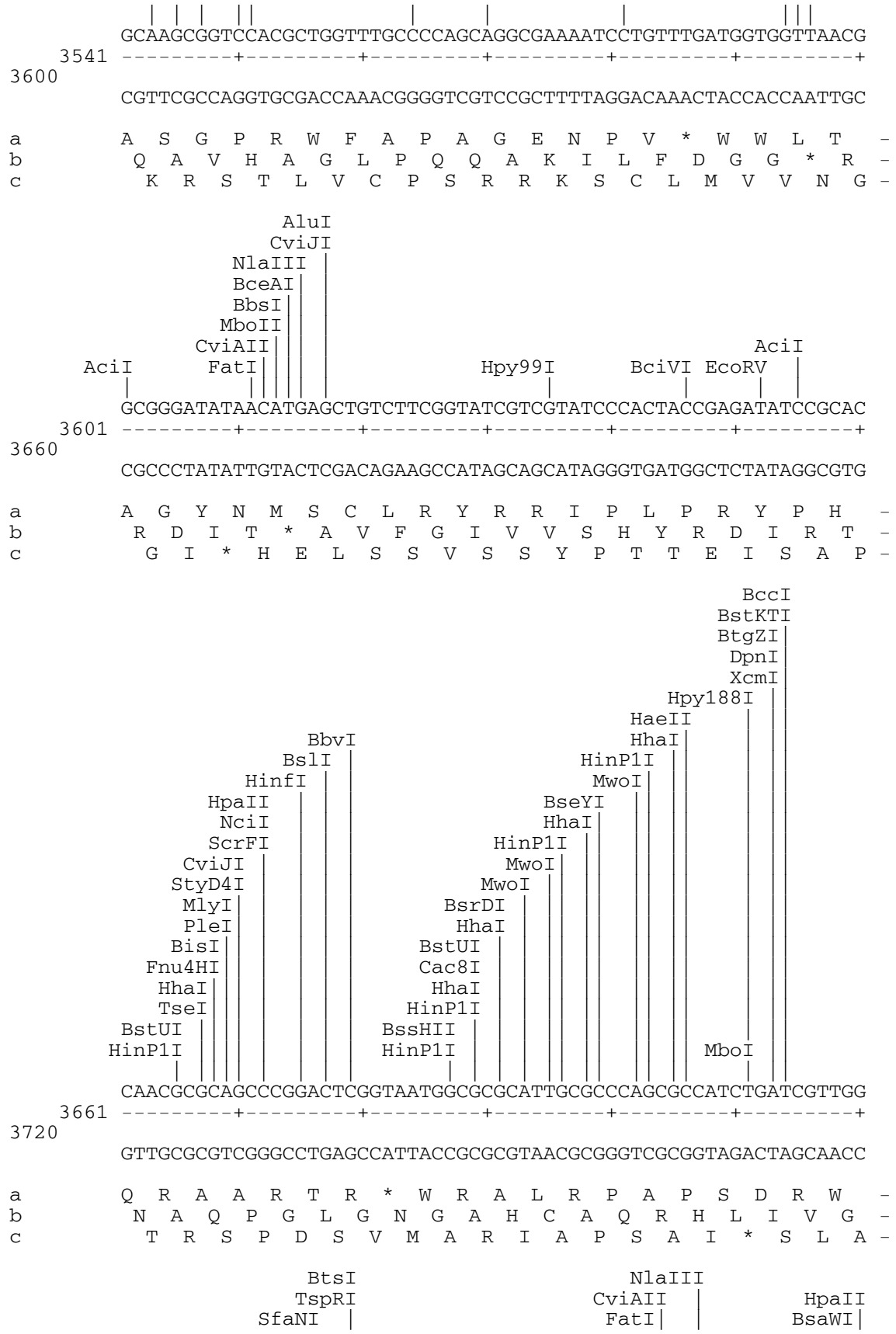

b N A S K T * P S A S A A M P A I M A C F -
c T P A R R S P A R R P P C R R * W P A S -
MaeIII
Tsp45I
BsrI
MnlI
FauI
NlaIV
SmlI
BpuEI
TaiI
AvaII
CviJI
HinfI
AclI
Sau96I
BsmFI
HpyCH4V
HpyCH4IV
AciI
TspRI
Cac8I
TfiI
CTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGAT
3061
3120
GAGCGGCTTTGCAAACCACCGCCCTGGTCACTGCTTCCGAACTCGCTCCCGCACGTTCTA
a L A E T F G G G T S D E G L S E G V Q D -
b S P K R L V A G P V T K A * A R A C K I -
c R R N V W W R D Q * R R L E R G R A R F -
BstKTI
DpnI
MboI
BpmI
HhaI
Eco57MI
BstUI
AvaII
Cac8I
CviJI
HinP1I
Sau96I
Hpy188I
AciI
HaeIII
Hpy99I
AciI
MwoI
TCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCC
3121
3180
AGGCTTATGGCGTTCGCTGTCCGGCTAGTAGCAGCGCGAGGTCGCTTTCGCCAGGAGCGG
a S E Y R K R Q A D H R R A P A K A V L A -
b P N T A S D R P I I V A L Q R K R S S P -
c R I P Q A T G R S S S R S S E S G P R R -
NlaIV
BanI
Cac8I
NaeI
HpaII
MwoI
BsrFI
NgoMIV
BisI
Fnu4HI
HaeII
HhaI
NlaIII
TseI
CviAII
BbvI
AfeI
FatI
BbsI
MnlI
HinP1I
HpyCH4V
HpyCH4III
GAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGT
3181
3240
CTTTTACTGGGTCTCGCGACGGCCGTGGACAGGATGCTCAACGTACTATTTCTTCTGTCA
a E N D P E R C R H L S Y E L H D K E D S -
b K M T Q S A A G T C P T S C M I K K T V -
c K * P R A L P A P V L R V A * * R R Q S -
BslI
BcgI

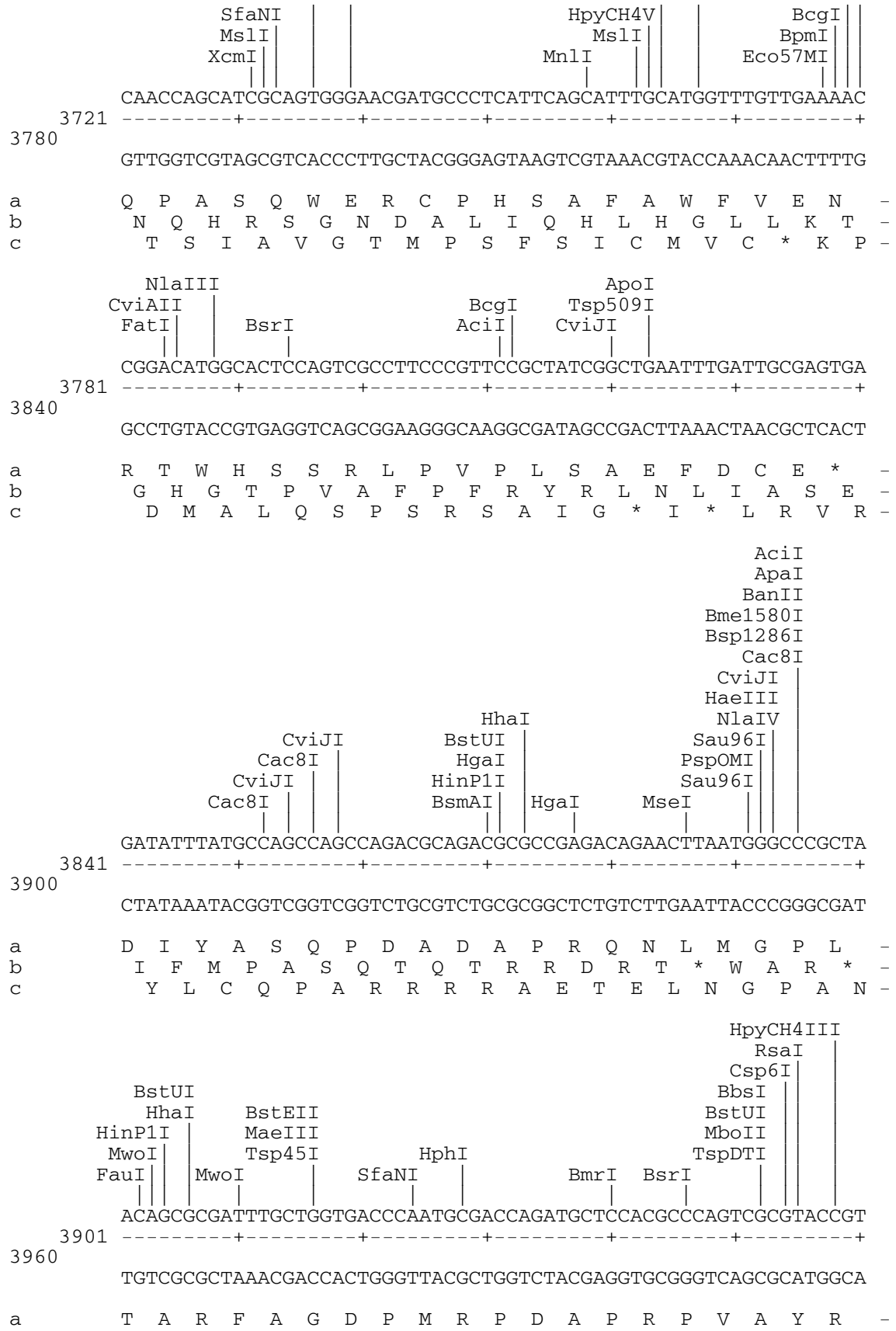

HpaII
BsaWI
FauI
HhaI
BstUI
HinP1I
AciI
NlaIII
CviAII
PshAI
Hpy99I
TauI
BisI
Fnu4HI
AciI
BcgI
MboII
FatI
BpuEI
AluI
CviJI
BsrI
BmrI
CATAAGTGCGGCGACGATAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAA
3241 ---------+---------+---------+---------+---------+---------+
3300
GTATTCACGCCGCTGCTATCAGTACGGGGCGCGGGTGGCCTTCCTCGACTGACCCAACTT
a H K C G D D S H A P R P P E G A D W V E -
b I S A A T I V M P R A H R K E L T G L K -
c * V R R R * S C P A P T G R S * L G * R -

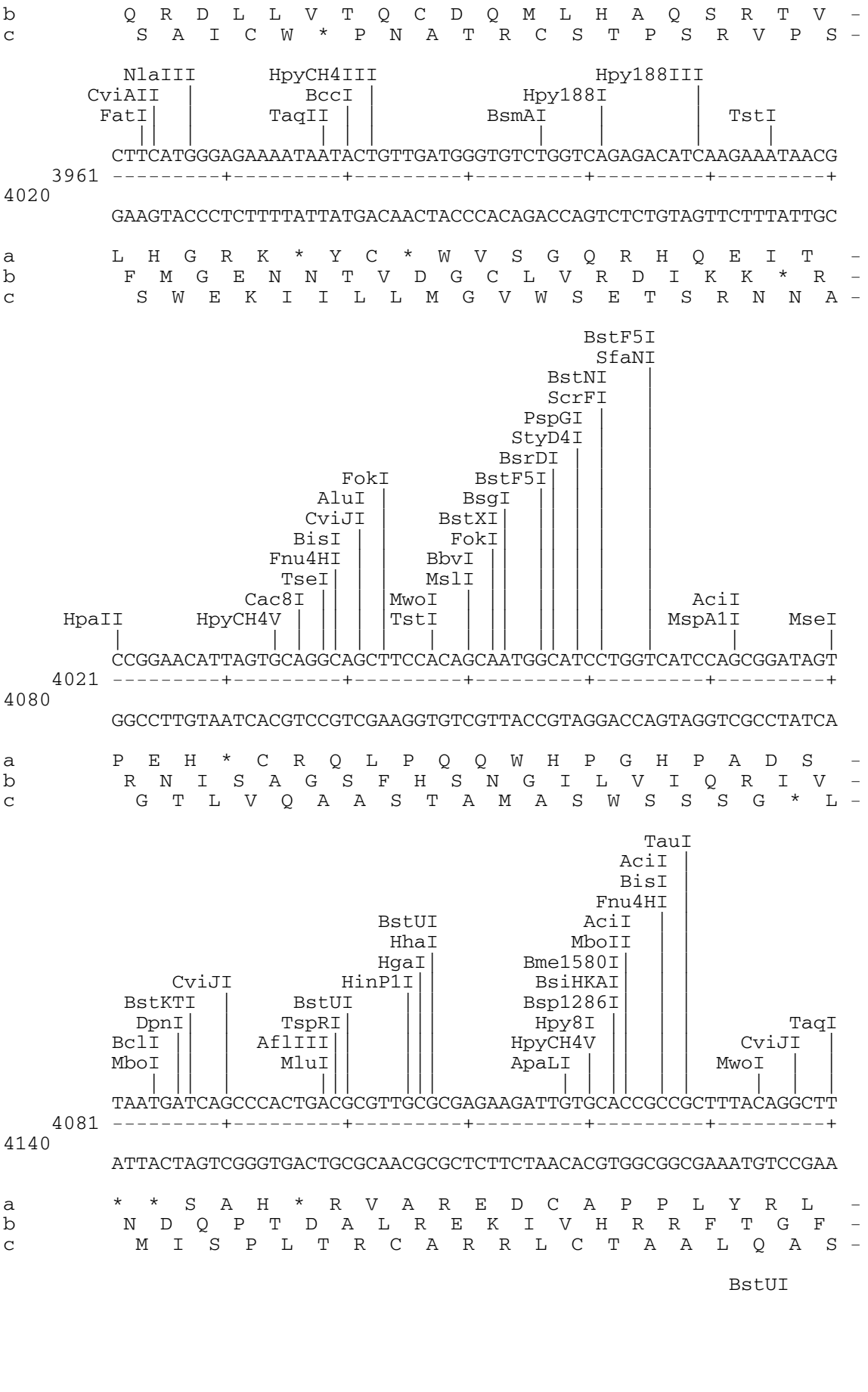

NlaIV
BanI
HpaII
NciI
ScrFI
BstKTI
StyD4I
DpnI
BstYI
MboI
SfaNI
Hpy188III
SmlI
BsiEI
TaqII
TaqI
CviJI
AlwI
AluI
CviJI
Tsp509I
AseI
MseI
GGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAAT
3301 ---------+---------+---------+---------+---------+---------+
3360
CCGAGAGTTCCCGTAGCCAGCTCTAGGGCCACGGATTACTCACTCGATTGAATGTAATTA
a G S Q G H R S R S R C L M S E L T Y I N -
b A L K G I G R D P G A * * V S * L T L I -
c L S R A S V E I P V P N E * A N L H * L -

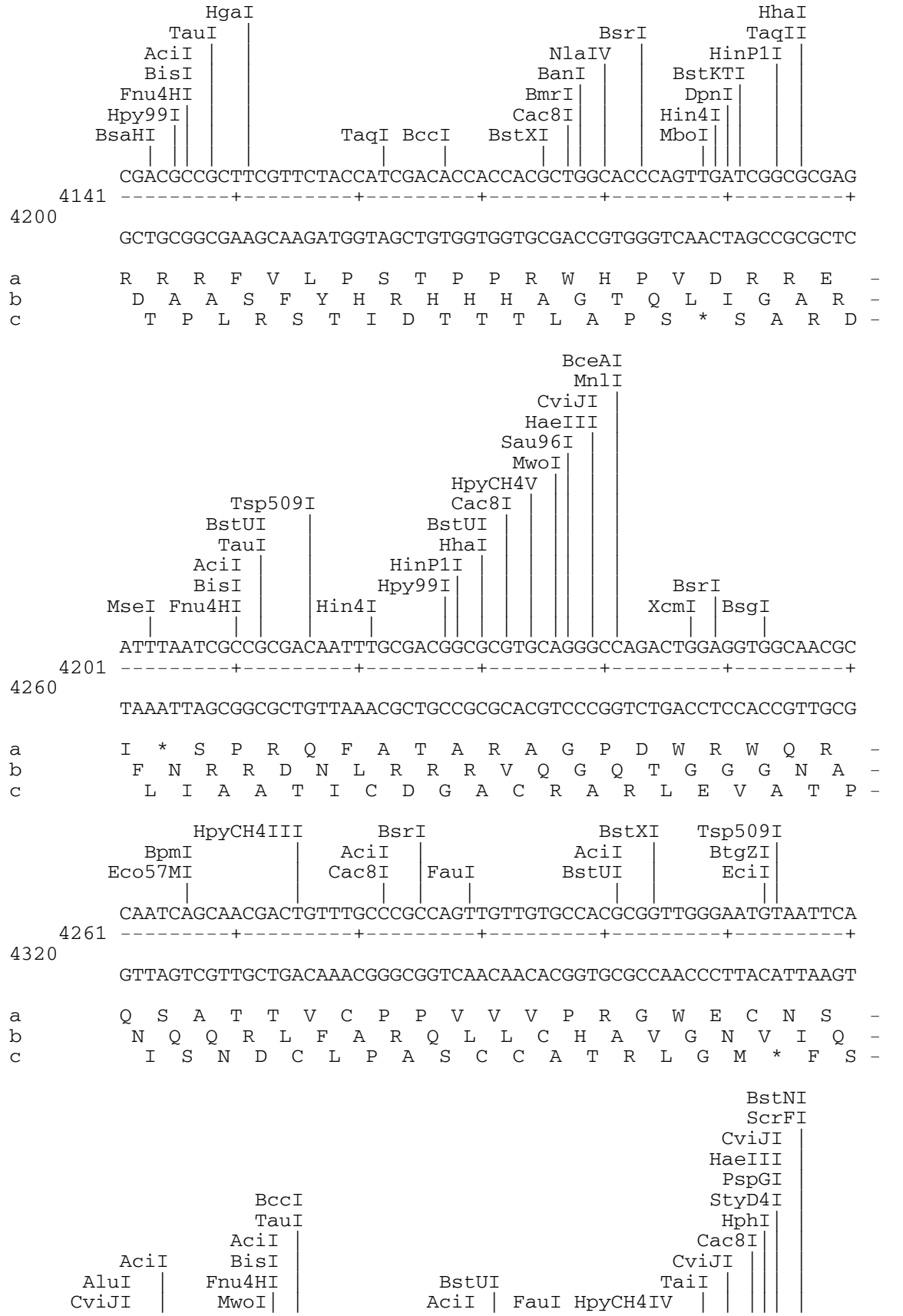

AseI
MseI
HpyCH4V
BisI
Fnu4HI
AluI
CviJI
MspA1I
PvuII
TseI
Cac8I
MwoI
BtsI
AciI
HhaI
Cac8I
HinP1I
TspRI
Hpy188III
BsrI
FauI
BbvI
HinfI
TfiI

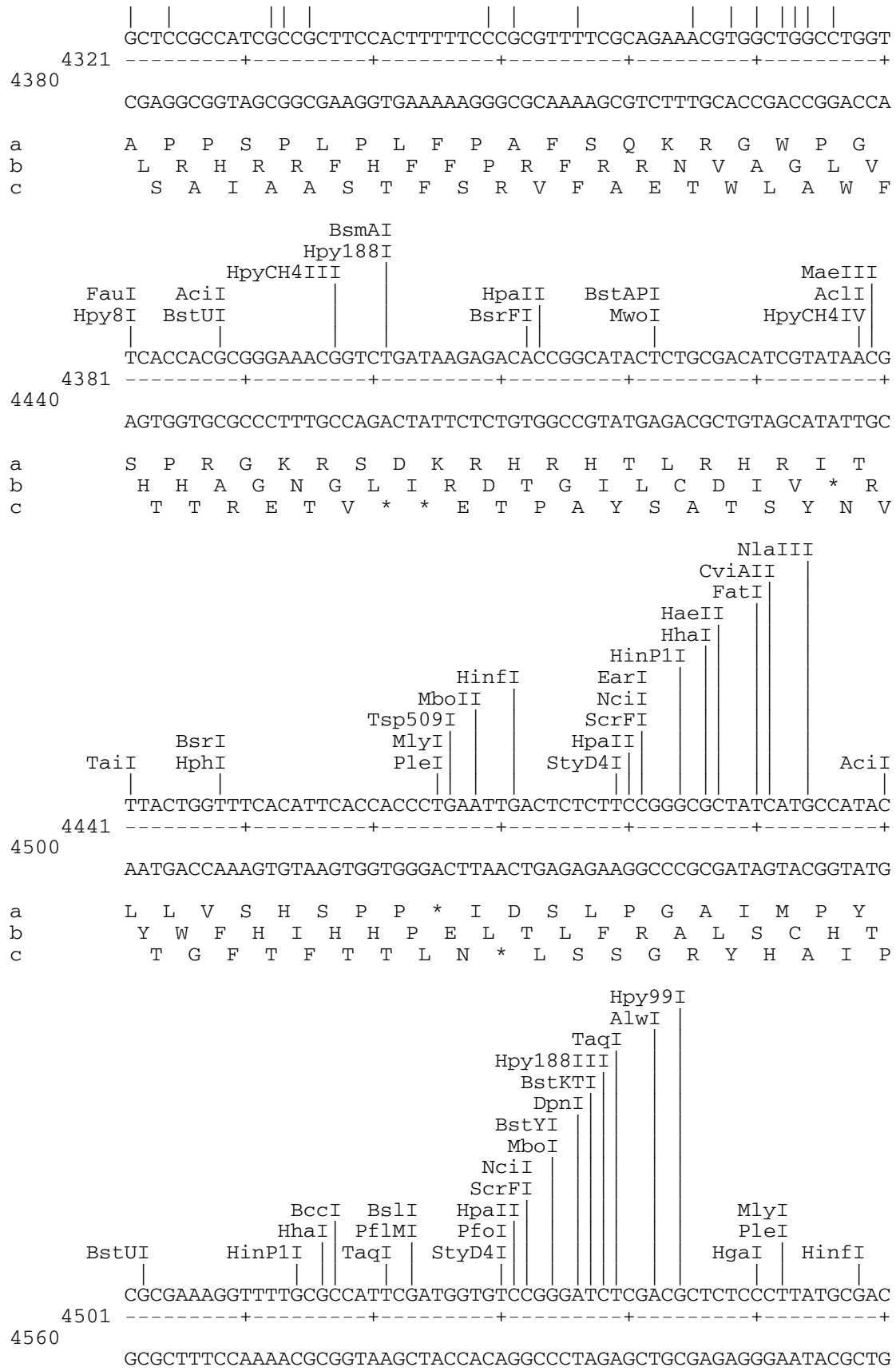

TGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATG
3361
3420
ACGCAACGCGAGTGACGGGCGAAAGGTCAGCCCTTTGGACAGCACGGTCGACGTAATTAC
a C V A L T A R F P V G K P V V P A A L M -
b A L R S L P A F Q S G N L S C Q L H * * -
c R C A H C P L S S R E T C R A S C I N E -
BstNI
ScrFI
BbeI
BsaJI
AciI
HaeII
BstUI
HhaI
HhaI
PspGI
MnlI
StyD4I
BslI
NlaIV
TspDTI
SfoI
BstUI
BsaHI
HinP1I
BsaXI
BsaXI
HinP1I
FauI
NarI
CviJI
BanI
HaeIII
KasI
EaeI
AciI
MwoI
HphI
AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTT
3421
3480
TTAGCCGGTTGCGCGCCCCTCTCCGCCAAACGCATAACCCGCGGTCCCACCAAAAAGAAA
a N R P T R G E R R F A Y W A P G W F F F -
b I G Q R A G R G G L R I G R Q G G F S F -
c S A N A R G E A V C V L G A R V V F L F -
DdeI
CviJI
HaeIII
Sau96I
BstNI
ScrFI
HphI
BspCNI
AluI
PspGI
BisI
BsrI
CviJI
StyD4I
Fnu4HI
BsmAI
MspA1I
BseMII
HpyCH4V
BsmBI
TspRI
PvuII
AciI
TseI
TCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCA
3481
3540
AGTGGTCACTCTGCCCGTTGTCGACTAACGGGAAGTGGCGGACCGGGACTCTCTCAACGT
a S P V R R A T A D C P S P P G P E R V A -
b H Q * D G Q Q L I A L H R L A L R E L Q -
c T S E T G N S * L P F T A W P * E S C S -
MwoI
BbvI
Hpy8I
HincII
AvaII
HpaI
Sau96I
Hpy8I
AciI
MseI
Cac8I
BseYI Cac8I
BccI
FauI

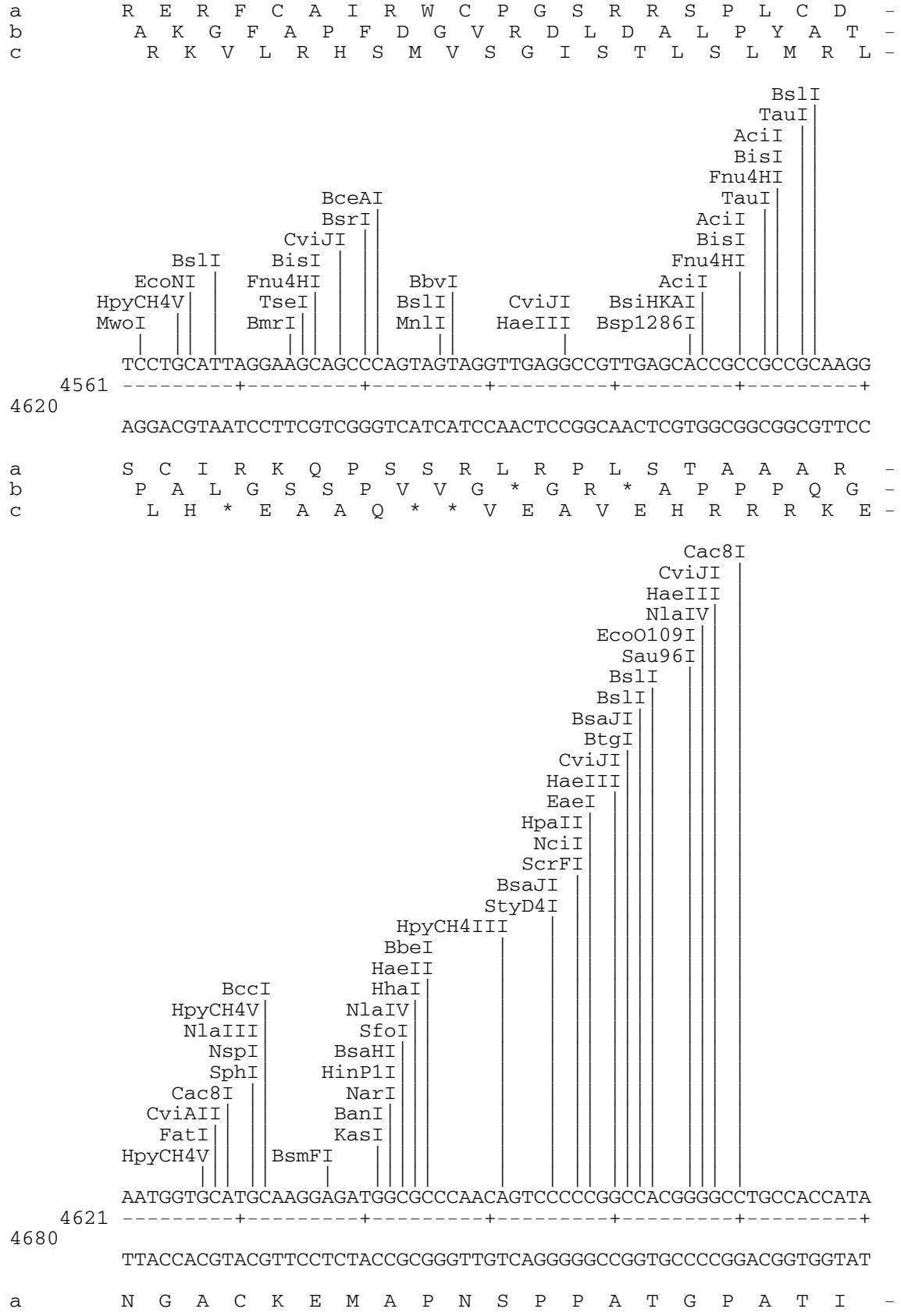

GCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACG
3541 ---------+---------+---------+---------+---------+---------+
3600
CGTTCGCCAGGTGCGACCAAACGGGGTCGTCCGCTTTTAGGACAAACTACCACCAATTGC
a A S G P R W F A P A G E N P V * W W L T -
b Q A V H A G L P Q Q A K I L F D G G * R -
c K R S T L V C P S R R K S C L M V V N G -
AluI
CviJI
NlaIII
BceAI
BbsI
MboII
CviAII
AciI
FatI
AciI
Hpy99I
BciVI
EcoRV
AciI
GCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCAC
3601 ---------+---------+---------+---------+---------+---------+
3660
CGCCCTATATTGTACTCGACAGAAGCCATAGCAGCATAGGGTGATGGCTCTATAGGCGTG
a A G Y N M S C L R Y R R I P L P R Y P H -
b R D I T * A V F G I V V S H Y R D I R T -
c G I * H E L S S V S S Y P T T E I S A P -
BccI
BstKTI
BtgZI
DpnI
XcmI
Hpy188I
HaeII
HhaI
BbvI
BslI
HinP1I
HinfI
MwoI
HpaII
BseYI
NciI
HhaI
ScrFI
HinP1I
CviJI
MwoI
StyD4I
MwoI
BsrDI
MlyI
HhaI
PleI
BstUI
BisI
Cac8I
Fnu4HI
HhaI
HhaI
TseI
HinP1I
BstUI
BssHII
MboI
HinP1I
HinP1I
CAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGG
3661 ---------+---------+---------+---------+---------+---------+
3720
GTTGCGCGTCGGGCCTGAGCCATTACCGCGCGTAACGCGGGTCGCGGTAGACTAGCAACC
a Q R A A R T R * W R A L R P A P S D R W -
b N A Q P G L G N G A H C A Q R H L I V G -
c T R S P D S V M A R I A P S A I * S L A -
BtsI
TspRI
SfaNI
NlaIII
CviAII
FatI
HpaII
BsaWI

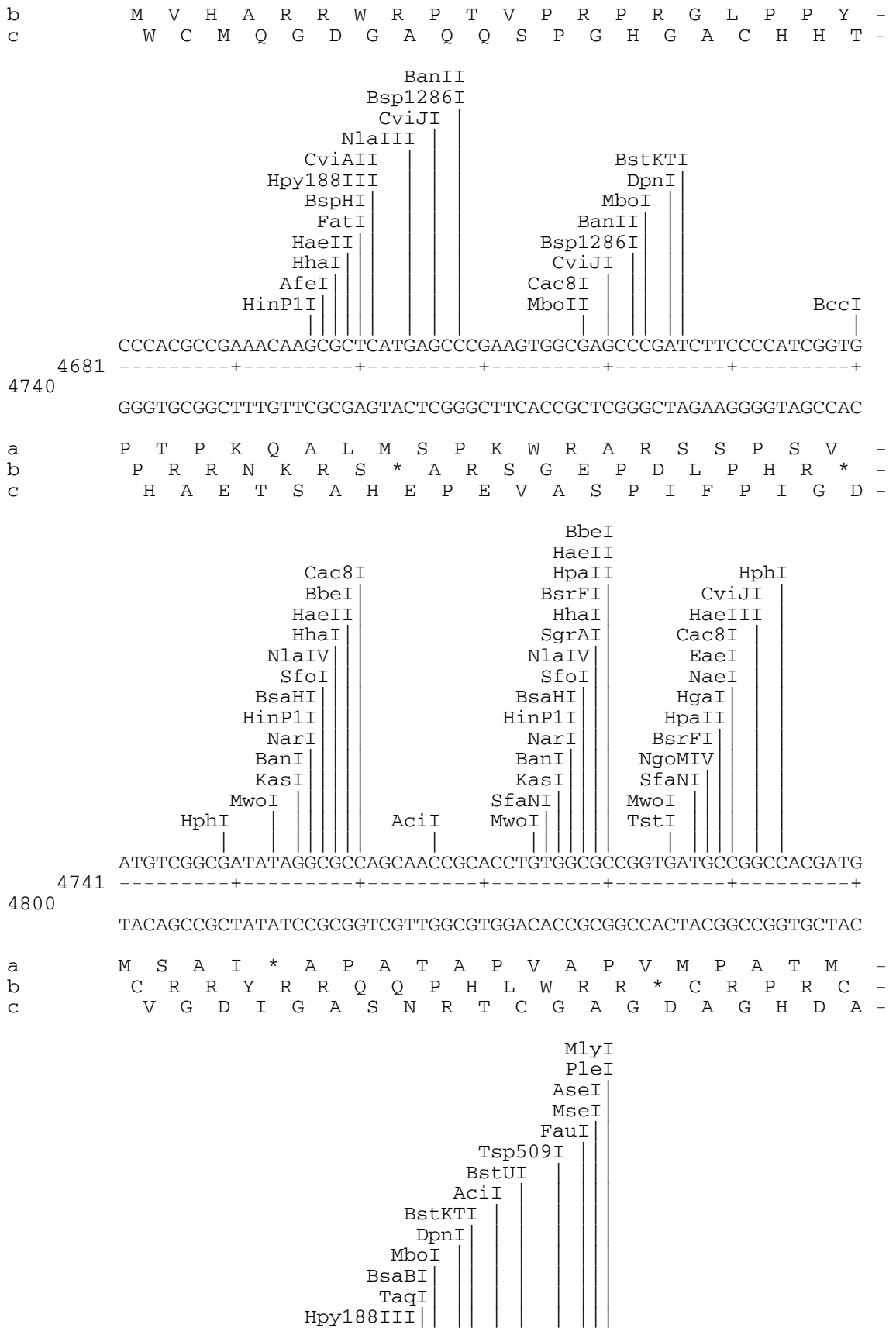

SfaNI
MslI
XcmI
HpyCH4V
MslI
MnlI
BcgI
BpmI
Eco57MI
CAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTTGTTGAAAAC
3721
3780
GTTGGTCGTAGCGTCACCCTTGCTACGGGAGTAAGTCGTAAACGTACCAAACAACTTTTG
a Q P A S Q W E R C P H S A F A W F V E N -
b N Q H R S G N D A L I Q H L H G L L K T -
c T S I A V G T M P S F S I C M V C * K P -
NlaIII
CviAII
FatI
BsrI
ApoI
BcgI
Tsp509I
AciI
CviJI
CGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGA
3781
3840
GCCTGTACCGTGAGGTCAGCGGAAGGGCAAGGCGATAGCCGACTTAAACTAACGCTCACT
a R T W H S S R L P V P L S A E F D C E * -
b G H G T P V A F P F R Y R L N L I A S E -
c D M A L Q S P S R S A I G * I * L R V R -
AciI
ApaI
BanII
Bme1580I
Bsp1286I
Cac8I
CviJI
HaeIII
NlaIV
HhaI
CviJI
BstUI
Sau96I
Cac8I
HgaI
PspOMI
CviJI
HinP1I
Sau96I
Cac8I
BsmAI
HgaI
MseI
GATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTA
3841
3900
CTATAAATACGGTCGGTCGGTCTGCGTCTGCGCGGCTCTGTCTTGAATTACCCGGGCGAT
a D I Y A S Q P D A D A P R Q N L M G P L -
b I F M P A S Q T Q T R R D R T * W A R * -
c Y L C Q P A R R R R A E T E L N G P A N -
HpyCH4III
RsaI
Csp6I
BbsI
BstUI
MboII
TspDTI
BstUI
HhaI
HinP1I
MwoI
FauI
BstEII
MaeIII
Tsp45I
MwoI
HphI
SfaNI
BmrI
BsrI
ACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGT
3901
3960
TGTCGCGCTAAACGACCACTGGGTTACGCTGGTCTACGAGGTGCGGGTCAGCGCATGGCA
a T A R F A G D P M R P D A P R P V A Y R -

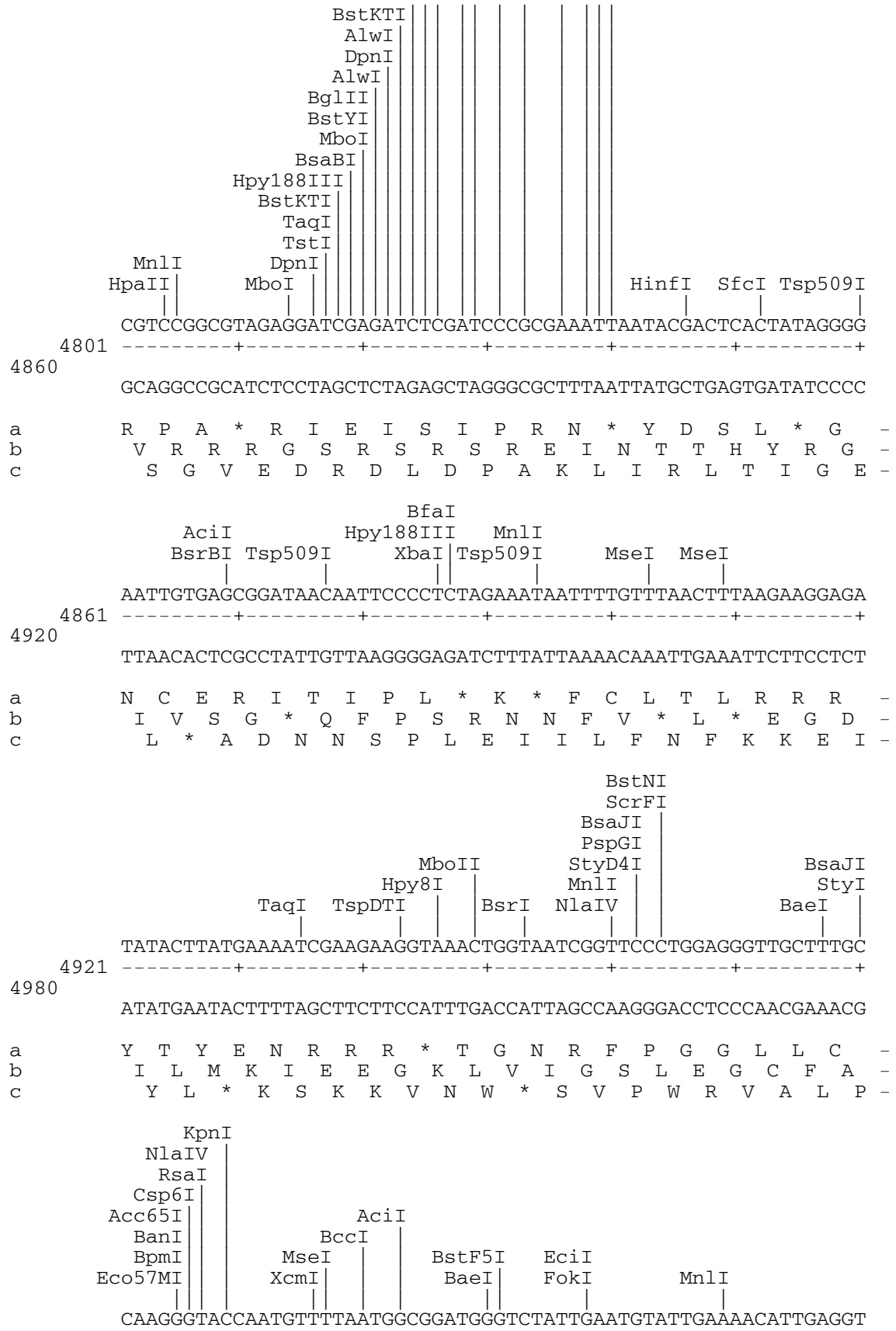

b Q R D L L V T Q C D Q M L H A Q S R T V -
c S A I C W * P N A T R C S T P S R V P S -
NlaIII
HpyCH4III
Hpy188III
CviAII
BccI
Hpy188I
FatI
TaqII
BsmAI
TstI
CTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACG
3961
4020
GAAGTACCCTCTTTTATTATGACAACTACCCACAGACCAGTCTCTGTAGTTCTTTATTGC
a L H G R K * Y C * W V S G Q R H Q E I T -
b F M G E N N T V D G C L V R D I K K * R -
c S W E K I I L L M G V W S E T S R N N A -
BstF5I
SfaNI
BstNI
ScrFI
PspGI
StyD4I
BsrDI
FokI
BstF5I
AluI
BsgI
CviJI
BstXI
BisI
FokI
Fnu4HI
BbvI
TseI
MslI
Cac8I
MwoI
AciI
HpaII
HpyCH4V
TstI
MspA1I
MseI
CCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGT
4021
4080
GGCCTTGTAATCACGTCCGTCGAAGGTGTCGTTACCGTAGGACCAGTAGGTCGCCTATCA
a P E H * C R Q L P Q Q W H P G H P A D S -
b R N I S A G S F H S N G I L V I Q R I V -
c G T L V Q A A S T A M A S W S S S G * L -
TauI
AciI
BisI
Fnu4HI
BstUI
AciI
HhaI
MboII
HgaI
Bme1580I
CviJI
HinP1I
BsiHKAI
BstKTI
BstUI
Bsp1286I
DpnI
TspRI
Hpy8I
TaqI
BclI
AflIII
HpyCH4V
CviJI
MboI
MluI
ApaLI
MwoI
TAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTT
4081
4140
ATTACTAGTCGGGTGACTGCGCAACGCGCTCTTCTAACACGTGGCGGCGAAATGTCCGAA
a * * S A H * R V A R E D C A P P L Y R L -
b N D Q P T D A L R E K I V H R R F T G F -
c M I S P L T R C A R R L C T A A L Q A S -
BstUI

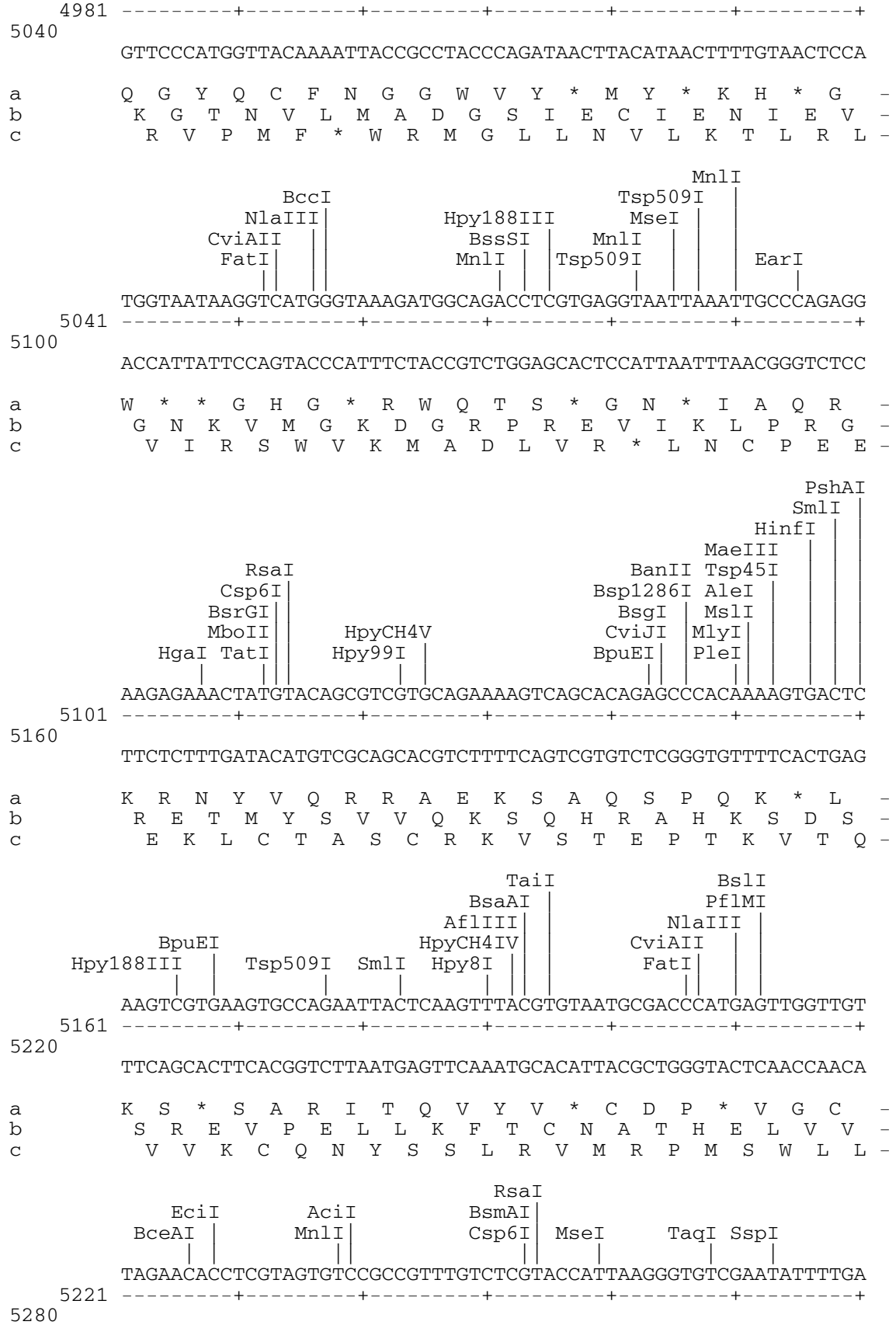

HgaI
TauI
AciI
BisI
Fnu4HI
Hpy99I
BsaHI
TaqI BccI
BstXI
BsrI
NlaIV
BanI
BmrI
Cac8I
HhaI
TaqII
HinP1I
BstKTI
DpnI
Hin4I
MboI
CGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAG
4141
4200
GCTGCGGCGAAGCAAGATGGTAGCTGTGGTGGTGCGACCGTGGGTCAACTAGCCGCGCTC
a R R R F V L P S T P P R W H P V D R R E -
b D A A S F Y H R H H H A G T Q L I G A R -
c T P L R S T I D T T T L A P S * S A R D -
BceAI
MnlI
CviJI
HaeIII
Sau96I
MwoI
HpyCH4V
Cac8I
Tsp509I
BstUI
TauI
AciI
BisI
MseI Fnu4HI
Hin4I
HinP1I
Hpy99I
BsrI
XcmI
BsgI
ATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGC
4201
4260
TAAATTAGCGGCGCTGTTAAACGCTGCCGCGCACGTCCCGGTCTGACCTCCACCGTTGCG
a I * S P R Q F A T A R A G P D W R W Q R -
b F N R R D N L R R R V Q G Q T G G G N A -
c L I A A T I C D G A C R A R L E V A T P -
HpyCH4III
BsrI
BstXI
Tsp509I
BpmI
AciI
AciI
BtgZI
Eco57MI
Cac8I
FauI
BstUI
EciI
CAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCA
4261
4320
GTTAGTCGTTGCTGACAAACGGGCGGTCAACAACACGGTGCGCCAACCCTTACATTAAGT
a Q S A T T V C P P V V V P R G W E C N S -
b N Q Q R L F A R Q L L C H A V G N V I Q -
c I S N D C L P A S C C A T R L G M * F S -
BstNI
ScrFI
CviJI
HaeIII
PspGI
StyD4I
HphI
Cac8I
CviJI
TaiI
BccI
TauI
AciI
BisI
Fnu4HI
MwoI
AciI
AluI
CviJI
BstUI
AciI
FauI
HpyCH4IV


GCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGT
4321
4380
CGAGGCGGTAGCGGCGAAGGTGAAAAAGGGCGCAAAAGCGTCTTTGCACCGACCGGACCA
a A P P S P L P L F P A F S Q K R G W P G -
b L R H R R F H F F P R F R R N V A G L V -
c S A I A A S T F S R V F A E T W L A W F -
BsmAI
Hpy188I
HpyCH4III
MaeIII
FauI
AciI
AclI
Hpy8I
BstUI
HpaII
BstAPI
BsrFI
MwoI
HpyCH4IV
TCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACG
4381
4440
AGTGGTGCGCCCTTTGCCAGACTATTCTCTGTGGCCGTATGAGACGCTGTAGCATATTGC
a S P R G K R S D K R H R H T L R H R I T -
b H H A G N G L I R D T G I L C D I V * R -
c T T R E T V * * E T P A Y S A T S Y N V -
NlaIII
CviAII
FatI
HaeII
HhaI
HinP1I
HinfI
EarI
MboII
NciI
Tsp509I
ScrFI
BsrI
MlyI
HpaII
TaiI
HphI
PleI
StyD4I
AciI
TTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATAC
4441
4500
AATGACCAAAGTGTAAGTGGTGGGACTTAACTGAGAGAAGGCCCGCGATAGTACGGTATG
a L L V S H S P P * I D S L P G A I M P Y -
b Y W F H I H H P E L T L F R A L S C H T -
c T G F T F T T L N * L S S G R Y H A I P -
Hpy99I
AlwI
TaqI
Hpy188III
BstKTI
DpnI
BstYI
MboI
NciI
ScrFI
BccI
BslI
HpaII
MlyI
HhaI
PflMI
PfoI
PleI
BstUI
HinP1I
TaqI
StyD4I
HgaI
HinfI
CGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGAC
4501
4560
GCGCTTTCCAAAACGCGGTAAGCTACCACAGGCCCTAGAGCTGCGAGAGGGAATACGCTG

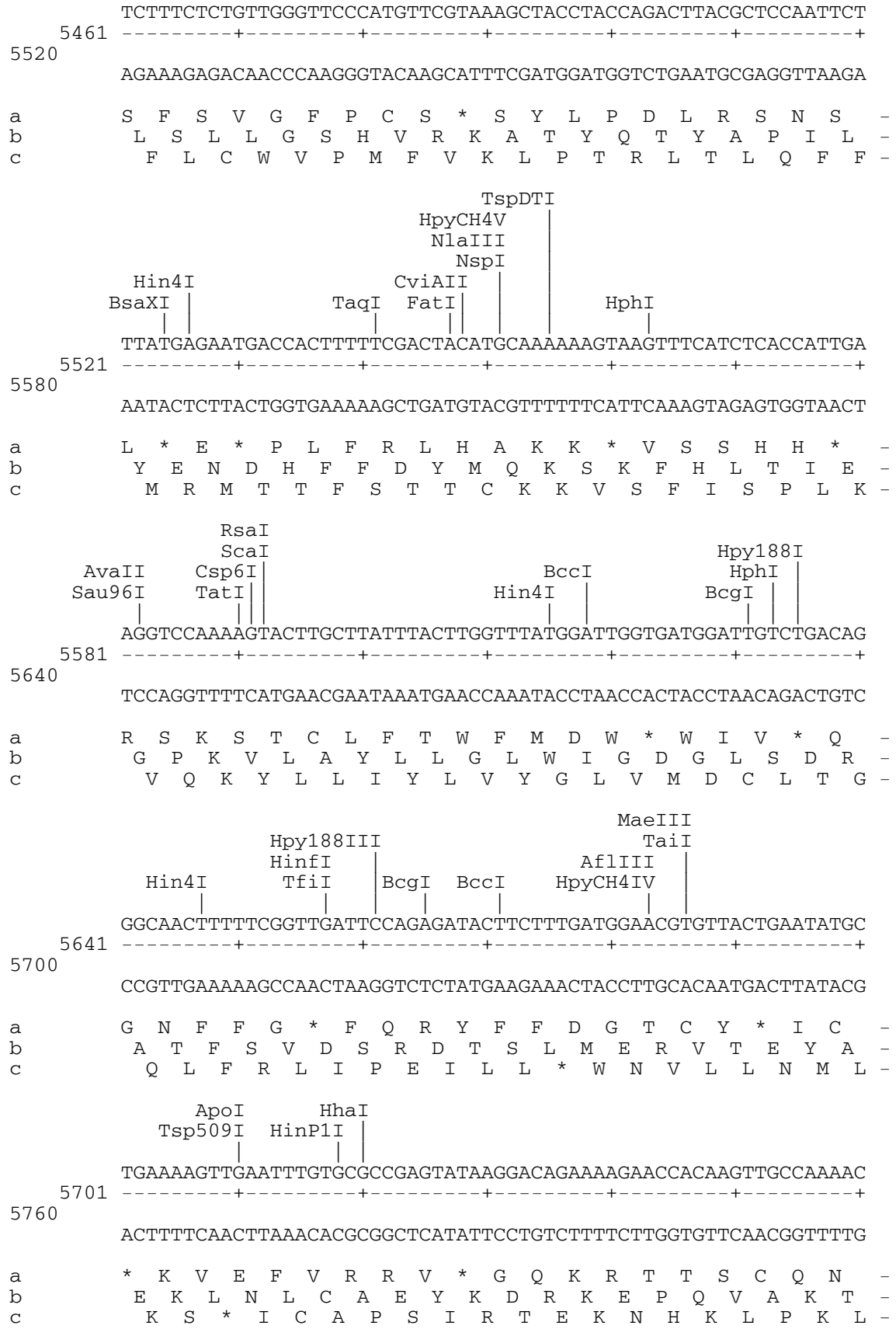

a R E R F C A I R W C P G S R R S P L C D -
b A K G F A P F D G V R D L D A L P Y A T -
c R K V L R H S M V S G I S T L S L M R L -
BslI
TauI
AciI
BisI
Fnu4HI
TauI
BceAI
BsrI
AciI
CviJI
BisI
BslI
BisI
Fnu4HI
EcoNI
Fnu4HI
AciI
BbvI
HpyCH4V
TseI
BslI
CviJI
BsiHKAI
MwoI
BmrI
MnlI
HaeIII
Bsp1286I
TCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGG
4561 ---------+---------+---------+---------+---------+---------+
4620
AGGACGTAATCCTTCGTCGGGTCATCATCCAACTCCGGCAACTCGTGGCGGCGGCGTTCC
a S C I R K Q P S S R L R P L S T A A A R -
b P A L G S S P V V G * G R * A P P P Q G -
c L H * E A A Q * * V E A V E H R R R K E -
Cac8I
CviJI
HaeIII
NlaIV
EcoO109I
Sau96I
BslI
BslI
BsaJI
BtgI
CviJI
HaeIII
EaeI
HpaII
NciI
ScrFI
BsaJI
StyD4I
HpyCH4III
BbeI
HaeII
BccI
HhaI
HpyCH4V
NlaIV
NlaIII
SfoI
NspI
BsaHI
SphI
HinP1I
Cac8I
NarI
CviAII
BanI
FatI
KasI
HpyCH4V
BsmFI
AATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATA
4621 ---------+---------+---------+---------+---------+---------+
4680
TTACCACGTACGTTCCTCTACCGCGGGTTGTCAGGGGGCCGGTGCCCCGGACGGTGGTAT
a N G A C K E M A P N S P P A T G P A T I -

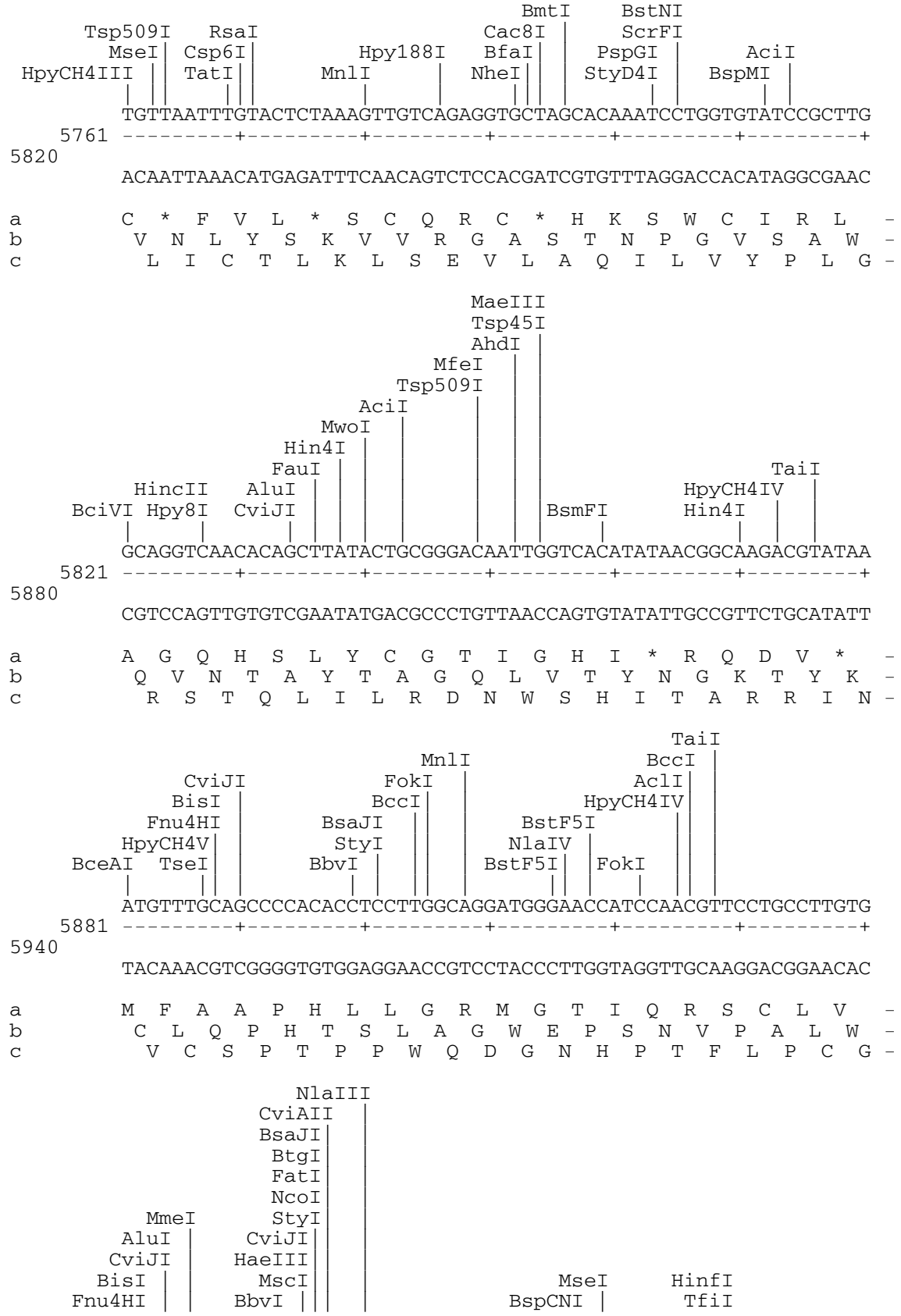

b M V H A R R W R P T V P R P R G L P P Y -
c W C M Q G D G A Q Q S P G H G A C H H T -
BanII
Bsp1286I
CviJI
NlaIII
CviAII
BstKTI
Hpy188III
DpnI
BspHI
MboI
FatI
BanII
HaeII
Bsp1286I
HhaI
CviJI
AfeI
Cac8I
HinP1I
MboII
BccI
CCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTG
4681 ---------+---------+---------+---------+---------+---------+
4740
GGGTGCGGCTTTGTTCGCGAGTACTCGGGCTTCACCGCTCGGGCTAGAAGGGGTAGCCAC
a P T P K Q A L M S P K W R A R S S P S V -
b P R R N K R S * A R S G E P D L P H R * -
c H A E T S A H E P E V A S P I F P I G D -
BbeI
HaeII
Cac8I
HpaII
HphI
BbeI
BsrFI
CviJI
HaeII
HhaI
HaeIII
HhaI
SgrAI
Cac8I
NlaIV
NlaIV
EaeI
SfoI
SfoI
NaeI
BsaHI
BsaHI
HgaI
HinP1I
HinP1I
HpaII
NarI
NarI
BsrFI
BanI
BanI
NgoMIV
KasI
KasI
SfaNI
MwoI
SfaNI
MwoI
HphI
AciI
MwoI
TstI
ATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATG
4741 ---------+---------+---------+---------+---------+---------+
4800
TACAGCCGCTATATCCGCGGTCGTTGGCGTGGACACCGCGGCCACTACGGCCGGTGCTAC
a M S A I * A P A T A P V A P V M P A T M -
b C R R Y R R Q Q P H L W R R * C R P R C -
c V G D I G A S N R T C G A G D A G H D A -
MlyI
PleI
AseI
MseI
FauI
Tsp509I
BstUI
AciI
BstKTI
DpnI
MboI
BsaBI
TaqI
Hpy188III

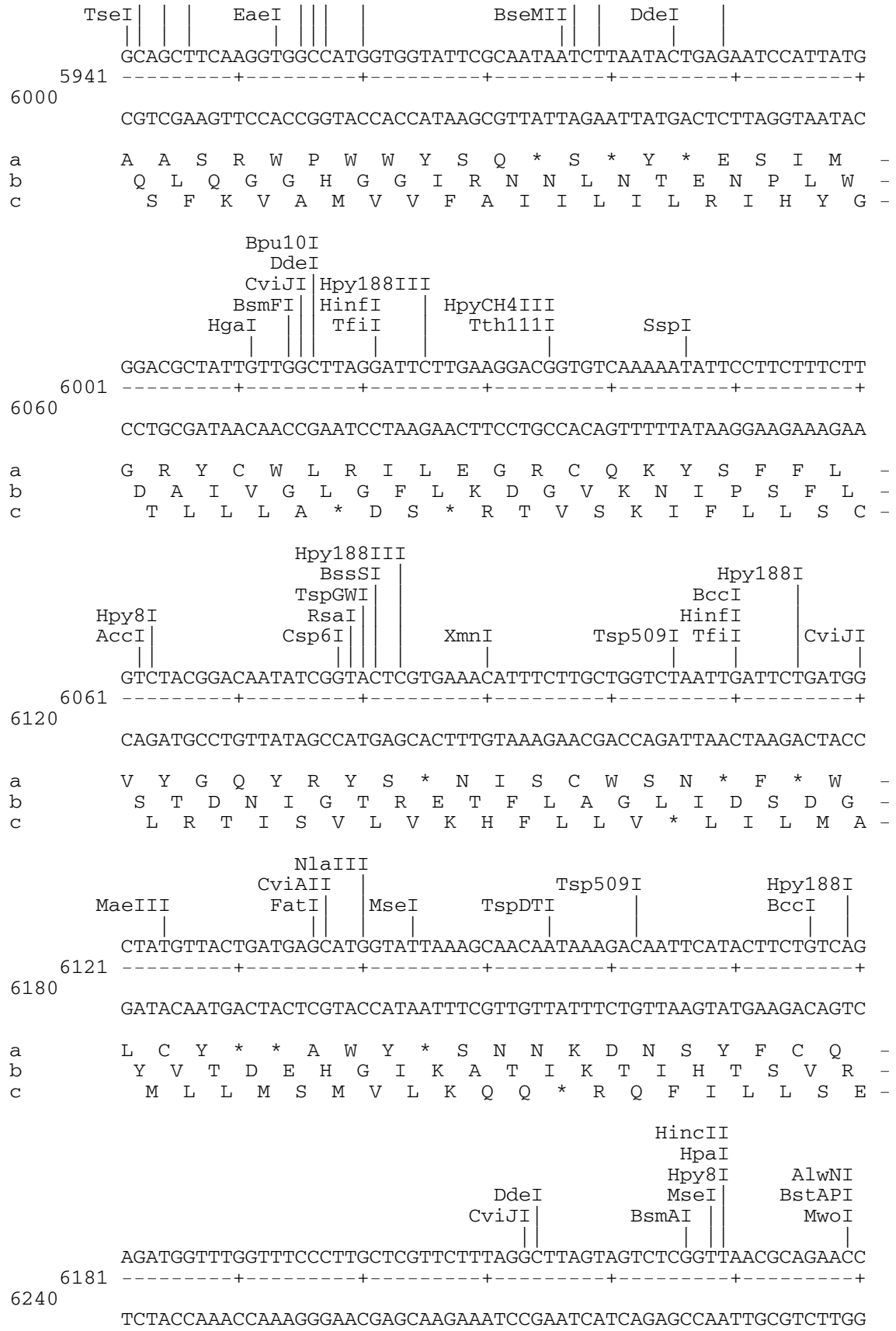

BstKTI
AlwI
DpnI
AlwI
BglII
BstYI
MboI
BsaBI
Hpy188III
BstKTI
TaqI
TstI
MnlI
DpnI
HpaII
MboI
HinfI
SfcI
Tsp509I
CGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGG
4801
4860
GCAGGCCGCATCTCCTAGCTCTAGAGCTAGGGCGCTTTAATTATGCTGAGTGATATCCCC
a R P A * R I E I S I P R N * Y D S L * G -
b V R R R G S R S R S R E I N T T H Y R G -
c S G V E D R D L D P A K L I R L T I G E -
BfaI
AciI
Hpy188III
MnlI
BsrBI
Tsp509I
XbaI
Tsp509I
MseI
MseI
AATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGA
4861
4920
TTAACACTCGCCTATTGTTAAGGGGAGATCTTTATTAAAACAAATTGAAATTCTTCCTCT
a N C E R I T I P L * K * F C L T L R R R -
b I V S G * Q F P S R N N F V * L * E G D -
c L * A D N N S P L E I I L F N F K K E I -
BstNI
ScrFI
BsaJI
PspGI
MboII
StyD4I
BsaJI
Hpy8I
MnlI
StyI
TaqI
TspDTI
BsrI
NlaIV
BaeI
TATACTTATGAAAATCGAAGAAGGTAAACTGGTAATCGGTTCCCTGGAGGGTTGCTTTGC
4921
4980
ATATGAATACTTTTAGCTTCTTCCATTTGACCATTAGCCAAGGGACCTCCCAACGAAACG
a Y T Y E N R R R * T G N R F P G G L L C -
b I L M K I E E G K L V I G S L E G C F A -
c Y L * K S K K V N W * S V P W R V A L P -
KpnI
NlaIV
RsaI
Csp6I
Acc65I
AciI
BanI
BccI
BpmI
MseI
BstF5I
EciI
Eco57MI
XcmI
BaeI
FokI
MnlI
CAAGGGTACCAATGTTTTAATGGCGGATGGGTCTATTGAATGTATTGAAAACATTGAGGT

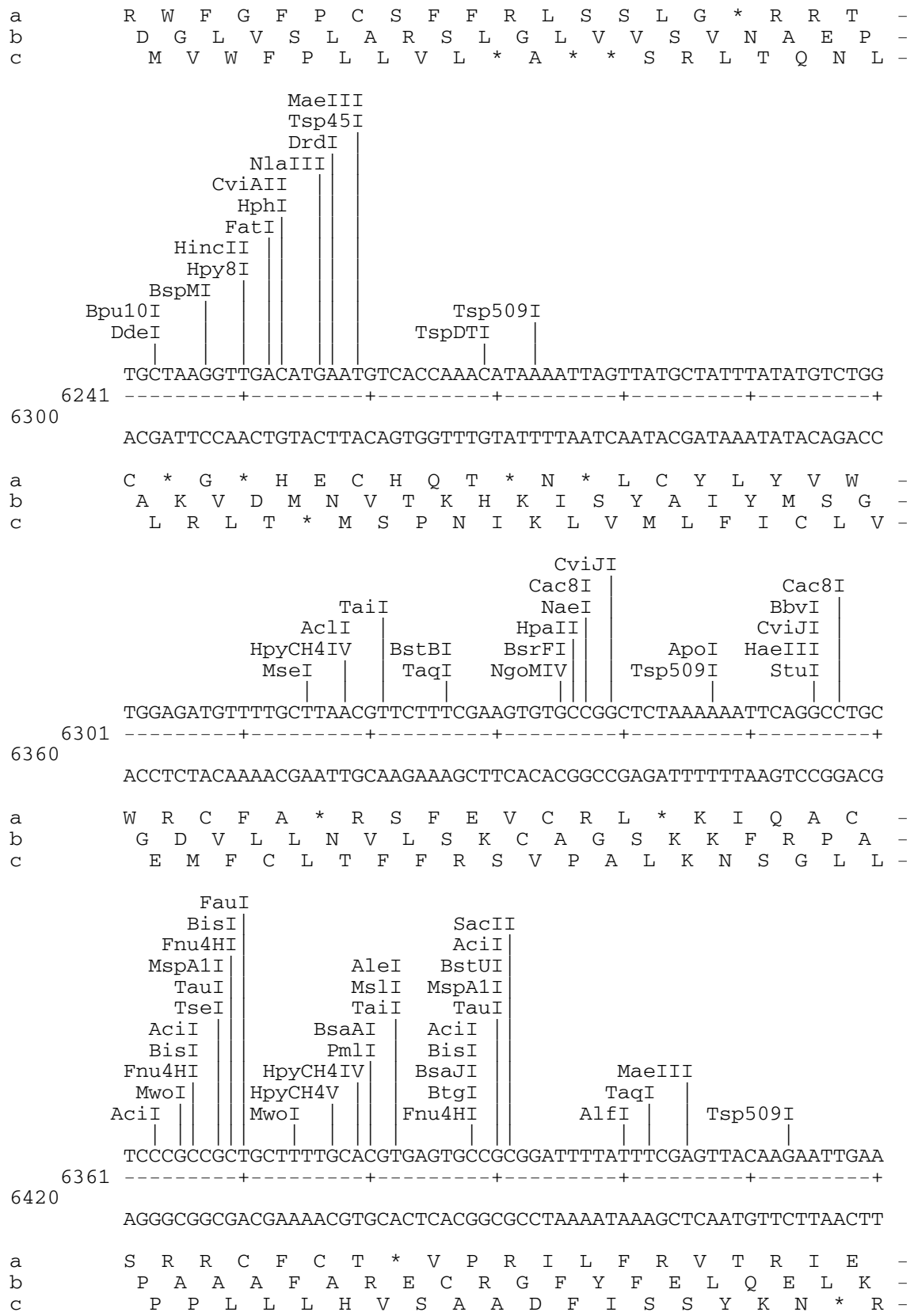

4981
5040
GTTCCCATGGTTACAAAATTACCGCCTACCCAGATAACTTACATAACTTTTGTAACTCCA
a Q G Y Q C F N G G W V Y * M Y * K H * G -
b K G T N V L M A D G S I E C I E N I E V -
c R V P M F * W R M G L L N V L K T L R L -
MnlI
Tsp509I
BccI
MseI
NlaIII
Hpy188III
BssSI
MnlI
CviAII
MnlI
Tsp509I
EarI
FatI
TGGTAATAAGGTCATGGGTAAAGATGGCAGACCTCGTGAGGTAATTAAATTGCCCAGAGG
5041
5100
ACCATTATTCCAGTACCCATTTCTACCGTCTGGAGCACTCCATTAATTTAACGGGTCTCC
a W * * G H G * R W Q T S * G N * I A Q R -
b G N K V M G K D G R P R E V I K L P R G -
c V I R S W V K M A D L V R * L N C P E E -
PshAI
SmlI
HinfI
MaeIII
RsaI
BanII
Tsp45I
Csp6I
Bsp1286I
AleI
BsrGI
BsgI
MslI
MboII
HpyCH4V
CviJI
MlyI
HgaI
TatI
Hpy99I
BpuEI
PleI
AAGAGAAACTATGTACAGCGTCGTGCAGAAAAGTCAGCACAGAGCCCACAAAAGTGACTC
5101
5160
TTCTCTTTGATACATGTCGCAGCACGTCTTTTCAGTCGTGTCTCGGGTGTTTTCACTGAG
a K R N Y V Q R R A E K S A Q S P Q K * L -
b R E T M Y S V V Q K S Q H R A H K S D S -
c E K L C T A S C R K V S T E P T K V T Q -
TaiI
BslI
BsaAI
PflMI
AflIII
NlaIII
BpuEI
HpyCH4IV
CviAII
Hpy188III
Tsp509I
SmlI
Hpy8I
FatI
AAGTCGTGAAGTGCCAGAATTACTCAAGTTTACGTGTAATGCGACCCATGAGTTGGTTGT
5161
5220
TTCAGCACTTCACGGTCTTAATGAGTTCAAATGCACATTACGCTGGGTACTCAACCAACA
a K S * S A R I T Q V Y V * C D P * V G C -
b S R E V P E L L K F T C N A T H E L V V -
c V V K C Q N Y S S L R V M R P M S W L L -
RsaI
EciI
AciI
BsmAI
BceAI
MnlI
Csp6I
MseI
TaqI
SspI
TAGAACACCTCGTAGTGTCCGCCGTTTGTCTCGTACCATTAAGGGTGTCGAATATTTTGA
5221
5280

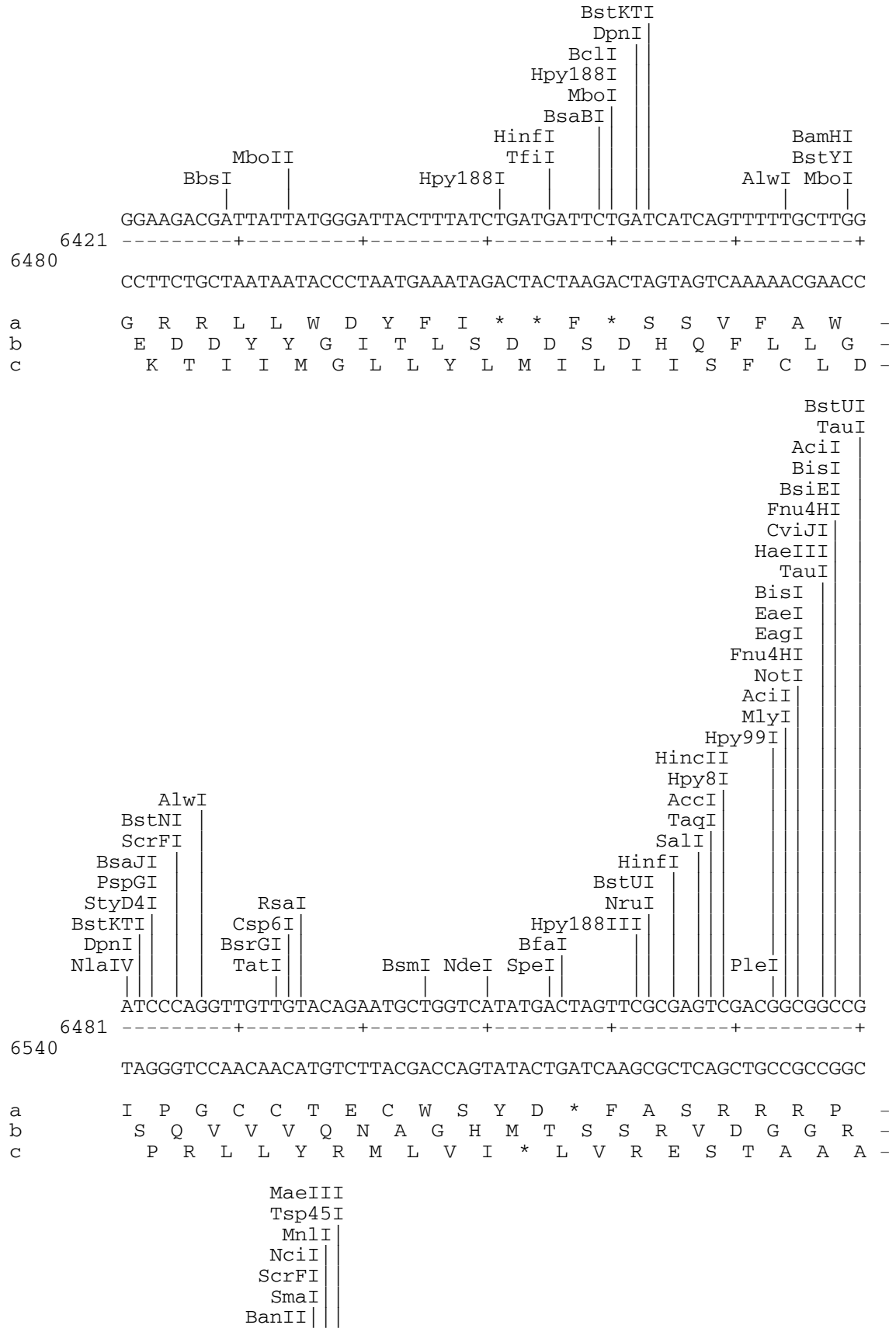
```
         ATCTTGTGGAGCATCACAGGCGGCAAACAGAGCATGGTAATTCCCACAGCTTATAAAACT

a        *  N  T  S  *  C  P  P  F  V  S  Y  H  *  G  C  R  I  F  *   -
b         R  T  P  R  S  V  R  R  L  S  R  T  I  K  G  V  E  Y  F  E  -
c          E  H  L  V  V  S  A  V  C  L  V  P  L  R  V  S  N  I  L  K -

                                                  Tsp509I
                          CviJI               Hpy99I   |
                         HaeIII            HpyCH4III   |      AluI
                BccI   Sau96I |      CviJI        |   |     CviJI
                   |        | |          |        |   |         |
         AGTTATTACTTTTGAGATGGGCCAAAAGAAAGCCCCCGACGGTAGAATTGTTGAGCTTGT
    5281 ---------+---------+---------+---------+---------+---------+
5340
         TCAATAATGAAAACTCTACCCGGTTTTCTTTCGGGGGCTGCCATCTTAACAACTCGAACA

a        S  Y  Y  F  *  D  G  P  K  E  S  P  R  R  *  N  C  *  A  C   -
b         V  I  T  F  E  M  G  Q  K  K  A  P  D  G  R  I  V  E  L  V  -
c          L  L  L  L  R  W  A  K  R  K  P  P  T  V  E  L  L  S  L  S -

                                                DdeI
                                             CviJI |
                                            HaeIII |
                                            NlaIV| |
                                        EcoO109I|| |
                                          Sau96I|| |
                                      BspCNI   ||| |
                                     Hpy188I   ||| |
                                     BseMII|   ||| |
                                 MnlI    ||   ||| |
                          BspCNI  |    ||   ||| |
                         BseMII|  |    ||   ||| |
                        AluI  ||  |    ||   ||| |       Tsp509I
                       CviJI  ||  |  DdeI|   ||| |   CviJI     |
                          |  ||  |    ||   ||| |       |     |
         CAAGGAAGTTTCAAAGAGCTACCCAATATCTGAGGGGCCTGAGAGAGCCAACGAATTAGT
    5341 ---------+---------+---------+---------+---------+---------+
5400
         GTTCCTTCAAAGTTTCTCGATGGGTTATAGACTCCCCGGACTCTCTCGGTTGCTTAATCA

a        Q  G  S  F  K  E  L  P  N  I  *  G  A  *  E  S  Q  R  I  S   -
b         K  E  V  S  K  S  Y  P  I  S  E  G  P  E  R  A  N  E  L  V  -
c          R  K  F  Q  R  A  T  Q  Y  L  R  G  L  R  E  P  T  N  *  * -

                                                                 DpnI
                                                              BglII |
                                   AluI                       BstYI |
      HinfI          CviJI       CviJI          MnlI     CviJI    | |
       TfiI SfcI     XmnI|   HindIII |        Hpy8I|    HaeIII MboI |
          |    |        ||         | |            ||         |    | |
         AGAATCCTATAGAAAGGCTTCAAATAAAGCTTATTTTGAGTGGACTATTGAGGCCAGAGA
    5401 ---------+---------+---------+---------+---------+---------+
5460
         TCTTAGGATATCTTTCCGAAGTTTATTTCGAATAAAACTCACCTGATAACTCCGGTCTCT

a        R  I  L  *  K  G  F  K  *  S  L  F  *  V  D  Y  *  G  Q  R   -
b         E  S  Y  R  K  A  S  N  K  A  Y  F  E  W  T  I  E  A  R  D  -
c          N  P  I  E  R  L  Q  I  K  L  I  L  S  G  L  L  R  P  E  I -

                           NlaIII      BsaXI
                        CviAII  |      Hin4I
                         FatI|  |     AluI |
    BstKTI          NlaIV  ||  |    CviJI |             Tsp509I
         |              |  ||  |        | |                  |
```

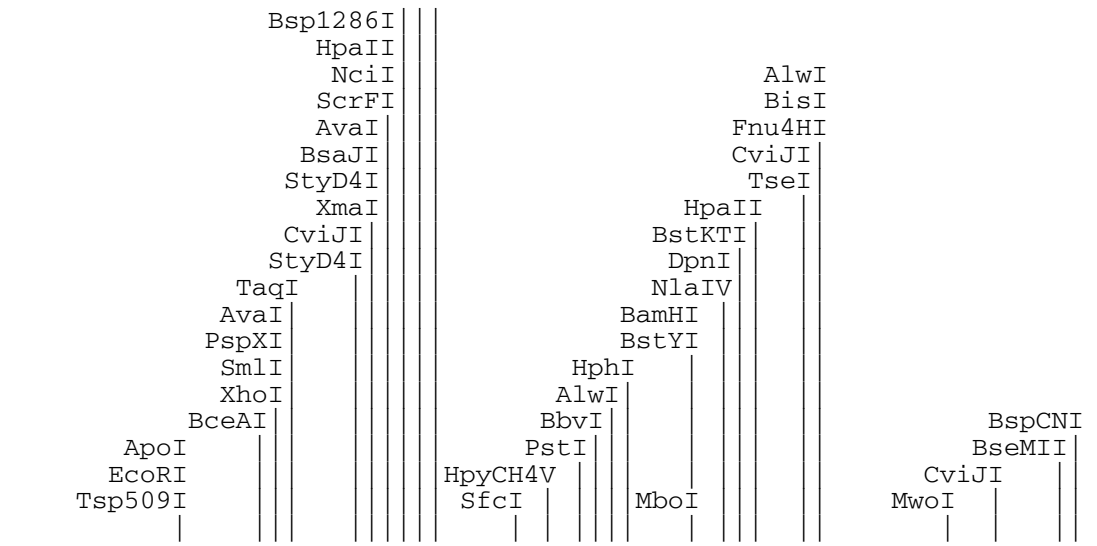

TCTTTCTCTGTTGGGTTCCCATGTTCGTAAAGCTACCTACCAGACTTACGCTCCAATTCT
5461
5520
AGAAAGAGACAACCCAAGGGTACAAGCATTTCGATGGATGGTCTGAATGCGAGGTTAAGA
a S F S V G F P C S * S Y L P D L R S N S -
b L S L L G S H V R K A T Y Q T Y A P I L -
c F L C W V P M F V K L P T R L T L Q F F -
TspDTI
HpyCH4V
NlaIII
NspI
Hin4I
CviAII
BsaXI
TaqI
FatI
HphI
TTATGAGAATGACCACTTTTTCGACTACATGCAAAAAAGTAAGTTTCATCTCACCATTGA
5521
5580
AATACTCTTACTGGTGAAAAAGCTGATGTACGTTTTTTCATTCAAAGTAGAGTGGTAACT
a L * E * P L F R L H A K K * V S S H H * -
b Y E N D H F F D Y M Q K S K F H L T I E -
c M R M T T F S T T C K K V S F I S P L K -
RsaI
ScaI
Hpy188I
AvaII
Csp6I
BccI
HphI
Sau96I
TatI
Hin4I
BcgI
AGGTCCAAAAGTACTTGCTTATTTACTTGGTTTATGGATTGGTGATGGATTGTCTGACAG
5581
5640
TCCAGGTTTTCATGAACGAATAAATGAACCAAATACCTAACCACTACCTAACAGACTGTC
a R S K S T C L F T W F M D W * W I V * Q -
b G P K V L A Y L L G L W I G D G L S D R -
c V Q K Y L L I Y L V Y G L V M D C L T G -
MaeIII
TaiI
Hpy188III
HinfI
AflIII
Hin4I
TfiI
BcgI
BccI
HpyCH4IV
GGCAACTTTTTCGGTTGATTCCAGAGATACTTCTTTGATGGAACGTGTTACTGAATATGC
5641
5700
CCGTTGAAAAAGCCAACTAAGGTCTCTATGAAGAAACTACCTTGCACAATGACTTATACG
a G N F F G * F Q R Y F F D G T C Y * I C -
b A T F S V D S R D T S L M E R V T E Y A -
c Q L F R L I P E I L L * W N V L L N M L -
ApoI
HhaI
Tsp509I
HinP1I
TGAAAAGTTGAATTTGTGCGCCGAGTATAAGGACAGAAAAGAACCACAAGTTGCCAAAAC
5701
5760
ACTTTTCAACTTAAACACGCGGCTCATATTCCTGTCTTTTCTTGGTGTTCAACGGTTTTG
a * K V E F V R R V * G Q K R T T S C Q N -
b E K L N L C A E Y K D R K E P Q V A K T -
c K S * I C A P S I R T E K N H K L P K L -

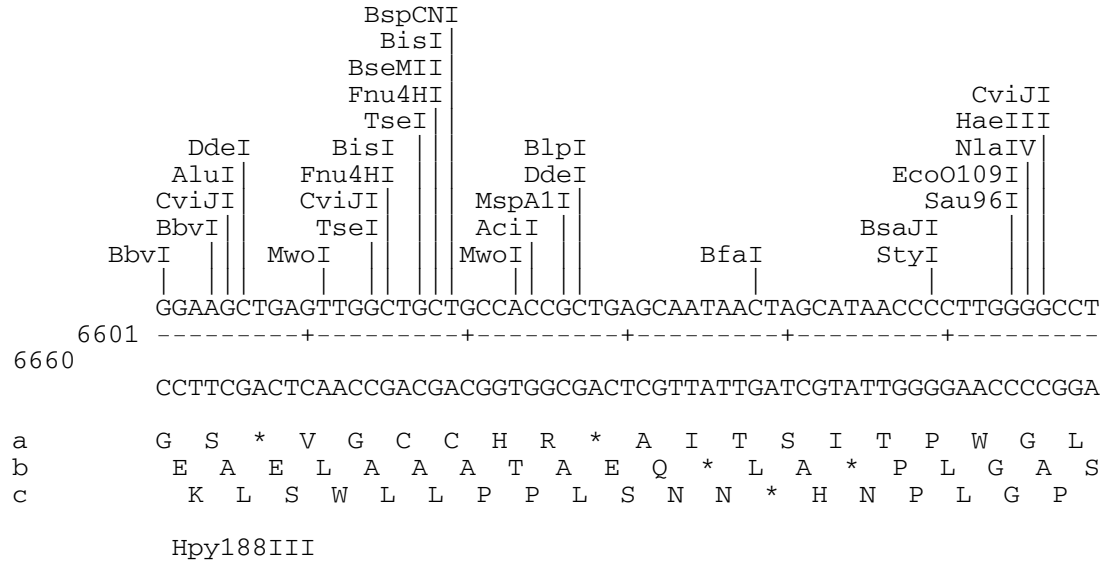

BmtI BstNI
Tsp509I RsaI Cac8I ScrFI
MseI Csp6I Hpy188I BfaI PspGI AciI
HpyCH4III TatI MnlI NheI StyD4I BspMI
TGTTAATTTGTACTCTAAAGTTGTCAGAGGTGCTAGCACAAATCCTGGTGTATCCGCTTG
5761
5820
ACAATTAAACATGAGATTTCAACAGTCTCCACGATCGTGTTTAGGACCACATAGGCGAAC
a C * F V L * S C Q R C * H K S W C I R L -
b V N L Y S K V V R G A S T N P G V S A W -
c L I C T L K L S E V L A Q I L V Y P L G -
MaeIII
Tsp45I
AhdI
MfeI
Tsp509I
AciI
MwoI
Hin4I
FauI
TaiI
HincII AluI HpyCH4IV
BciVI Hpy8I CviJI BsmFI Hin4I
GCAGGTCAACACAGCTTATACTGCGGGACAATTGGTCACATATAACGGCAAGACGTATAA
5821
5880
CGTCCAGTTGTGTCGAATATGACGCCCTGTTAACCAGTGTATATTGCCGTTCTGCATATT
a A G Q H S L Y C G T I G H I * R Q D V * -
b Q V N T A Y T A G Q L V T Y N G K T Y K -
c R S T Q L I L R D N W S H I T A R R I N -
TaiI
MnlI BccI
CviJI FokI AclI
BisI BccI HpyCH4IV
Fnu4HI BsaJI BstF5I
HpyCH4V StyI NlaIV
BceAI TseI BbvI BstF5I FokI
ATGTTTGCAGCCCCACACCTCCTTGGCAGGATGGGAACCATCCAACGTTCCTGCCTTGTG
5881
5940
TACAAACGTCGGGGTGTGGAGGAACCGTCCTACCCTTGGTAGGTTGCAAGGACGGAACAC
a M F A A P H L L G R M G T I Q R S C L V -
b C L Q P H T S L A G W E P S N V P A L W -
c V C S P T P P W Q D G N H P T F L P C G -
NlaIII
CviAII
BsaJI
BtgI
FatI
NcoI
MmeI StyI
AluI CviJI
CviJI HaeIII
BisI MscI MseI HinfI
Fnu4HI BbvI BspCNI TfiI

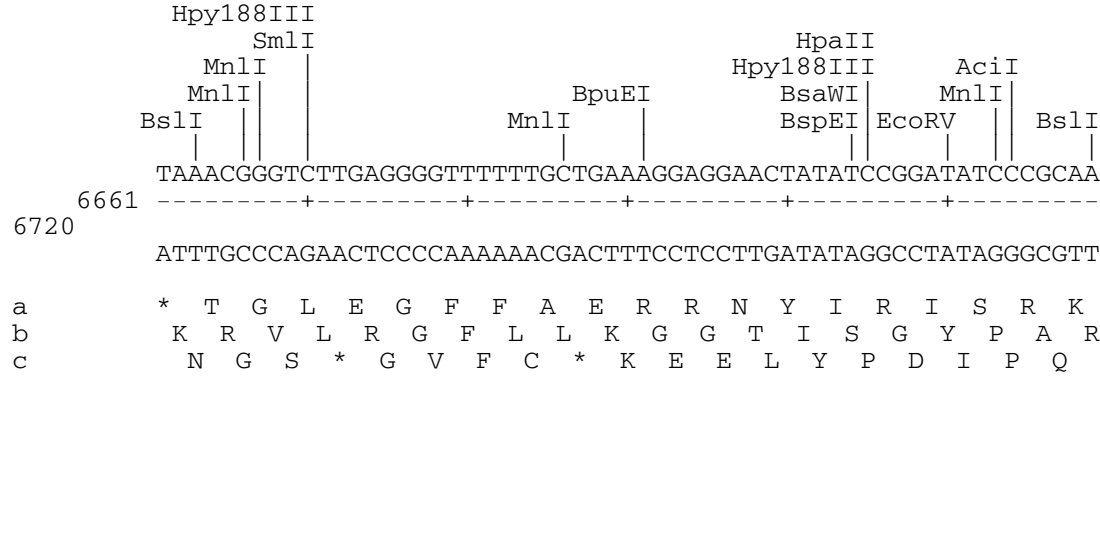

TseI
EaeI
BseMII
DdeI
GCAGCTTCAAGGTGGCCATGGTGGTATTCGCAATAATCTTAATACTGAGAATCCATTATG
5941
6000
CGTCGAAGTTCCACCGGTACCACCATAAGCGTTATTAGAATTATGACTCTTAGGTAATAC
a A A S R W P W W Y S Q * S * Y * E S I M -
b Q L Q G G H G G I R N N L N T E N P L W -
c S F K V A M V V F A I I L I L R I H Y G -
Bpu10I
DdeI
CviJI
Hpy188III
BsmFI
HinfI
HpyCH4III
HgaI
TfiI
Tth111I
SspI
GGACGCTATTGTTGGCTTAGGATTCTTGAAGGACGGTGTCAAAAATATTCCTTCTTTCTT
6001
6060
CCTGCGATAACAACCGAATCCTAAGAACTTCCTGCCACAGTTTTTATAAGGAAGAAAGAA
a G R Y C W L R I L E G R C Q K Y S F F L -
b D A I V G L G F L K D G V K N I P S F L -
c T L L L A * D S * R T V S K I F L L S C -
Hpy188III
BssSI
TspGWI
RsaI
Csp6I
Hpy8I
AccI
XmnI
Hpy188I
BccI
HinfI
Tsp509I
TfiI
CviJI
GTCTACGGACAATATCGGTACTCGTGAAACATTTCTTGCTGGTCTAATTGATTCTGATGG
6061
6120
CAGATGCCTGTTATAGCCATGAGCACTTTGTAAAGAACGACCAGATTAACTAAGACTACC
a V Y G Q Y R Y S * N I S C W S N * F * W -
b S T D N I G T R E T F L A G L I D S D G -
c L R T I S V L V K H F L L V * L I L M A -
NlaIII
CviAII
FatI
MseI
TspDTI
Tsp509I
Hpy188I
BccI
MaeIII
CTATGTTACTGATGAGCATGGTATTAAAGCAACAATAAAGACAATTCATACTTCTGTCAG
6121
6180
GATACAATGACTACTCGTACCATAATTTCGTTGTTATTTCTGTTAAGTATGAAGACAGTC
a L C Y * * A W Y * S N N K D N S Y F C Q -
b Y V T D E H G I K A T I K T I H T S V R -
c M L L M S M V L K Q Q * R Q F I L L S E -
HincII
HpaI
Hpy8I
AlwNI
DdeI
MseI
BstAPI
CviJI
BsmAI
MwoI
AGATGGTTTGGTTTCCCTTGCTCGTTCTTTAGGCTTAGTAGTCTCGGTTAACGCAGAACC
6181
6240
TCTACCAAACCAAAGGGAACGAGCAAGAAATCCGAATCATCAGAGCCAATTGCGTCTTGG

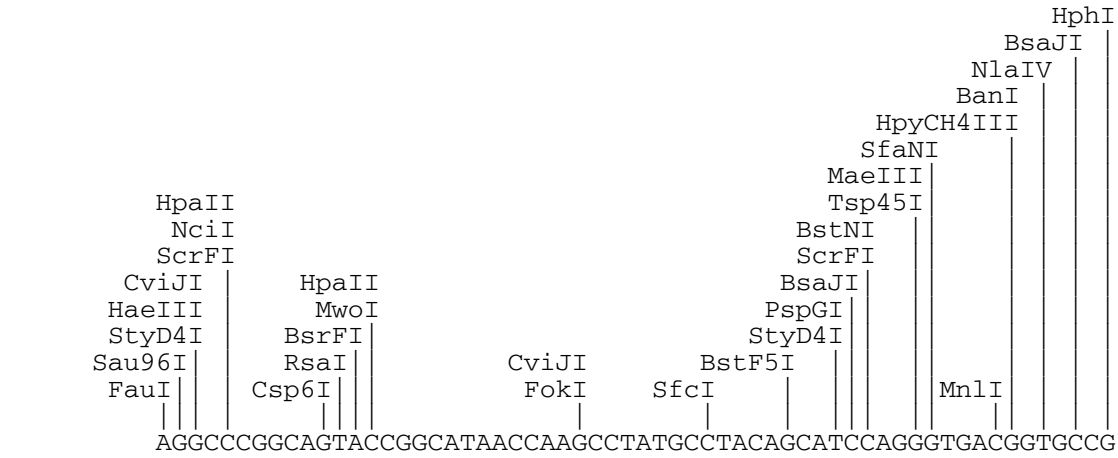

a R W F G F P C S F F R L S S L G * R R T -
b D G L V S L A R S L G L V V S V N A E P -
c M V W F P L L V L * A * * S R L T Q N L -
MaeIII
Tsp45I
DrdI
NlaIII
CviAII
HphI
FatI
HincII
Hpy8I
BspMI
Bpu10I
DdeI
Tsp509I
TspDTI
TGCTAAGGTTGACATGAATGTCACCAAACATAAAATTAGTTATGCTATTTATATGTCTGG
6241 ---------+---------+---------+---------+---------+---------+
6300
ACGATTCCAACTGTACTTACAGTGGTTTGTATTTTAATCAATACGATAAATATACAGACC
a C * G * H E C H Q T * N * L C Y L Y V W -
b A K V D M N V T K H K I S Y A I Y M S G -
c L R L T * M S P N I K L V M L F I C L V -
CviJI
Cac8I
Cac8I
TaiI
NaeI
BbvI
AclI
HpaII
CviJI
HpyCH4IV
BstBI
BsrFI
ApoI
HaeIII
MseI
TaqI
NgoMIV
Tsp509I
StuI
TGGAGATGTTTTGCTTAACGTTCTTTCGAAGTGTGCCGGCTCTAAAAAATTCAGGCCTGC
6301 ---------+---------+---------+---------+---------+---------+
6360
ACCTCTACAAAACGAATTGCAAGAAAGCTTCACACGGCCGAGATTTTTTAAGTCCGGACG
a W R C F A * R S F E V C R L * K I Q A C -
b G D V L L N V L S K C A G S K K F R P A -
c E M F C L T F F R S V P A L K N S G L L -
FauI
BisI
SacII
Fnu4HI
AciI
MspA1I
AleI
BstUI
TauI
MslI
MspA1I
TseI
TaiI
TauI
AciI
BsaAI
AciI
BisI
PmlI
BisI
Fnu4HI
HpyCH4IV
BsaJI
MaeIII
MwoI
HpyCH4V
BtgI
TaqI
AciI
MwoI
Fnu4HI
AlfI
Tsp509I
TCCCGCCGCTGCTTTTGCACGTGAGTGCCGCGGATTTTATTTCGAGTTACAAGAATTGAA
6361 ---------+---------+---------+---------+---------+---------+
6420
AGGGCGGCGACGAAAACGTGCACTCACGGCGCCTAAAATAAAGCTCAATGTTCTTAACTT
a S R R C F C T * V P R I L F R V T R I E -
b P A A A F A R E C R G F Y F E L Q E L K -
c P P L L L H V S A A D F I S S Y K N * R -

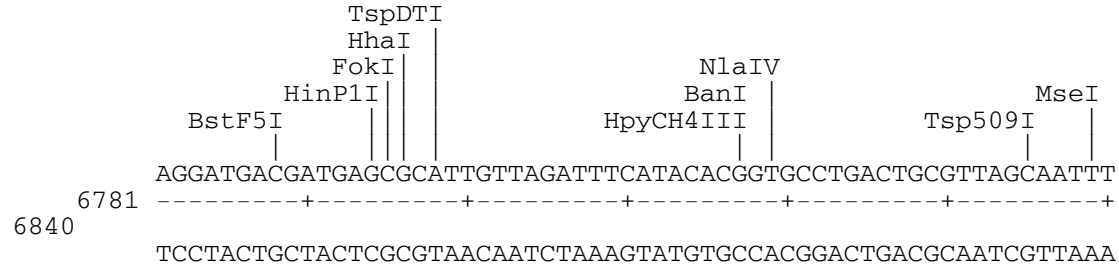

BstKTI
DpnI
BclI
Hpy188I
MboI
BsaBI
HinfI
TfiI
BamHI
BstYI
MboII
BbsI
Hpy188I
AlwI
MboI
GGAAGACGATTATTATGGGATTACTTTATCTGATGATTCTGATCATCAGTTTTTGCTTGG
6421
6480
CCTTCTGCTAATAATACCCTAATGAAATAGACTACTAAGACTAGTAGTCAAAAACGAACC
a G R R L L W D Y F I * * F * S S V F A W -
b E D D Y Y G I T L S D D S D H Q F L L G -
c K T I I M G L L Y L M I L I I S F C L D -
BstUI
TauI
AciI
BisI
BsiEI
Fnu4HI
CviJI
HaeIII
TauI
BisI
EaeI
EagI
Fnu4HI
NotI
AciI
MlyI
Hpy99I
HincII
Hpy8I
AccI
TaqI
SalI
HinfI
BstUI
NruI
Hpy188III
BfaI
PleI
AlwI
BstNI
ScrFI
BsaJI
PspGI
StyD4I
BstKTI
DpnI
NlaIV
RsaI
Csp6I
BsrGI
TatI
BsmI
NdeI
SpeI
ATCCCAGGTTGTTGTACAGAATGCTGGTCATATGACTAGTTCGCGAGTCGACGGCGGCCG
6481
6540
TAGGGTCCAACAACATGTCTTACGACCAGTATACTGATCAAGCGCTCAGCTGCCGCCGGC
a I P G C C T E C W S Y D * F A S R R R P -
b S Q V V V Q N A G H M T S S R V D G G R -
c P R L L Y R M L V I * L V R E S T A A A -
MaeIII
Tsp45I
MnlI
NciI
ScrFI
SmaI
BanII

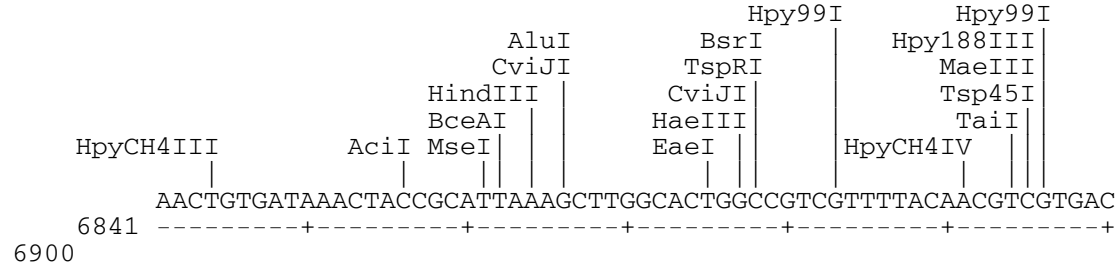

Bsp1286I
HpaII
NciI
ScrFI
AvaI
BsaJI
StyD4I
XmaI
CviJI
StyD4I
TaqI
AvaI
PspXI
SmlI
XhoI
BceAI
ApoI
EcoRI
Tsp509I
AlwI
BisI
Fnu4HI
CviJI
TseI
HpaII
BstKTI
DpnI
NlaIV
BamHI
BstYI
HphI
AlwI
BbvI
PstI
HpyCH4V
SfcI
MboI
BspCNI
BseMII
CviJI
MwoI
CGAATTCCTCGAGCCCGGGTGACTGCAGGAAGGGGATCCGGCTGCTAACAAAGCCCGAAA
6541
6600
GCTTAAGGAGCTCGGGCCCACTGACGTCCTTCCCCTAGGCCGACGATTGTTTCGGGCTTT
a R I P R A R V T A G R G S G C * Q S P K -
b E F L E P G * L Q E G D P A A N K A R K -
c N S S S P G D C R K G I R L L T K P E R -

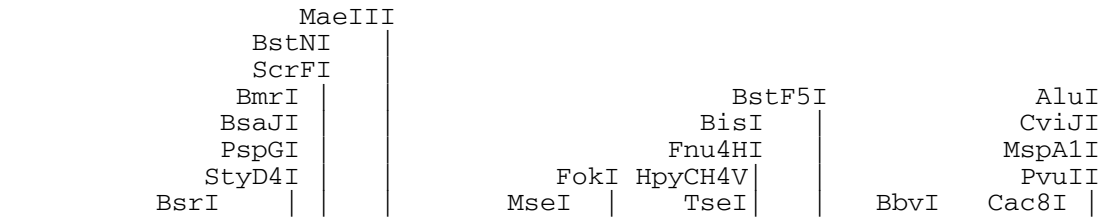

BspCNI
BisI
BseMII
Fnu4HI
TseI
CviJI
HaeIII
NlaIV
EcoO109I
Sau96I
BsaJI
StyI
DdeI
AluI
CviJI
BbvI
BbvI
BisI
Fnu4HI
CviJI
TseI
MwoI
BlpI
DdeI
MspA1I
AciI
MwoI
BfaI
GGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTC
6601
6660
CCTTCGACTCAACCGACGACGGTGGCGACTCGTTATTGATCGTATTGGGGAACCCCGGAG
a G S * V G C C H R * A I T S I T P W G L -
b E A E L A A A T A E Q * L A * P L G A S -
c K L S W L L P P L S N N * H N P L G P L -


Hpy188III
SmlI
MnlI
MnlI
BslI
BpuEI
MnlI
HpaII
Hpy188III
BsaWI
BspEI
AciI
MnlI
EcoRV
BslI
TAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCCGCAAG
6661
6720
ATTTGCCCAGAACTCCCCAAAAAACGACTTTCCTCCTTGATATAGGCCTATAGGGCGTTC
a * T G L E G F F A E R R N Y I R I S R K -
b K R V L R G F L L K G G T I S G Y P A R -
c N G S * G V F C * K E E L Y P D I P Q E -

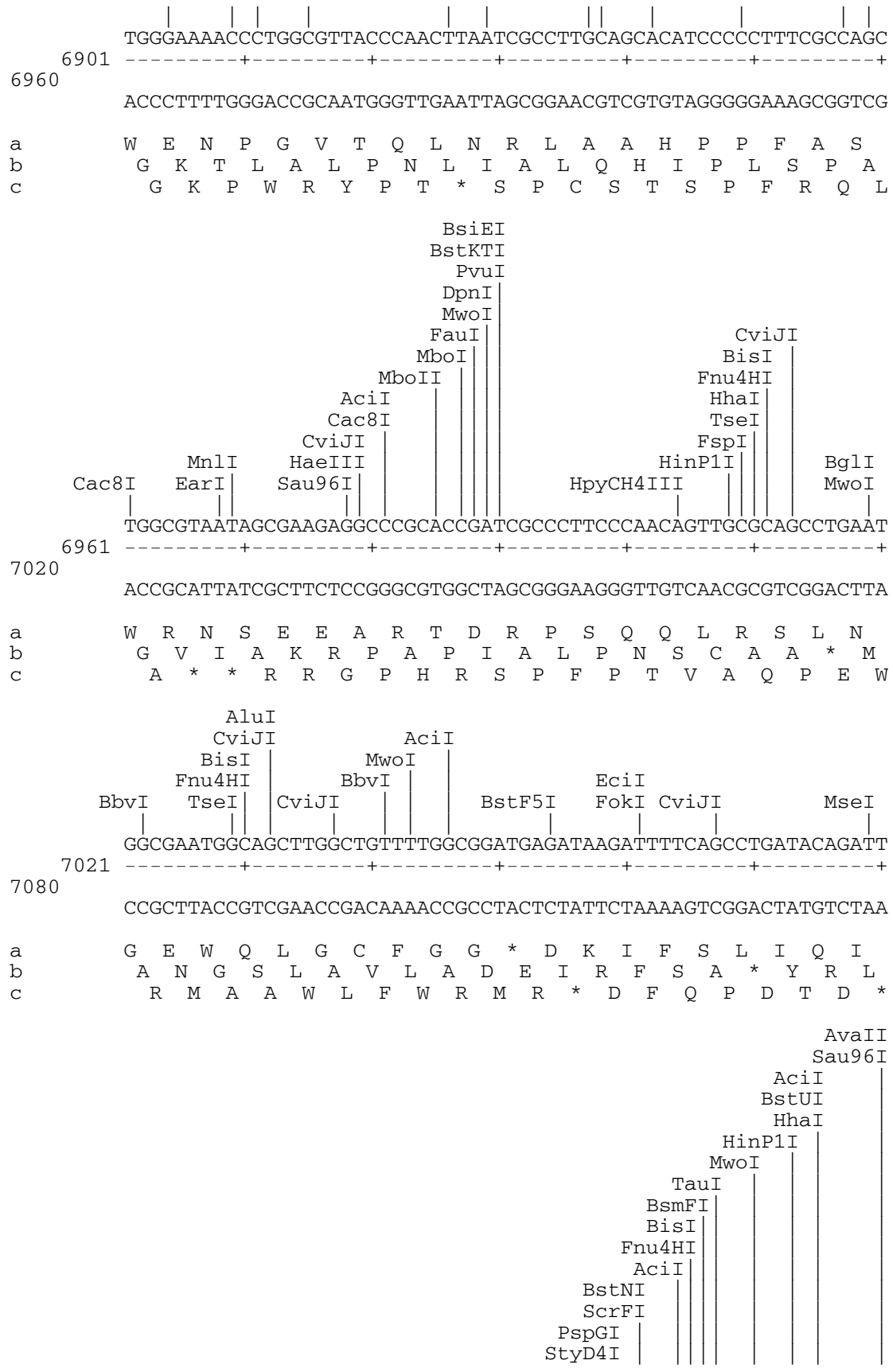


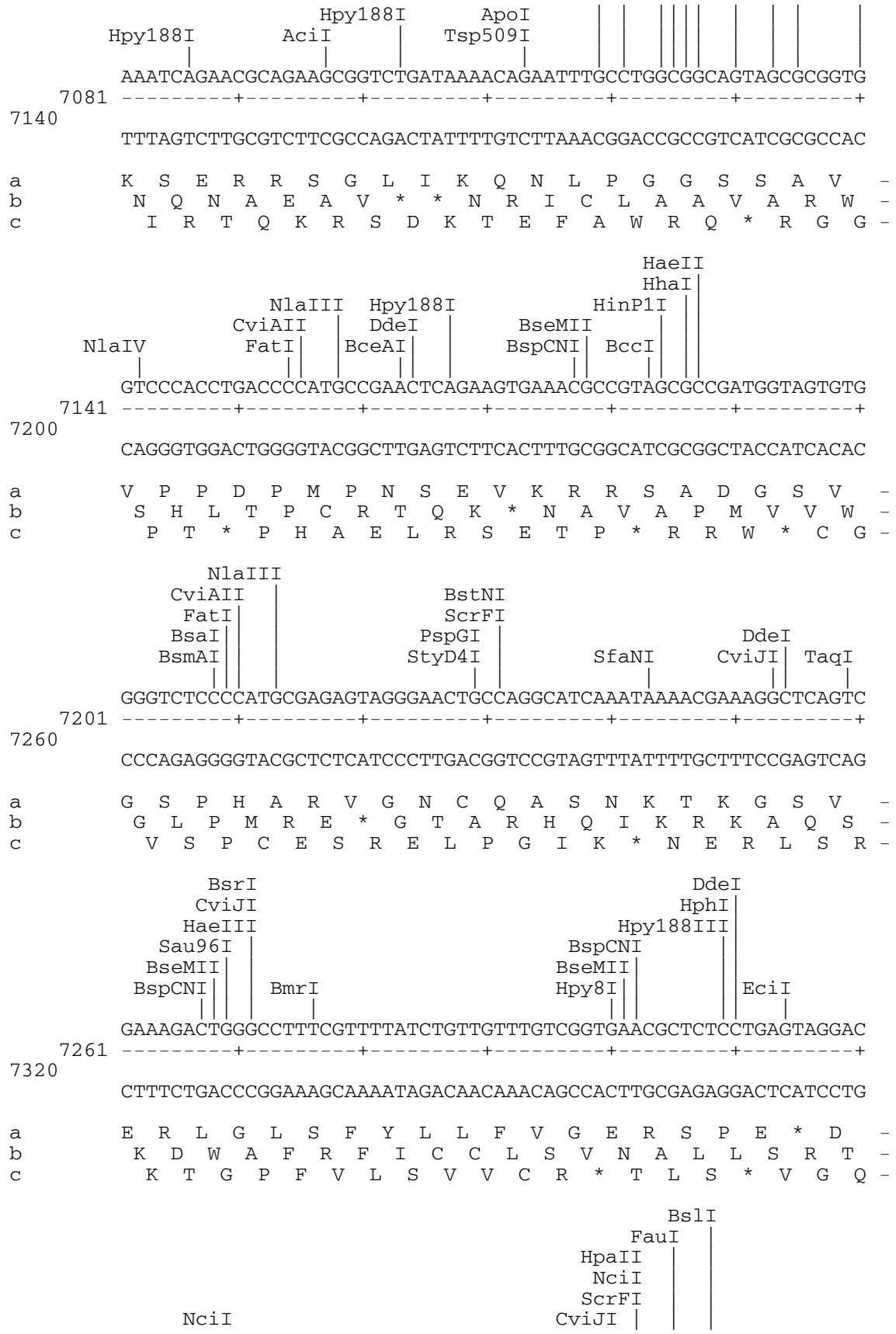


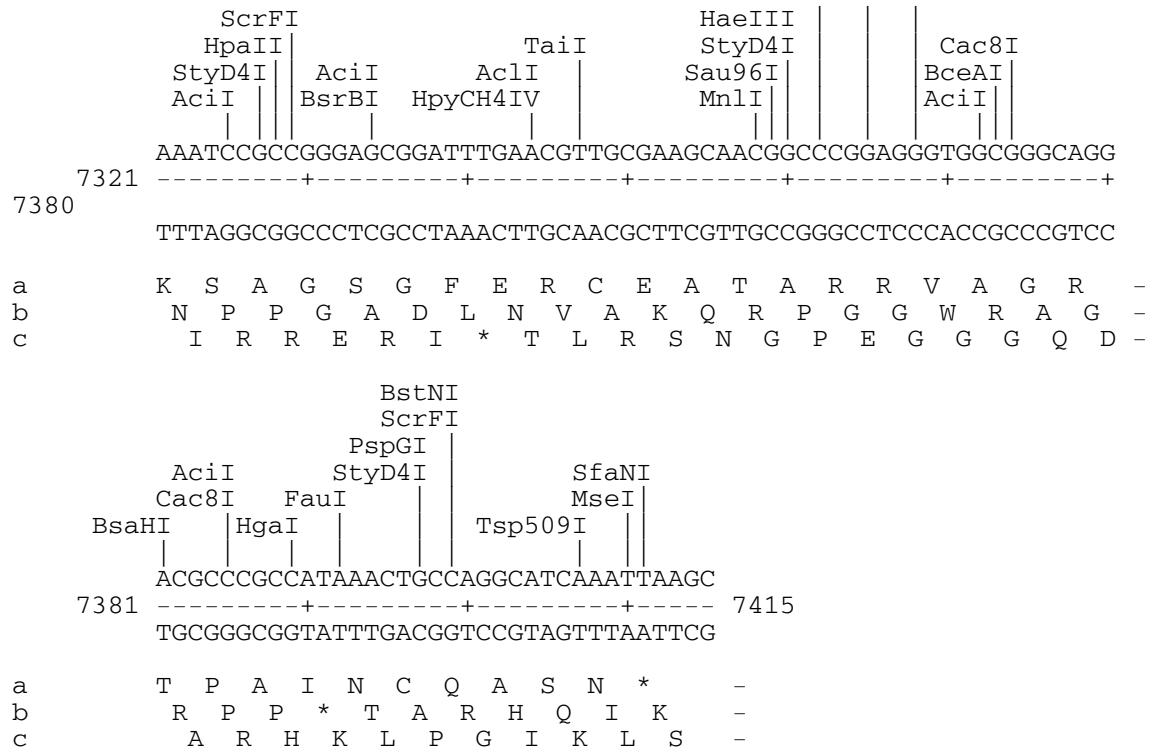




```
                                                                                HphI
                                                                             BsaJI |
                                                                          NlaIV | |
                                                                         BanI | | |
                                                                   HpyCH4III | | |
                                                                  SfaNI    | | | |
                                                               MaeIII|    | | | |
        HpaII                                                  Tsp45I|    | | | |
         NciI                                                BstNI  ||    | | | |
        ScrFI                                                ScrFI  ||    | | | |
      CviJI |    HpaII                                     BsaJI|  ||    | | | |
     HaeIII |     MwoI                                   PspGI||  ||    | | | |
     StyD4I |   BsrFI|                                  StyD4I||  ||    | | | |
    Sau96I| |   RsaI||        CviJI       BstF5I  |||  ||    | | | |
     FauI|| | Csp6I|||         FokI    SfcI    |  |||  ||MnlI| | | |
        ||| |     ||||            |       |    |  |||  ||   || | | |
        AGGCCCGGCAGTACCGGCATAACCAAGCCTATGCCTACAGCATCCAGGGTGACGGTGCCG
    6721 ---------+---------+---------+---------+---------+---------+
6780
        TCCGGGCCGTCATGGCCGTATTGGTTCGGATACGGATGTCGTAGGTCCCACTGCCACGGC

a       R  P  G  S  T  G  I  T  K  P  M  P  T  A  S  R  V  T  V  P   -
b        G  P  A  V  P  A  *  P  S  L  C  L  Q  H  P  G  *  R  C  R  -
c         A  R  Q  Y  R  H  N  Q  A  Y  A  Y  S  I  Q  G  D  G  A  E -

                    TspDTI
                    HhaI |
                   FokI| |                NlaIV
                HinP1I|| |               BanI |                MseI
          BstF5I     ||| |          HpyCH4III |         Tsp509I   |
               |     ||| |                  | |                |   |
        AGGATGACGATGAGCGCATTGTTAGATTTCATACACGGTGCCTGACTGCGTTAGCAATTT
    6781 ---------+---------+---------+---------+---------+---------+
6840
        TCCTACTGCTACTCGCGTAACAATCTAAAGTATGTGCCACGGACTGACGCAATCGTTAAA

a       R  M  T  M  S  A  L  L  D  F  I  H  G  A  *  L  R  *  Q  F   -
b        G  *  R  *  A  H  C  *  I  S  Y  T  V  P  D  C  V  S  N  L  -
c         D  D  D  E  R  I  V  R  F  H  T  R  C  L  T  A  L  A  I  * -

                                                  Hpy99I       Hpy99I
                                AluI        BsrI    |   Hpy188III|
                               CviJI       TspRI    |      MaeIII|
                          HindIII |      CviJI|    |      Tsp45I|
                          BceAI | |     HaeIII|    |       TaiI||
    HpyCH4III        AciI MseI| | |     EaeI ||    |HpyCH4IV  |||
            |            |    || | |        | ||    |       |  |||
        AACTGTGATAAACTACCGCATTAAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGAC
    6841 ---------+---------+---------+---------+---------+---------+
6900
        TTGACACTATTTGATGGCGTAATTTCGAACCGTGACCGGCAGCAAAATGTTGCAGCACTG

a       N  C  D  K  L  P  H  *  S  L  A  L  A  V  V  L  Q  R  R  D   -
b        T  V  I  N  Y  R  I  K  A  W  H  W  P  S  F  Y  N  V  V  T  -
c         L  *  *  T  T  A  L  K  L  G  T  G  R  R  F  T  T  S  *  L -

                 MaeIII
              BstNI   |
              ScrFI   |
             BmrI |   |                     BstF5I             AluI
            BsaJI |   |                   BisI   |            CviJI
            PspGI |   |                 Fnu4HI   |           MspA1I
           StyD4I |   |           FokI HpyCH4V|   |            PvuII
        BsrI    | |   |        MseI  |    TseI|   |   BbvI   Cac8I |
```


TGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC
6901 ---------+---------+---------+---------+---------+---------+
6960
ACCCTTTTGGGACCGCAATGGGTTGAATTAGCGGAACGTCGTGTAGGGGGAAAGCGGTCG
a W E N P G V T Q L N R L A A H P P F A S -
b G K T L A L P N L I A L Q H I P L S P A -
c G K P W R Y P T * S P C S T S P F R Q L -
BsiEI
BstKTI
PvuI
DpnI
MwoI
FauI
MboI
MboII
AciI
Cac8I
CviJI
MnlI
HaeIII
Cac8I
EarI
Sau96I
HpyCH4III
CviJI
BisI
Fnu4HI
HhaI
TseI
FspI
HinP1I
BglI
MwoI
TGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAAT
6961 ---------+---------+---------+---------+---------+---------+
7020
ACCGCATTATCGCTTCTCCGGGCGTGGCTAGCGGGAAGGGTTGTCAACGCGTCGGACTTA
a W R N S E E A R T D R P S Q Q L R S L N -
b G V I A K R P A P I A L P N S C A A * M -
c A * * R R G P H R S P F P T V A Q P E W -
AluI
CviJI
BisI
Fnu4HI
AciI
MwoI
BbvI
BbvI
TseI
CviJI
EciI
BstF5I
FokI
CviJI
MseI
GGCGAATGGCAGCTTGGCTGTTTTGGCGGATGAGATAAGATTTTCAGCCTGATACAGATT
7021 ---------+---------+---------+---------+---------+---------+
7080
CCGCTTACCGTCGAACCGACAAAACCGCCTACTCTATTCTAAAAGTCGGACTATGTCTAA
a G E W Q L G C F G G * D K I F S L I Q I -
b A N G S L A V L A D E I R F S A * Y R L -
c R M A A W L F W R M R * D F Q P D T D * -
AvaII
Sau96I
AciI
BstUI
HhaI
HinP1I
MwoI
TauI
BsmFI
BisI
Fnu4HI
AciI
BstNI
ScrFI
PspGI
StyD4I


Hpy188I
ApoI
Hpy188I
AciI
Tsp509I
AAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTG
7081
7140
TTTAGTCTTGCGTCTTCGCCAGACTATTTTGTCTTAAACGGACCGCCGTCATCGCGCCAC
a K S E R R S G L I K Q N L P G G S S A V -
b N Q N A E A V * * N R I C L A A V A R W -
c I R T Q K R S D K T E F A W R Q * R G G -
HaeII
HhaI
NlaIII
Hpy188I
HinP1I
CviAII
DdeI
BseMII
NlaIV
FatI
BceAI
BspCNI
BccI
GTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTG
7141
7200
CAGGGTGGACTGGGGTACGGCTTGAGTCTTCACTTTGCGGCATCGCGGCTACCATCACAC
a V P P D P M P N S E V K R R S A D G S V -
b S H L T P C R T Q K * N A V A P M V V W -
c P T * P H A E L R S E T P * R R W * C G -
NlaIII
CviAII
BstNI
FatI
ScrFI
BsaI
PspGI
DdeI
BsmAI
StyD4I
SfaNI
CviJI
TaqI
GGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTC
7201
7260
CCCAGAGGGGTACGCTCTCATCCCTTGACGGTCCGTAGTTTATTTTGCTTTCCGAGTCAG
a G S P H A R V G N C Q A S N K T K G S V -
b G L P M R E * G T A R H Q I K R K A Q S -
c V S P C E S R E L P G I K * N E R L S R -
BsrI
DdeI
CviJI
HphI
HaeIII
Hpy188III
Sau96I
BspCNI
BseMII
BseMII
BspCNI
Hpy8I
BmrI
EciI
GAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGAC
7261
7320
CTTTCTGACCCGGAAAGCAAAATAGACAACAAACAGCCACTTGCGAGAGGACTCATCCTG
a E R L G L S F Y L L F V G E R S P E * D -
b K D W A F R F I C C L S V N A L L S R T -
c K T G P F V L S V V C R * T L S * V G Q -
BslI
FauI
HpaII
NciI
ScrFI
NciI
CviJI



Enzymes that do cut:

AccI	Acc65I	AciI	AclI	AcuI	AfeI	AflIII	AhdI
AleI	AlfI	AloI	AluI	AlwI	AlwNI	ApaI	ApaLI
ApoI	AseI	AvaI	AvaII	BaeI	BamHI	BanI	BanII
BbeI	BbsI	BbvI	BccI	BceAI	BcgI	BciVI	BclI
BfaI	BglI	BglII	BisI	BlpI	Bme1580I	BmrI	BmtI
BpmI	Bpu10I	BpuEI	BsaI	BsaAI	BsaBI	BsaHI	BsaJI
BsaWI	BsaXI	BseMII	BseYI	BsgI	BsiEI	BsiHKAI	BslI
BsmI	BsmAI	BsmBI	BsmFI	Bsp1286I	BspCNI	BspEI	BspHI
BspMI	BsrI	BsrBI	BsrDI	BsrFI	BsrGI	BssHII	BssSI
BstAPI	BstBI	BstEII	BstF5I	BstKTI	BstNI	BstUI	BstXI
BstYI	BstZ17I	BtgI	BtgZI	BtsI	Cac8I	Csp6I	CviAII
CviJI	DdeI	DpnI	DraI	DraIII	DrdI	EaeI	EagI
EarI	EciI	Eco57MI	EcoNI	EcoO109I	EcoRI	EcoRV	FatI
FauI	Fnu4HI	FokI	FspI	HaeII	HaeIII	HgaI	HhaI
Hin4I	HinP1I	HincII	HindIII	HinfI	HpaI	HpaII	HphI
Hpy8I	Hpy99I	Hpy188I	Hpy188III	HpyCH4III	HpyCH4IV	HpyCH4V	KasI
KpnI	MaeIII	MboI	MboII	MfeI	MluI	MlyI	MmeI
MnlI	MscI	MseI	MslI	MspA1I	MwoI	NaeI	NarI
NciI	NcoI	NdeI	NgoMIV	NheI	NlaIII	NlaIV	NotI
NruI	NspI	PciI	PflMI	PfoI	PleI	PmlI	PpiI
PshAI	PsiI	PspGI	PspOMI	PspXI	PstI	PvuI	PvuII
RsaI	SacII	SalI	Sau96I	ScaI	ScrFI	SfaNI	SfcI
SfoI	SgrAI	SmaI	SmlI	SpeI	SphI	SspI	StuI
StyI	StyD4I	SwaI	TaiI	TaqI	TaqII	TatI	TauI
TfiI	TseI	Tsp45I	Tsp509I	TspDTI	TspGWI	TspRI	TstI
Tth111I	XbaI	XcmI	XhoI	XmaI	XmnI		

Enzymes that do not cut:

AarI	AatII	AflII	AgeI	AscI	AsiSI	AvrII	BbvCI
BfrBI	BmgBI	BplI	BseRI	BsiWI	Bsu36I	ClaI	CspCI
EcoICRI	FalI	FseI	FspAI	NsiI	PacI	PasI	PmeI

PpuMI PsrI RsrII SacI SanDI SapI SbfI SexAI
SfiI SnaBI SrfI ZraI