POTENTIAL FOR GENETIC ENHANCEMENT OF BACTERIAL DEGRADATIVE PROCESSES IN WASTEWATER

FINAL REPORT

TO

NATIONAL WATER RESEARCH INSTITUTE
AND
IRVINE RANCH WATER DISTRICT

RESEARCH PROJECT NUMBER RAN 91-05

Department of Environmental Analysis and Design,
and
Department of Biochemical Engineering

UNIVERSITY OF CALIFORNIA AT IRVINE
CA 92717

February 1994

POTENTIAL FOR ENHANCEMENT FOR BACTERIAL
DEGRADATIVE PROCESSES IN WASTEWATER

Administrative Report for the Period Ending

December 31, 1993

Principal Investigator: Betty Olson/Oladele Ogunseitan
Agency: National Water Research Institute
Total Award: $100,000 (47.5% Indirect Cost)
Project Period: January 1, 1992 - December 31, 1993
Agency Grant Number: RAN-91-05
UCI Account/Fund Number: 447670-51526

EXPENSE DETAIL

Expense Categories	Allocation	Expenses for December	Expended To Date	Projected Expenses	End of Project Balance
Salaries	$30,940	$0	$27,872	$0	$3,068
Supplies & Expense	$15,000	$162	$16,885	$0	($1,885)
Equipment	$18,587	$0	$20,729	$0	($2,142)
Travel	$3,000	$0	$1,000	$0	$2,000
Benefits	$5,570	$0	$6,628	$0	($1,058)
Other	$685	$0	$22	$0	$663
Indirect	$26,218	$77	$26,863	$0	($645)
Total	$100,000	$238	$100,000	$0	$0

EXECUTIVE SUMMARY

There were two active research periods during the conduction of this project. The initial period was March to December 1992 and the second period was July to December 1993. During the first phase, the primary objective of the project, namely the use of molecular and physiological methods to identify microorganisms carrying out denitrification and phosphorus accumulation in wastewater, was fulfilled. During the second period, the other major objective of the research, to clone the genes for nitrite reductase and polyphosphate kinase into high level expression vectors was also fulfilled. We are in possession of the recombinant *Escherichia coli* strains haboring the genes for denitrification and phosphorus removal from wastewater. However, because of time limitations and the break in project schedule, the third objective of the research, to evaluate the performance of the recombinant strains in bioreactors simulating the wastewater treatment process was not conducted to completion. Additional sources of funding to further the research to include scale-up applications are presently sought.

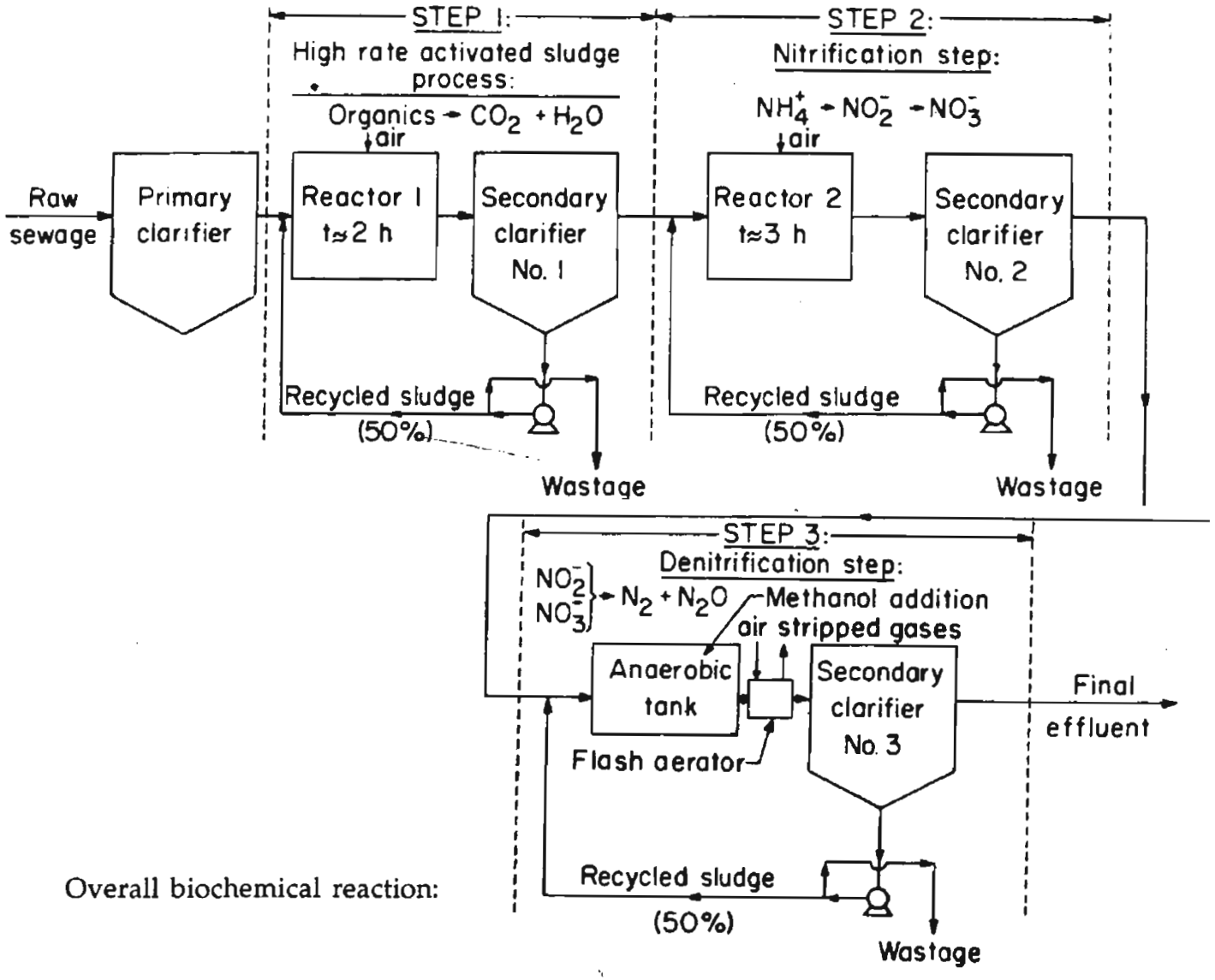
INTRODUCTION

The disposal and fate of nitrate (NO_3) and inorganic phosphorus (P_i) compounds in aquatic systems are important aspects of wastewater treatment and environmental water resources conservation. When the crucial processes affecting these nutrients are not controlled, excessive accumulation leads to eutrophication of creeks, lakes, and estuaries, with significant adverse ecological consequences (Ramalho, 1983). The removal of nitrate and phosphorus from wastewaters is expensive and usually conducted by manipulating aeration or by addition of chemicals such as lime, alum, and ferric chloride (Ramalho, 1983; Toerien *et al.*, 1990). Because of the expense and time consuming nature of traditional processes aimed at nitrate and phosphorus removal, biotechnological alternatives to the chemical methods have been under constant investigation, and are the subject of the present research effort.

Many organisms are known to be capable of nitrification and denitrification, and of phosphorus accumulation from aqueous media. The process of nitrate removal is tightly linked to the well studied nitrogen cycle consisting mainly of nitrification and denitrification reactions (Figure 1). The goal of biotechnological denitrification processes is to manipulate the direction and dynamics of the nitrogen cycle such that the formation of gaseous nitrogen compounds is favored. These compounds with sufficient partial pressures are then released into the atmospheric pool of nitrogenous compounds, thereby resulting in a lower concentration in the aqueous phase (i.e. dissimilatory nitrate reduction).

In the case of phosphorus removal from aquatic environments, biotechnological strategies depend on the ability of certain bacteria including

Figure 1. An example of a three-step process for nitrification and dissimilatory denitrification in a wastewater treatment process. (After Ramalho, 1983)



Overall biochemical reaction:

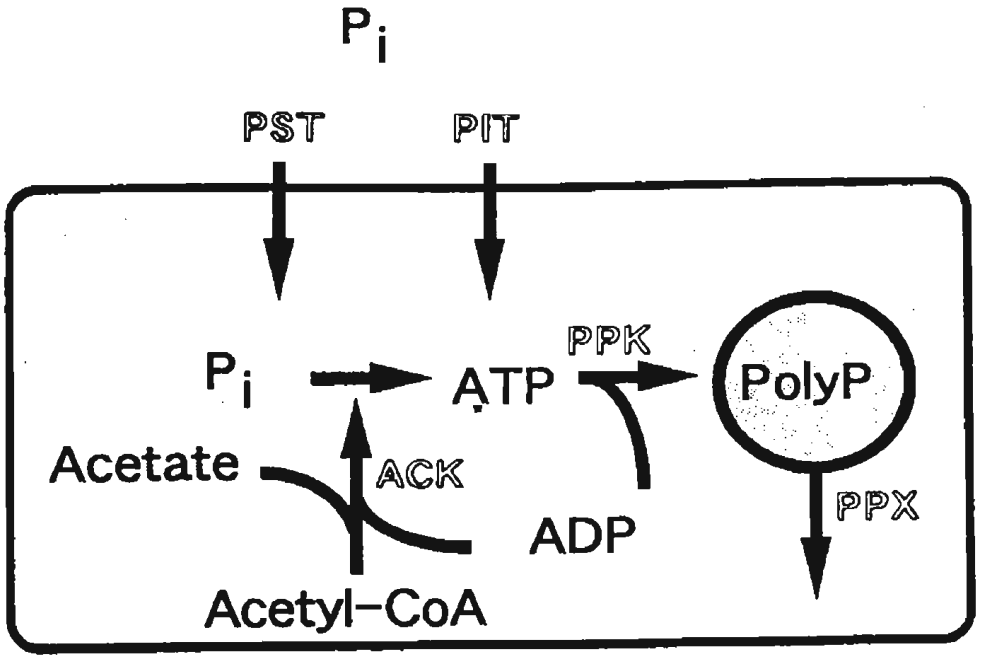
$2NO_3^-$ ----->4e- ---> $2NO_2^-$ ---2e- ---> 2NO ----->2e- ---> N_2O ----2e- --->N_2

(gas) (gas) (gas)

nitrate reductase	nitrite reductase	nitric oxide reductase	nitrous oxide reductase

Escherichia coli to accumulate polyphosphate granules in their cells during growth under certain nutritional conditions. The physiological functions of polyphosphate in cells are not completely understood, but there is sufficient evidence that it may serve as a source of energy because the reaction leading to its formation is reversible, leading to the production of adenosine triphosphate (ATP) which supports many biochemical reactions (Kato *et al.*, 1993). There are many enzymes involved in the accumulation and degradation of polyphosphate granules, but the one amenable to biotechnological manipulations is polyphosphate kinase (PPK; see Figure 2), and is the targeted enzyme in the study reported here.

Figure 2. Process by which phosphorus accumulation might be improved through genetic engineering of polyphosphate kinase. (From Kato *et al.*, 1993).



MATERIALS AND METHODS

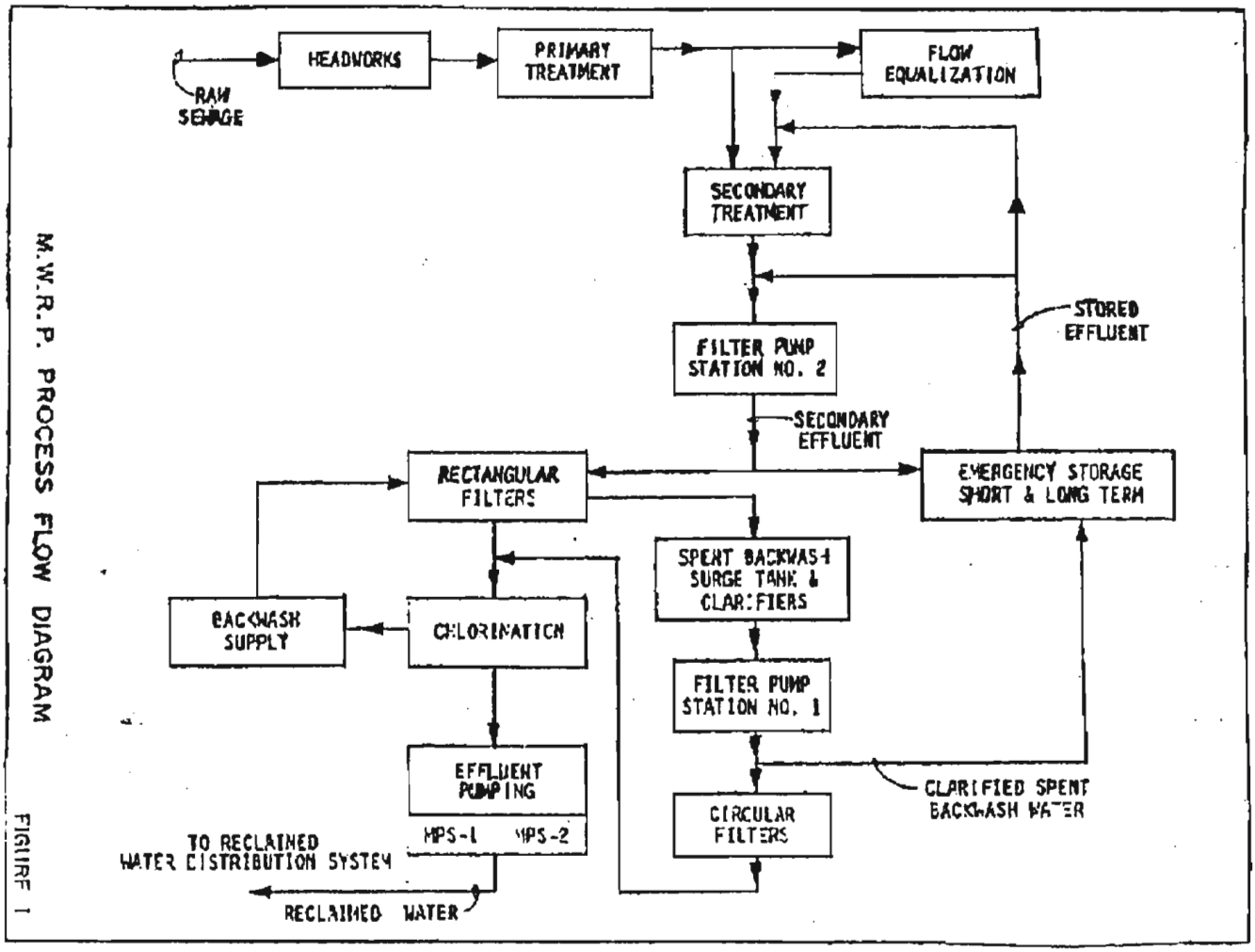
WASTEWATER SAMPLES

Throughout the project, wastewater samples were obtained from the Michelson wastewater treatment Plant operated by the Irvine Ranch Water District (IRWD). Samples were routinely collected from points representing the raw sewage inlet, primary and secondary settling outlflows and the treated effluent (Figure 3). One liter water samples were collected in polystyrene vessels and immediately placed on ice. Water samples were routinely analyzed within two hours of collection.

DENITRIFICATION

The first step taken for biotechnological enhancement of denitrification in wastewater was to determine the population densities of bacteria capable of carrying out denitrification (phenotypic assay), and of bacteria containing DNA sequences homologous to the segment cloned from *Pseudomonas stutzeri* encoding for nitrite reductase (Smith and Tiedje, 1992). The second phase involved the cloning of the nitrite reductase gene into a high expression level plasmid. The resulting recombinant plasmid was then transformed into *E. coli*, and the relative expression level of the cloned gene was checked through protein analysis by performing a denaturing polyacrlyamide gel electrophoresis (SDS-PAGE). The third planned phase of the project was comparative performance of the recombinant nitrite reductase-producing *E. coli* and and the control of the naturally-occuring denitrifying bacterial population through manipulation of aeration levels in bioreactors simulating wastewater treatment processes.

Figure 3. Sampling points in the Michelson wastewater treatment Plant.



PHOSPHORUS REMOVAL

The design of experiments aimed at biotechnological enhancement of phosphorus removal from wastewater was similar to those for denitrification. An initial assessment of phenotype (ability of wastewater bacteria to accumulate polyphosphate granules) through staining techniques and microscopy was conducted. The phenotype assay was followed by a genotype assay where wastewater bacteria containing DNA sequences homologous to the cloned fragment from *E. coli* encoding for polyphosphate kinase were detected through colony hybridization. The second phase of work regarding biotechnological phosphorus removal was the cloning of the *E. coli* polyphosphate gene (ppk) into a high level expression plamid for increased production of the crucial enzyme. Finally, the efficiency of phosphorus removal was to be compared between the recombinant strain containing the high level expression ppk-plasmid and the natural induction of polyphosphate accumulating potential of wastewater bacteria.

NEW METHOD DEVELOPMENT FOR MOLECULAR MICROBIAL ECOLOGY

In order to facilitate quantitative analysis of gene expression and consequential enzyme production in natural microbial communities, a method was developed for electrophoretic resolution of proteins extracted directly from natural or assembled microbial communities. The results of the method development and preliminary applications to monitoring polyphosphate kinase production was the first published article derived directly from the project (see appended manuscript).

RESULTS AND DISCUSSION

Tables 1 through 5 summarize the data on selected bacterial population densities in the wastewater samples investigated during this project. By comparing data on phenotypes with the data on DNA hybridization, the results indicate the possibility of increasing the expression of latent genetic potential. For example, in the primary effluent, 32% of the heterotrophic bacterial population contained DNA sequences homologous to the nitrite reductase gene, whereas only 15% of the bacteria in the same sample showed phenotypic ability to convert nitrate directly to gaseous products (Tables 2 and 3).

However, in the case of phosphorus accumulation the data indicate that there is significant diversity in the genetic potential responsible for the function of polyphosphate kinase in the wastewater microbial community. This is because, for example, 43.5% of total heterotrophic bacteria were able to accumulate phosphate granules according to the microscopic assay (Table 4), whereas only 1% of the bacteria in the same sample contained DNA sequences homologous the the *E. coli* polyphosphate kinase gene (Table 5).

These are important results because they indicate that the biotechnological strategy of choice for ehancing denitrification might be genetic induction of the latent genetic potential by means of aerobic-anaerobic system manipulation (e.g. with *E. coli* membranes) or with certain amino acids. That strategy will likely not work for phosphorus removal because the prototype genetic system does not account for the majority of phosphorus accumulating bacteria in the wastewater samples analyzed. Therefore, installing bioreactors containing genetically-engineered high expression leve *E. coli* strains containing the polyphosphate kinase gene might be a better strategy. We subsequently

decided to explore both strategies for denitrification and the reactor strategy for phosphorus removal.

TABLE 1. Aerobic heterotrophic bacterial population densities in IRWD Michelson Plant wastewater.

Wastewater Sample	Colony forming units (CFU) per ml
Raw influent	3.1 +/- 2.0 x 10^5
Primary effluent	2.6 +/- 1.2 x 10^5
Secondary effluent	9.6 +/- 3.2 x 10^4
Treated effluent	less than one

TABLE 2. Population densities of bacteria exhibiting denitrification phenotype in IRWD Michelson Plant wastewater.

Wastewater sample	CFU/ml	Proportion of total (%)
Raw influent	2.8 +/- 1.4 x 10^4	9.0
Primary effluent	3.9 +/-2.3 x 10^4	15.0
Secondary effluent	2.9 +/- 1.1 x 10^3	3.0
Treated effluent	not detected	

TABLE 3. Population densities of bacteria containing DNA sequences homologous to *E. coli* nitrite reductase gene in the IRWD Michelson Plant wastewater.

Wastewater sample	CFU/ml	Proportion of total (%)
Raw influent	1.6 +/- 0.8 x 10^4	37.0
Primary effluent	4.4 +/- 2.1 x 10^4	32.0
Secondary effluent	2.1 +/-1.2 x 10^4	19.0
Treated effluent	not detected	

TABLE 4. Population densities of bacteria exhibiting phenotypic polyphosphate accumulation in the IRWD Michelson Plant wastewater.

Wastewater sample	CFU/ml	Proportion of total (%)
Raw influent	4.1 +/- 2.3 x 10^4	13.2
Primary effluent	2.4 +/-1.4 x 10^4	9.2
Secondary effluent	4.2 +/- 1.1 x 10^4	43.8
Treated effluent	not detected	

TABLE 5. Population densities of bacteria containing DNA sequences homologous to the E. coli polyphosphate kinase (ppk) in the IRWD Michelson Plant wastewater.

Wastewater sample	CFU/ml	Proportion of total (%)
Raw influent	9.3 +/- 3.3 x 10^3	3.0
Primary effluent	2.6 +/- 1.4 x 10^3	1.0
Secondary effluent	9.6 +/- 2.7 x 10^2	1.0
Treated effluent	not detected	

Figure 4 shows the data for *in situ* induction of the denitrification potential for natural wastewater bacteria. The data showed a significant enhancement of denitrification in the wastewater achieved through the addition of *E. coli* membranes (Oxyrase®) . Simple prevention of oxygen diffusion into the batch reactors by sealing was not enough to account for any signigficant removal of nitrate beyound the levels achieved in constantly aerated reactors.

Figure 5 shows the data from experiments conducted to detect the increase in the production of specific protein molecules, presumable enzyme important in phosphate accumulation, in wastewater bacteria, compared to an induced strain of *Acinetobacter calcoaceticus*, a noted accumulator of phosphate in many sewage treatment operations. Again the data suggest a highly diverse

genetic system for phosphate accumulation in the particular wastewater investigated.

GENETIC CLONING

Similar molecular cloning process was undertaken for introducing the polyphosphate kinase (ppk) and the nitrite reductase (nir) genes into a high expression level plasmid vector. Therefore the process will be described in genral terms. The ppk gene was obtained from professor Kornberg's laboratory at Stanford University, and it was supplied as plasmid pBC29 (Figure 6). The hyper-expression plasmid vector used was pET-15b (Figure 7). The scheme for integrating ppk with pET-15b is illustrated in the flow diagram in Figure 8. Generally, this procedure leads to a 30 - 40% increase in specific enzyme production. The exact level, with IPTG induction can be determined by performing a polyacrylamide gel electrophoresis, as shown in Figure 9. At present we have generated two independent recombinant *E. coli* strains containing the *ppk* gene and the *nir* gene respectively. We were able to conduct only preliminary experiments on the maintenance of these strains in fermentors by the end of the project. One of the outstanding goals is to incorporate the two genes into one strain of *E. coli*. Presumably, this will improve the efficiency of biotechnological enhancement of wastewater treatment with recombinant organisms.

Figure 4.

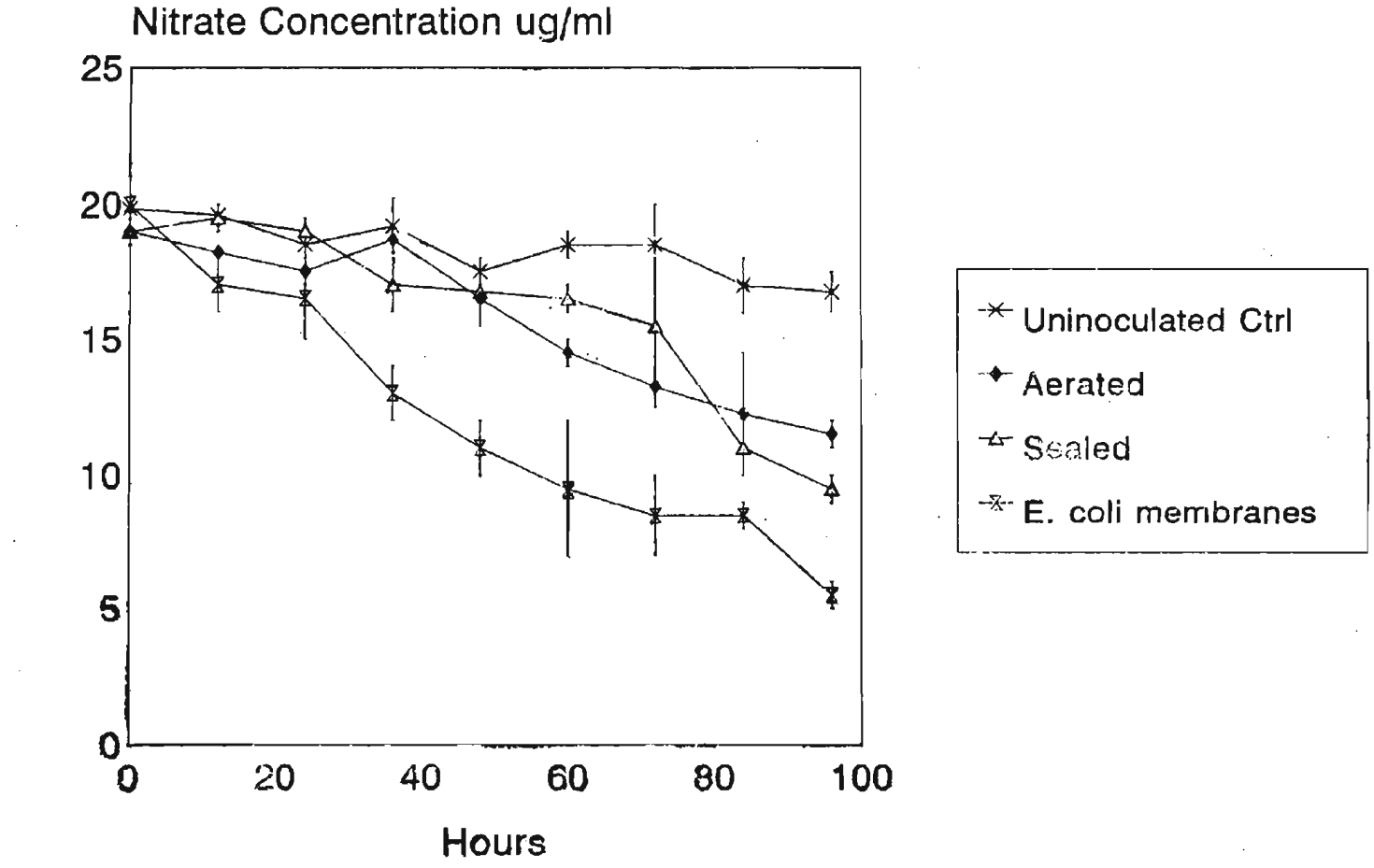

Induction of Nitrate Removal from Batch Wastewater Microcosms
Nitrate Concentration ug/ml
25
20
15
10
5
0
0
20
40
60
80
100
Hours
Uninoculated Ctrl
Aerated
Sealed
E. coli membranes

Figure 5. Polypeptides extracted from induced A. calcoaceticus and wastewater bacteria for phosphorus removal. Lanes 1 and 2; influent;3 and 4, primary effluent; 1 and 3: induced; 2 and 4:uninduced. compared to lane M.

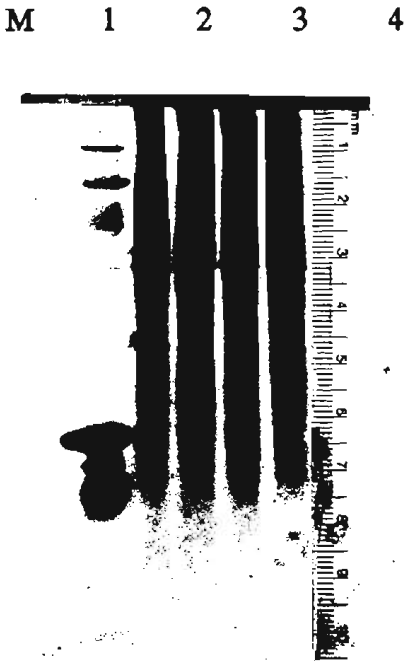


Figure 6. Plasmid pBC29 containing ppk from Kornberg laboratory at Stanford.

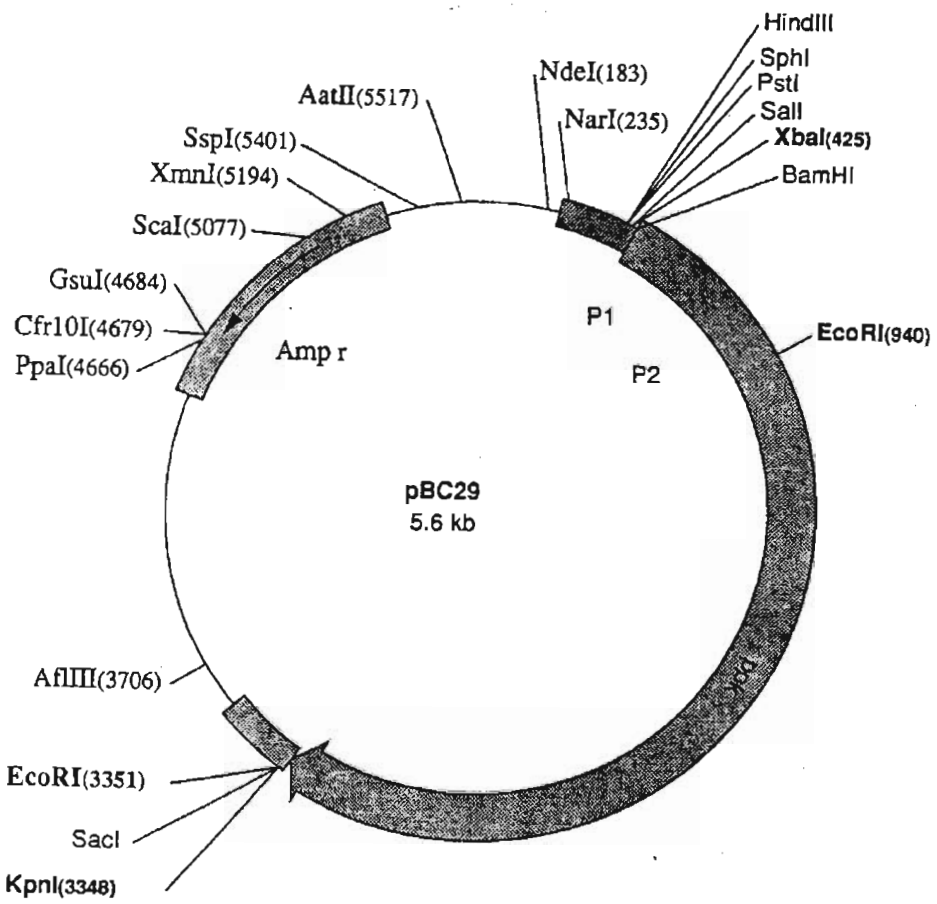


Figure 7. Cloning vector pET-15b

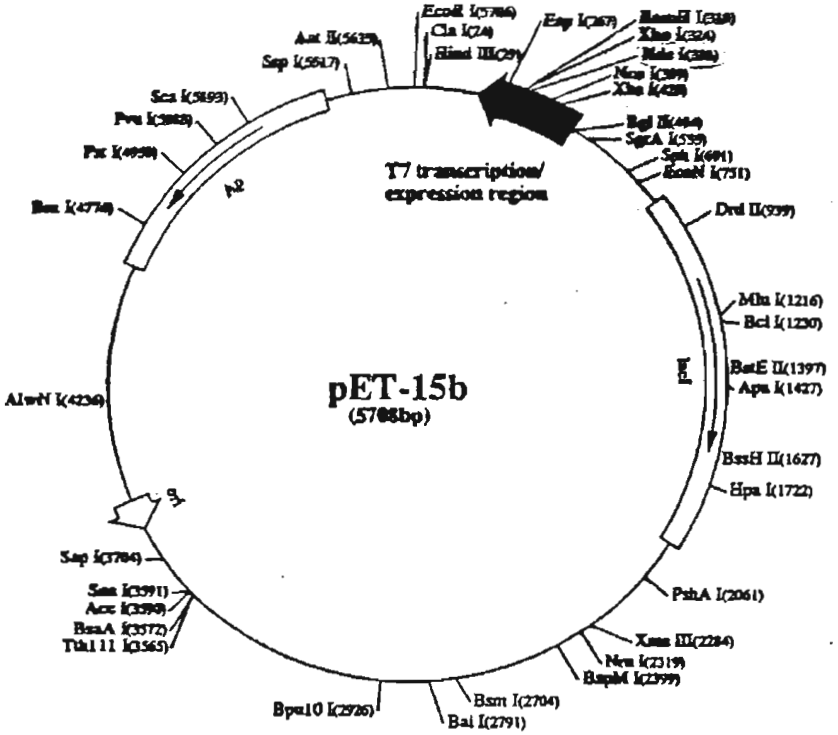


Figure 8. Construction of recombinant plasmid containing ppk.

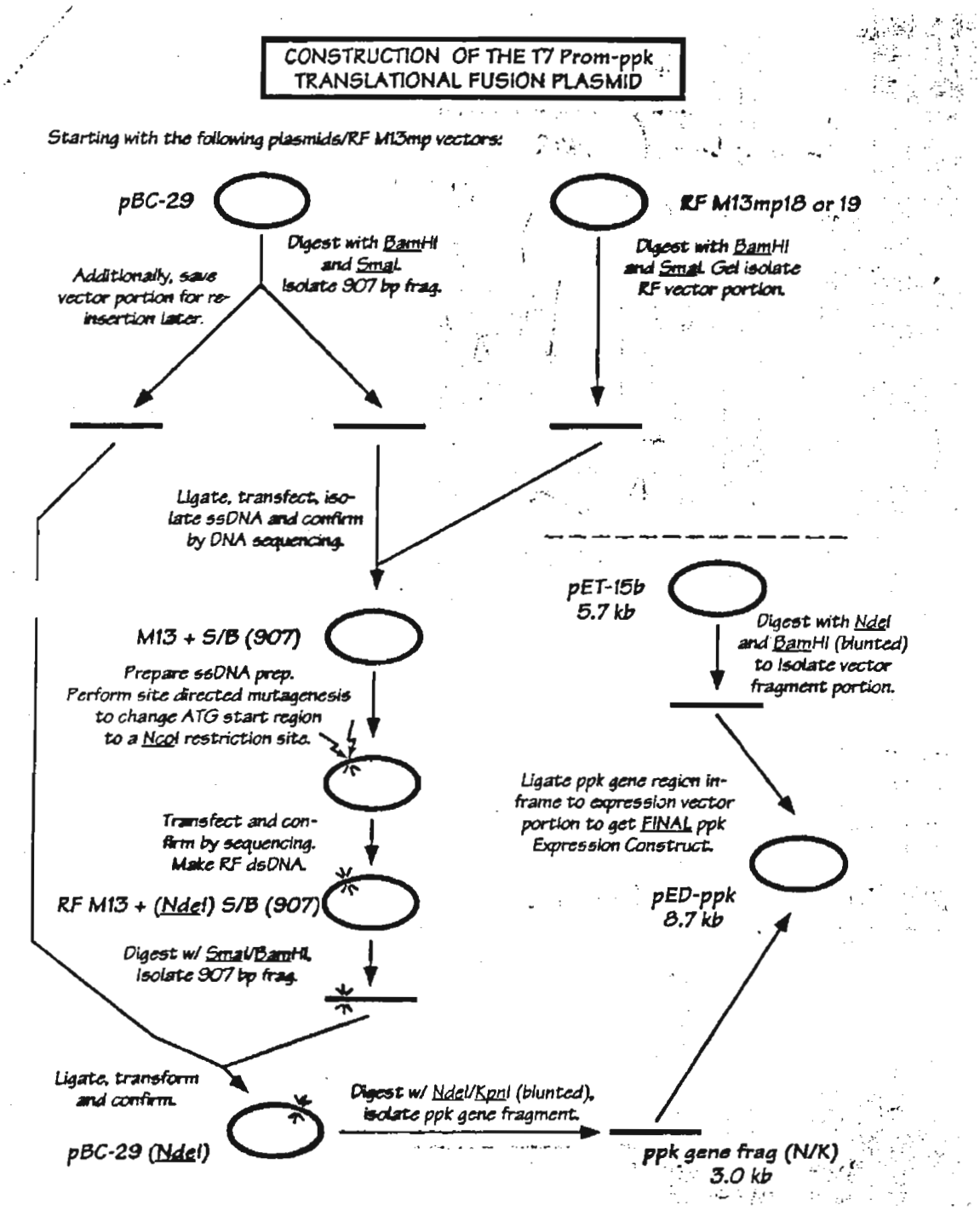
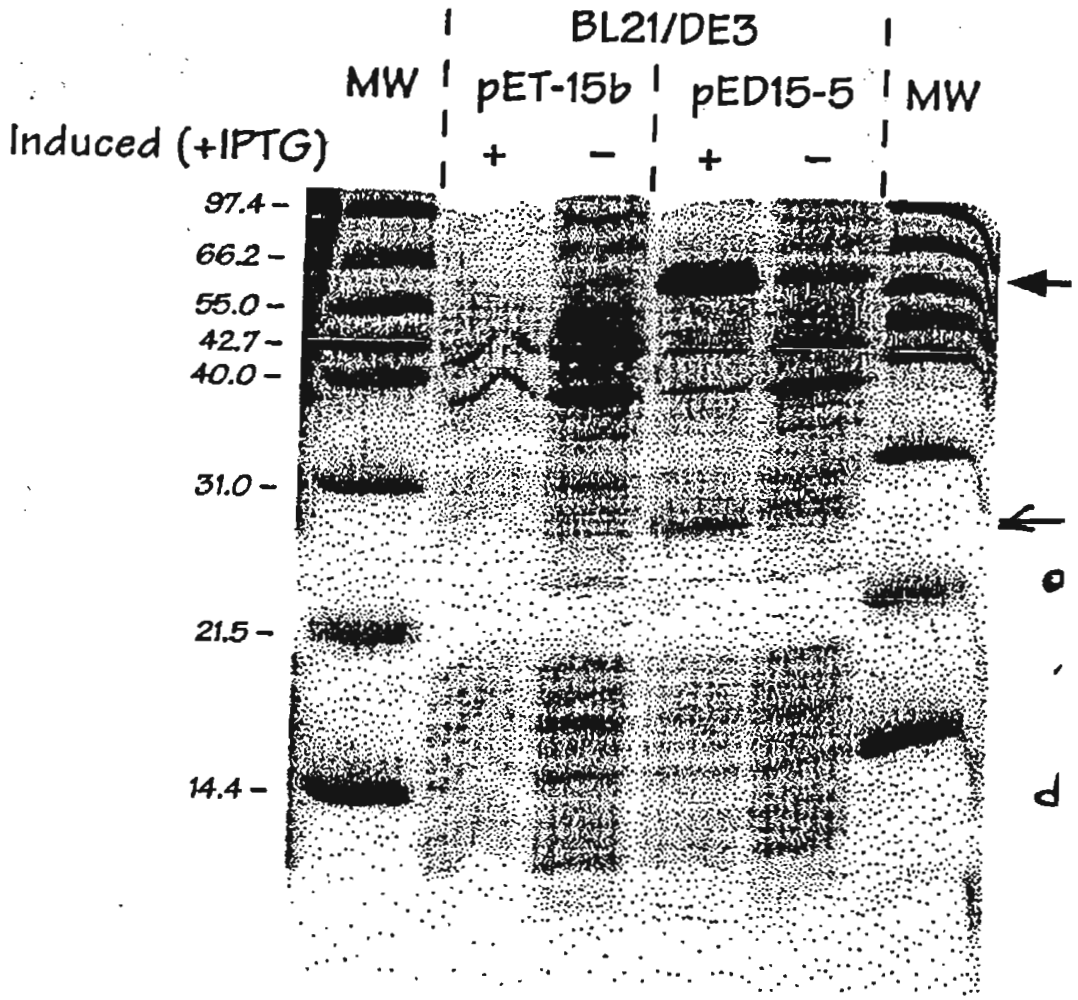


Figure 9. Picture of SDS-PAGE gel showing increased level of enzyme production by strain containing recombinant plasmid pED15-5 induced with IPTG.



BIBLIOGRAPHY

Ahn, K. and A. Kornberg. 1990. Polyphosphate kinase from Escherichia coli. J. biol. Chem. 265:11734-11739

Groenestijn, J., M. Bentvelsen, M. Deinema, and A.J.B. Zehnder. 1989. Polyphosphate-degrading enzymes in *Acinetobacter* spp. and activated sludge. Appl. Env. Microbiol. 55:219-223.

Hochstein, L.I., and G.A. Tomlinson. 1988. The enzymes associated with denitrification. Ann. Rev. Microbiol. 42:231-261.

Kato, J. K. Yamada, A. Muramatsu, H. and H. Ohtake. 1993. Genetic improvement of *Escherichia coli* for enhanced biological removal of phosphate from wastewater. Appl. Environ. Microbiol. 59:3744-3749.

Ogunseitan, O.A. 1993. Direct extraction of proteins from environmental samples. J. Microbiol. Methods. 17:273-281.

Ramalho, R.S. 1983. Introduction to wastewater treatment processes. 580 pages. Academic Press, New York.

Smith, G.B., and J.M. Tiedje. 1992. Isolation and characterization of a nitrite reductase gene and its use as a probe for identifying bacteria. Appl. Environ. Microbiol. 58:376-384.

Toerien, D.F., A. Gerber, L.H. Lotter, and T.E. Cloete. 1990. Enhanced biological phosphorus removal in activated sludge systems. Adv. Microb. Ecol. 9K.C. Marshall, ed.) 173-230.

APPENDIX

(attached)

DNA sequences desribing the polyphosphate kinase gene and the nitrite reductase gene used.

Publication from the study. In addition to the published article, we are preparing another manuscript describing the cloning work. Presentatins were also made to the 1992 and 1993 Annual Meetings of the American Society for Microbiology.

BECKMAN CENTER

Department of Biochemistry
School of Medicine

Stanford University Medical Center
Stanford, California 94305-5307
(415) 723 6161
FAX (415) 723 6783

January 16, 1992

Dr. Oladele Ogunseitan
Enviromental Microbiology and
Genetics Laboratory
Enviromental Analysis and Design Program
in Social Ecology
University of California
Irvine, Ca 92717

Dear Dr. Oladele Ogunseitan:

Thank you for your letter of December 14, 1991. We are sending you the clone pBC29 and the DNA sequence of the *ppk* gene. In pBC29, a 3-kb *Bgl*I-*Kpn*I fragment encoding the *ppk* gene is inserted into *Sma*I-*Kpn*I sites of the vector pUC18. The 3-kb fragment can be isolated by digesting pBC29 with *Kpn*I-*Bam*HI or *Kpn*I-*Hin*dIII. The plasmid DNA (10 µg, also indicated on the tube) was ethanol precipitated and dried; add an appropriate amount of TE buffer before your use.

We have not yet published the cloning of the *ppk* gene, nor its DNA sequence. We understand that you will limit the use of the sequence data and the plasmid for identifying other polyphosphate solubilizing bacteria in activated sludge. We would appreciate receiving your reprints related to polyphosphate metabolism.

Sincerely yours,

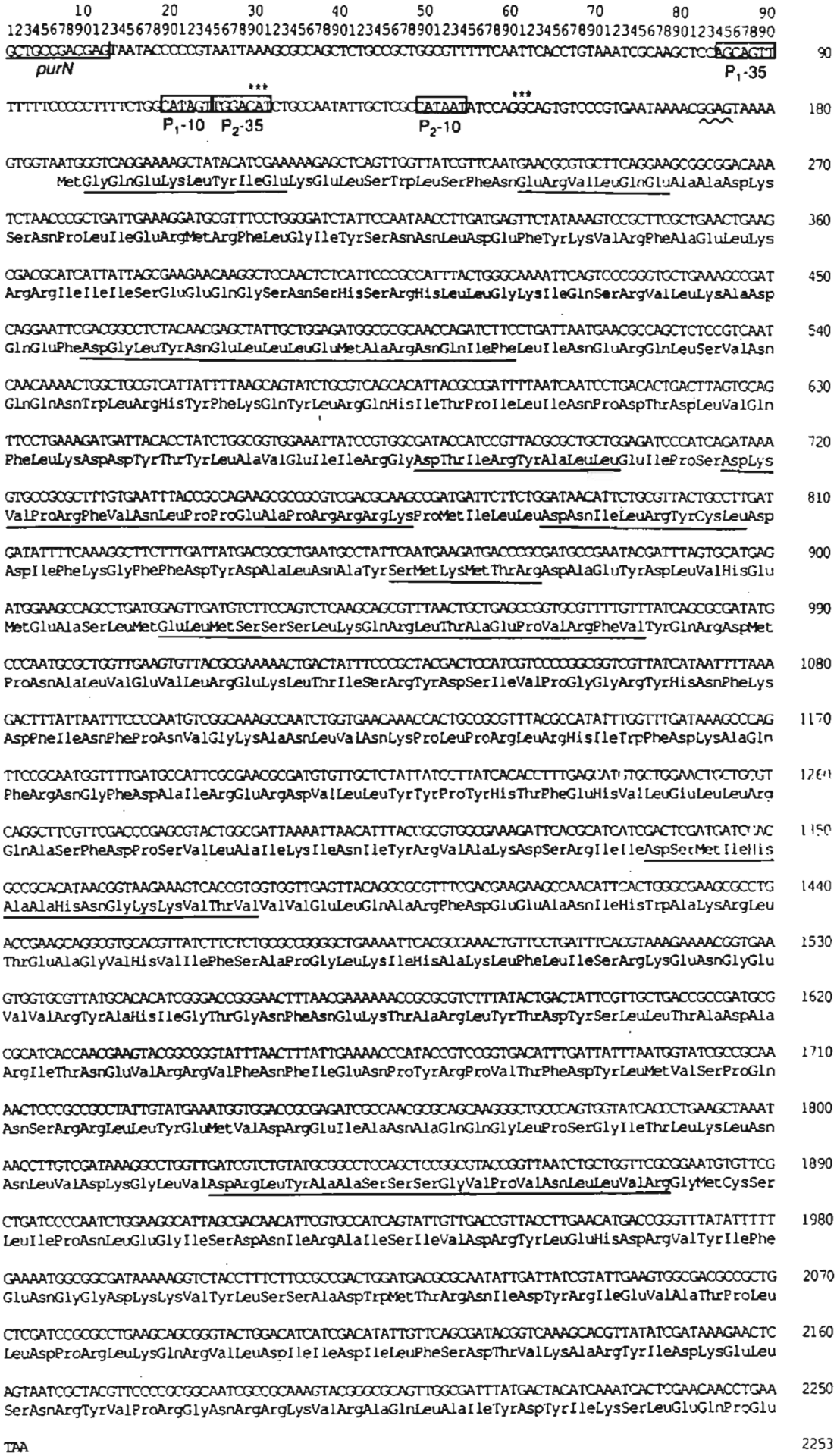
Masahiro Akiyama

Elliott Crooke

cc: Arthur Kornberg

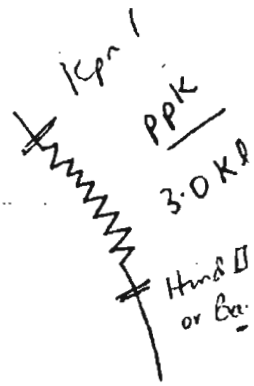
Nucleotide sequence of the 2.8-kb region of *ppk*. The nucleotide sequence of the noncoding strand of the *ppk* gene is given from 5' to 3', starting at nt-187 and terminating at 2250. The transcription start sites are marked by asterics; the putative promoters (P_1 and P_2) are boxed; a possible Shine-Delgarno sequence for translation initiation is marked by a wavy line. The predicted amino acid sequence is also shown. Analyzed internal peptide sequences were in agreement with the deduced amino acid sequence, and are underlined.

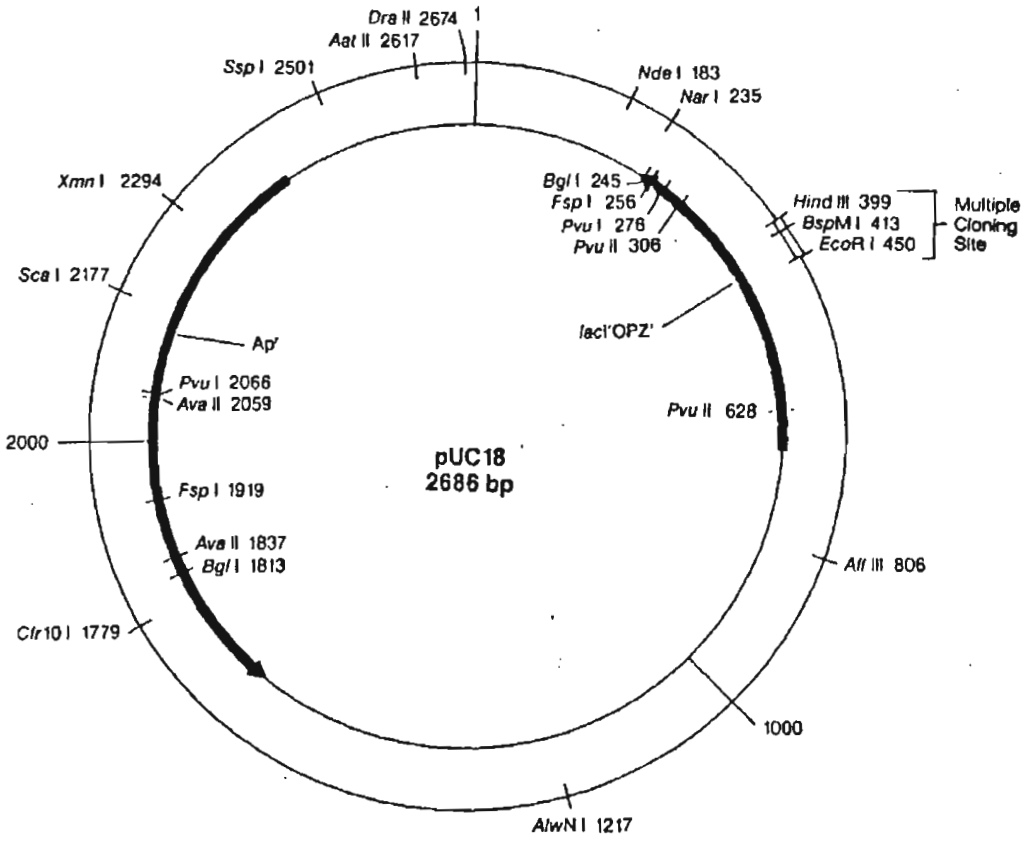
Fig 3
6


10 20 30 40 50 60 70 80 90
purN
P1-35
P1-10
P2-35
P2-10
MetGlyGlnGluLysLeuTyrIleGluLysGluLeuSerTrpLeuSerPheAsnGluArgValLeuGlnGluAlaAlaAspLys
SerAsnProLeuIleGluArgMetArgPheLeuGlyIleTyrSerAsnAsnLeuAspGluPheTyrLysValArgPheAlaGluLeuLys
ArgArgIleIleIleSerGluGluGlnGlySerAsnSerHisSerArgHisLeuLeuGlyLysIleGlnSerArgValLeuLysAlaAsp
GlnGluPheAspGlyLeuTyrAsnGluLeuLeuLeuGluMetAlaArgAsnGlnIlePheLeuIleAsnGluArgGlnLeuSerValAsn
GlnGlnAsnTrpLeuArgHisTyrPheLysGlnTyrLeuArgGlnHisIleThrProIleLeuIleAsnProAspThrAspLeuValGln
PheLeuLysAspAspTyrThrTyrLeuAlaValGluIleIleArgGlyAspThrIleArgTyrAlaLeuLeuGluIleProSerAspLys
ValProArgPheValAsnLeuProProGluAlaProArgArgArgLysProMetIleLeuLeuAspAsnIleLeuArgTyrCysLeuAsp
AspIlePheLysGlyPhePheAspTyrAspAlaLeuAsnAlaTyrSerMetLysMetThrArgAspAlaGluTyrAspLeuValHisGlu
MetGluAlaSerLeuMetGluLeuMetSerSerSerLeuLysGlnArgLeuThrAlaGluProValArgPheValTyrGlnArgAspMet
ProAsnAlaLeuValGluValLeuArgGluLysLeuThrIleSerArgTyrAspSerIleValProGlyGlyArgTyrHisAsnPheLys
AspPheIleAsnPheProAsnValGlyLysAlaAsnLeuValAsnLysProLeuProArgLeuArgHisIleTrpPheAspLysAlaGln
PheArgAsnGlyPheAspAlaIleArgGluArgAspValLeuLeuTyrTyrProTyrHisThrPheGluHisValLeuGluLeuLeuArg
GlnAlaSerPheAspProSerValLeuAlaIleLysIleAsnIleTyrArgValAlaLysAspSerArgIleIleAspSerMetIleHis
AlaAlaHisAsnGlyLysLysValThrValValValGluLeuGlnAlaArgPheAspGluGluAlaAsnIleHisTrpAlaLysArgLeu
ThrGluAlaGlyValHisValIlePheSerAlaProGlyLeuLysIleHisAlaLysLeuPheLeuIleSerArgLysGluAsnGlyGlu
ValValArgTyrAlaHisIleGlyThrGlyAsnPheAsnGluLysThrAlaArgLeuTyrThrAspTyrSerLeuLeuThrAlaAspAla
ArgIleThrAsnGluValArgArgValPheAsnPheIleGluAsnProTyrArgProValThrPheAspTyrLeuMetValSerProGln
AsnSerArgArgLeuLeuTyrGluMetValAspArgGluIleAlaAsnAlaGlnGlnGlyLeuProSerGlyIleThrLeuLysLeuAsn
AsnLeuValAspLysGlyLeuValAspArgLeuTyrAlaAlaSerSerSerGlyValProValAsnLeuLeuValArgGlyMetCysSer
LeuIleProAsnLeuGluGlyIleSerAspAsnIleArgAlaIleSerIleValAspArgTyrLeuGluHisAspArgValTyrIlePhe
GluAsnGlyGlyAspLysLysValTyrLeuSerSerAlaAspTrpMetThrArgAsnIleAspTyrArgIleGluValAlaThrProLeu
LeuAspProArgLeuLysGlnArgValLeuAspIleIleAspIleLeuPheSerAspThrValLysAlaArgTyrIleAspLysGluLeu
SerAsnArgTyrValProArgGlyAsnArgArgLysValArgAlaGlnLeuAlaIleTyrAspTyrIleLysSerLeuGluGlnProGlu

January 21, 1

pBC 29





E. coli HB101 = lac Y

(2)

misc_feature 2029..2382
/note="Open reading frame with homology to tetraheme which has been shown to be downstream of the nitrite reductase gene in Pseudomonas stutzeri [2]"
/citation=[2]
misc_feature 2194..2208
/note="Heme c binding site"
misc_feature 2284..2298
/note="Heme c binding site"
misc_feature 2377..2382
/note="Cla I site"

BASE COUNT 549 a 782 c 674 g 377 t
ORIGIN

1 AAGCTTGATT ACGGTCAAGT CCCGCTTGAA ATGGCACCTT CTCGAGGCCG TGCACCCGCC
61 AGCAGCGCAG GCCGTACTCG GGCTCGCCGA GCTTTCGTTT CTGCATACAA CCCTCGCCCG
121 CACTGGCCTC CCCGCGACAC GACGTCACGT TCGGGGGGTC GCCGAAAGAC TCTTGACTGC
181 CATCAAGCGC GTTCGCAAGA GGCCCACCTA GGATGCAAAC CGCGCACAAG AAGAAAGCAC
241 CCGGGAACTG CCACATCGCC GTAACTAGCA GGAGCCGCCC CAAGCGCTCC AAAACGAGAG
301 ACATCCATGA GTACCATTGG TAAACCTGTG ATCGGCCTGT TCGCCGGCAT GTCGAATCTG
361 CTCGGCATGG CGGTCGCCCA TGCCGCCGCA CCGGACATGA CGGCGGAAGA AAAAGAGGCC
421 GCCAAGAAGA TCTACTTCGA GCGCTGGGCC GGCTGTCACG GTGTTCTGCG CAAGGGCGCC
481 ACGGGCAAGA ACCTCGAACC GCACTGGGAA AAGACCGAAG ACGGCAAGAA AATCGAAGGC
541 GGCACCCTGA AGCTGGGCAC CAAGCGCCTG GAGAACATCA TTGCCTTCGG TACCGAAGGC
601 GGCATGGTCA ACTACGACGA CATCCTGACC GCCGAAGAAA TCAACCTGAT GGCGCGCTAT
661 ATCCAGCACA CGCCCGACAT TCCCCCAGAG TTCTCTCTGC ACGACATGAA GGACAGCTGG
721 AACCTGATCG TTCCGGTGGA ACGCCGAAGA CAGATGAACA AGGTCAACCT CGAGAACGTG
781 TTCGCCATCA CCCTGCGTGA CGCGCAGCTC TGGGACGGTG ATACCCACGA GATCTGGAAG
841 ATCCTCGATA CCGGCTACGC GGTGCACATC TCGCGTCTGT CCGCCTCCGG CCGTATGTCT
901 ACACCGTCGG CCGGATGGCT GACCACCATC ATCGACATGT CGTATCCCGA ACCCACCACC
961 GTCGCGACCG TTCGCCTGGG TCCGATCCGC TCGGTGGACG TCTCTAAGTT CAAGGGCTAC
1021 GAAGACAAGT ACCTGATCGG TGGCACCTAC TGGCCGCCAC AGTACTCGAT CATGGACGGC
1081 GAGACTCTGG AACCGATGAA AGTCGTCTCC ACCCGCGGCC AGACCGTCGA TGGCGACTAC
1141 CACCCCGAGC CGCGCGTGGC GTCCATCGTC GCCTCGCACA TCAAGCCCGA GTGGGTGGTG
1201 AACGTGAAGG AGACCGGGCA GATCATGCTG GTCGACTACA CCGACATCAA GAACCTCAAG
1261 ACCACCACCA TCGAATCCGC CAAGTTCCTG CACGACGGCG GCTGGGATGC CTCCCATCGC
1321 TACTTCATGG TCGCCGCCAA CGCCTCCAAC AAGGCTGCGC CTGCAGTCGA TACCAAGACC
1381 GGTAAGCTGG CAGCTCTGAT CGATACCGCG AAGATCCGCA CCCGGACGCG CAACTTCGTG
1441 CACCCGCAGT TCGCCCCCCT ATGCTCCACC GGCCACCTGG GCGACGACGT GGTGTCCCTC
1501 ATCTCCACGC CTTCGGATGA ATCCAAGTAC GCCAAGTACA AGGAGCACAA CTGGAAGGTG

(1)

```
LOCUS       NIrauoth1.   2382 bp DNA                              21-OCT-1991
DEFINITION  Pseudomonas stutzeri JM 300 Nitrite Reductase.
KEYWORDS    Nitrite Reductase; nitrite reductase.
SOURCE      Pseudomonas stutzeri, 2.4 KB Cla I fragment.
  ORGANISM  Pseudomonas stutzeri
REFERENCE   1   (bases 1 to 2382)
  AUTHORS   Smith,C.B. and Tiedje,J.M.
  TITLE     Isolation and Characterization of a Nitrite Reductase Gene
            and its Use as a Probe for Denitrifying Bacteria
  JOURNAL   Applied and Environmental Microbiology, 58 (1), 376-384 (1992)
  STANDARD  full automatic
REFERENCE   2   (bases 1 to 2382)
  AUTHORS   Jungst,A., Wakabayashi,S., Matsubara,H. and Zumft,W.G.
  TITLE     The nirSTBM region coding for cytochrome cd1-depedent
            nitrite respiration of Pseudomonas stutzeri consists of a
            cluster of mono-, di-, and tetraheme proteins
  JOURNAL   FEBS Letters 279 (2), 205-209 (1991)
  COMMENT
            Data kindly submitted in computer readable form by:
            Michael MGL LaMontagne
            Biology, Marine Program
            Boston University
            DUMP/MBL
            Woods Hole Ma. 02543
            USA
            Phone:  508-548-3705 X 516
            Fax:    508-548-7295
            [1] Author requested hold until 11-SEP-1991
            Corresponding Author for Reference #1:
            Geoffrey GBS Smith Phd
            Biology
            New Mexico State University
            Las Cruces New Mexico 88003
            Phone:  505-646-3611
            Corresponding Author for Reference #2:
FEATURES             Location/Qualifiers
     promoter        1..306
     RBS             294..299
                     /note="Shine-Dalgarno sequence"
     RBS             307..309
                     /note="initiation codon"
     promoter        join(5..9, 14..18)
                     /note="fnr box"
                     /citation=[1]
     promoter        join(173..177, 182..186)
                     /citation=[1]
                     /note="fnr box"
     sig_peptide     307..384
     CDS             385..1962
                     /product="Nitrite Reductase"
     mat_peptide     385..1959
                     /product="nitrite reductase"
     misc_feature    445..459
                     /function="Heme c binding site"
     RBS             2015..2019
                     /note="Shine-Dalgarno sequence"
     RBS             2029
                     /note="initiation codon"
```

```
LOCUS       PSENITRED    1872 bp ds-DNA              BCT       21-MAY-1991
DEFINITION  Pseudomonas aeruginosa gene for nitrite reductase (EC 1.6.6.4)
ACCESSION   X16452
KEYWORDS    nitrite reductase (NAD(P)H).
SOURCE      Pseudomonas aeruginosa DNA.
  ORGANISM  Pseudomonas aeruginosa
            Prokaryota; Bacteria; Gracilicutes; Scotobacteria;
            Pseudomonadaceae.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Silvestrini,M.C., Galeotti,C.L., Gervais,M., Schinina,E., Barra,D.,
            Bossa,F. and Brunori,M.
  TITLE     Nitrite reductase from Pseudomonas aeruginosa: sequence of the gene
            and the protein
  JOURNAL   FEBS Lett. 254, 33-38 (1989)
  STANDARD  full automatic
COMMENT     From EMBL 26   entry PANITRED;  dated 27-MAR-1991.
FEATURES             Location/Qualifiers
     mRNA            <1..1872
                     /gene="nitrit reductase gene"
                     /evidence=EXPERIMENTAL
     CDS             82..1788
                     /EC_number="1.6.6.4"
                     /product="nitrite reductase (NAD(P)H)"
                     /gene="nitrit reductase gene"
                     /codon_start=1
BASE COUNT      402 A    651 C    555 G    264 T
ORIGIN
        1 GAATTCCCGG GAGTTCCCGA CGCAGCCACC CCCAAAACAC TGCTAAGGGA GCGCCTCGCA
       61 GGGCTCCTGA GGAGATAGAC CATGCCATTT GGCAAGCCAC TGGTGGGCAC CTTGCTCGCC
      121 TCGCTGACGC TGCTGGGCCT GGCCACCGCT CACGCCAAGG ACGACATGAA AGCCGCCGAG
      181 CAATACCAGG GTGCCGCTTC CGCCGTCGAT CCCGCTCACG TGGTGCGCAC CAACGGCGCT
      241 CCCGACATGA GTGAAAGCGA GTTCAACGAG GCCAAGCAGA TCTACTTCCA ACGCTGCGCC
      301 GGTTGCCACG GCGTCCTGCG CAAGGGCGCC ACCGGCAAGC CGCTGACCCC GGACATCACC
      361 CAGCAACGCG GCCAGCAATA CCTGGAAGCG CTGATCACCT ACGGCACCCC GCTGGGCATG
      421 CCGAACTGGG GCAGCTCCGG CGAGCTGAGC AAGGAACAGA TCACCCTGAT GGCCAAGTAC
      481 ATCCAGCACA CCCCGCCGCA ACCGCCGGAG TGGGGCATGC CGGAGATGCG CGAATCGTGG
      541 AAGGTGCTGG TGAAGCCGGA GGACCGGCCG AAGAAACAGC TCAACGACCT CGACCTGCCC
      601 AACCTGTTCT CGGTGACCCT GCGCGACGCC GGGCAGATCG CCCTGGTCGA CGGCGACAGC
      661 AAAAAGATCG TCAAGGTCAT CGATACCGGC TATGCCGTGC ATATCTCGCG GATGTCCGCT
      721 TCCGGCCGCT ACCTGCTGGT GATCGGCCGC GACGCGCGGA TCGACATGAT CGACCTGTGG
      781 GCCAAGGAGC CGACCAAGGT CGCCGAGATC AAGATCGGCA TCGAGGCGCG CTCGGTGGAA
      841 AGCTCCAAGT TCAAGGGCTA CGAGGACCGC TACACCATCG CCGGCGCCTA CTGGCCGCCG
      901 CAGTTCGCGA TCATGGACGG CGAGACCCTG GAACCGAAGC AGATCGTCTC CACCCGCGGC
      961 ATGACCGTAG ACACCCAGAC CTACCACCCG GAACCGCGCG TGGCGGCGAT CATCGCCTCC
     1021 CACGAGCACC CCGAGTTCAT CGTCAACGTG AAGGAGACCG GCAAGGTCCT GCTGGTCAAC
     1081 TACAAGGATA TCGACAACCT CACCGTCACC AGCATCGGTG CGGCGCCGTT CCTCCACGAC
     1141 GGCGGCTGGG ACAGCAGCCA CCGCTACTTC ATGACCGCCG CCAACAACTC CAACAAGGTT
     1201 GCCGTGATCG ACTCCAAGGA CCGTCGCCTG TCGGCCCTGG TCGACGTCGG CAAGACCCCG
     1261 CACCCGGGGC GTGGCGCCAA CTTCGTGCAT CCCAAGTACG GCCCGGTGTG GAGCACCAGC
     1321 CACCTGGGCG ACGGCAGCAT CTCGCTGATC GGCACCGATC CGAAGAACCA TCCGCAGTAC
     1381 GCCTGGAAGA AAGTCGCCGA ACTACAGGGC CAGGGCGGCG GCTCGCTGTT CATCAAGACC
     1441 CATCCGAAGT CCTCGCACCT CTACGTCGAC ACCACCTTCA ACCCCGACGC CAGGATCAGC
     1501 CAGAGCGTCG CGGTGTTCGA CCTGAAGAAC CTCGACGCCA AGTACCAGGT GCTGCCGATC
     1561 GCCGAATGGG CCGATCTCGG CGAAGGCGCC AAGCGGGTGG TGCAGCCCGA GTACAACAAG
     1621 CGCGGCGATG AAGTCTGGTT CTCGGTGTGG AACGGCAAGA ACGACAGCTC CGCGCTGGTG
     1681 GTGGTGGACG ACAAGACCCT GAAGCTCAAG GCCGTGGTCA AGGACCCGCG GCTGATCACC
     1741 CCGACCGGTA AGTTCAACGT CTACAACACC CAGCACGACG TGTACTGAGA CCCGCGTGCG
     1801 GGGCACGCCC CGCACGCTCC CCCCTACGAG GAACCGTGAT GAAACCGTAC GCACTGCTTT
     1861 CGCTGCTCGC CA
```

DNA Strider 1.0 ### Friday, July 24, 1992 7:41:59 AM

[redacted] (P. stutzeri) -> [redacted]

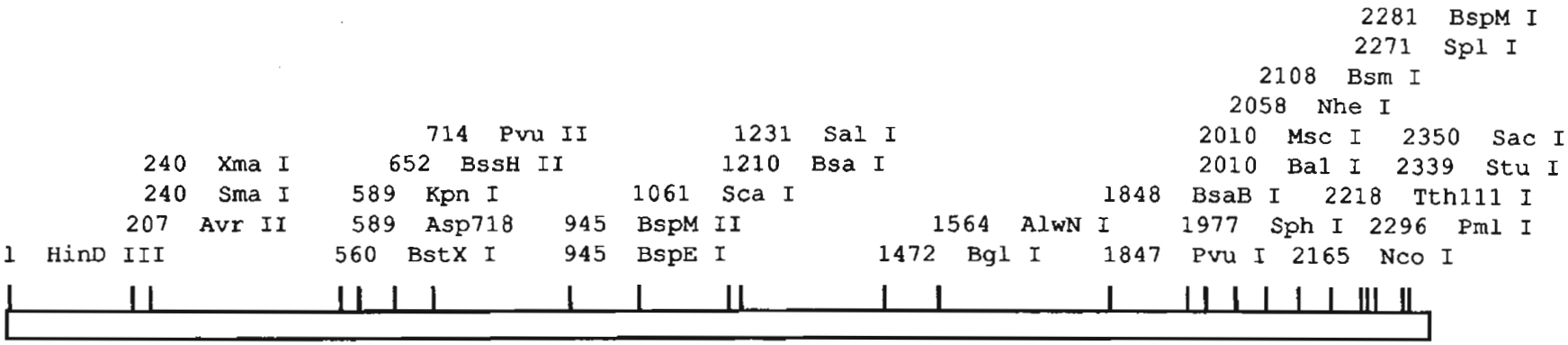
DNA sequence 2382 b.p. AAGCTTGATTAC ... GGCACCATCGAT linear

Enzyme	Site	<--	Pos.	-->
HinD III	a/agctt	0	1	2382
Avr II	c/ctagg	206	207	2176
Sma I	ccc/ggg	239	240	2143
Xma I	c/ccggg	239	240	2143
BstX I	ccannnnn/ntgg	559	560	1823
Asp718	g/gtacc	588	589	1794
Kpn I	ggtac/c	588	589	1794
BssH II	g/cgcgc	651	652	1731
Pvu II	cag/ctg	713	714	1669
BspE I	t/ccgga	944	945	1438
BspM II	t/ccgga	944	945	1438
Sca I	agt/act	1060	1061	1322
Bsa I	ggtctc 1/5	1209	1210	1173
Sal I	g/tcgac	1230	1231	1152
Bgl I	gccnnnn/nggc	1471	1472	911
AlwN I	cagnnn/ctg	1563	1564	819
Pvu I	cgat/cg	1846	1847	536
BsaB I	gatnn/nnatc	1847	1848	535
Sph I	gcatg/c	1976	1977	406
Bal I	tgg/cca	2009	2010	373
Msc I	tgg/cca	2009	2010	373
Nhe I	g/ctagc	2057	2058	325
Bsm I	gaatgc 1/-1	2107	2108	275
Nco I	c/catgg	2164	2165	218
Tth111 I	gacn/nngtc	2217	2218	165
Spl I	c/gtacg	2270	2271	112
BspM I	acctgc 4/8	2280	2281	102
Pml I	cac/gtg	2295	2296	87
Stu I	agg/cct	2338	2339	44
Sac I	gagct/c	2349	2350	33

DNA Strider 1.0 ### Friday, July 24, 1992 7:40:13 AM

Nitrite redux. (P. stutzeri) -> **Graphic Map**

DNA sequence 2382 b.p. AAGCTTGATTAC ... GGCACCATCGAT linear



Nitrite redux. (P. stutzeri) 2382 base pairs **Unique Sites**

DNA Strider 1.0 ### Friday, July 24, 1992 7:41:03 AM

Nitrite redux. (P. stutzeri) -> All Sites

DNA sequence 2382 b.p. AAGCTTGATTAC ... GGCACCATCGAT linear

AlwN I	cagnnn/ctg	1	1	1563	1	1564	819	2			
Asp718	g/gtacc	1	1	588	2	589	1794	1			
Avr II	c/ctagg	1	1	206	2	207	2176	1			
Bal I	tgg/cca	1	1	2009	1	2010	373	2			
Bgl I	gccnnnn/nggc	1	1	1471	1	1472	911	2			
Bsa I	ggtctc 1/5	1	1	1209	1	1210	1173	2			
BsaB I	gatnn/nnatc	1	1	1847	1	1848	535	2			
Bsm I	gaatgc 1/-1	1	1	2107	1	2108	275	2			
BspE I	t/ccgga	1	1	944	2	945	1438	1			
BspM I	acctgc 4/8	1	1	2280	1	2281	102	2			
BspM II	t/ccgga	1	1	944	2	945	1438	1			
BssH II	g/cgcgc	1	1	651	2	652	1731	1			
BstX I	ccannnnn/ntgg	1	1	559	2	560	1823	1			
HinD III	a/agctt	1	1	0	2	1	2382	1			
Kpn I	ggtac/c	1	1	588	2	589	1794	1			
Msc I	tgg/cca	1	1	2009	1	2010	373	2			
Nco I	c/catgg	1	1	2164	1	2165	218	2			
Nhe I	g/ctagc	1	1	2057	1	2058	325	2			
Pml I	cac/gtg	1	1	2295	1	2296	87	2			
Pvu I	cgat/cg	1	1	1846	1	1847	536	2			
Pvu II	cag/ctg	1	1	713	2	714	1669	1			
Sac I	gagct/c	1	1	2349	1	2350	33	2			
Sal I	g/tcgac	1	1	1230	1	1231	1152	2			
Sca I	agt/act	1	1	1060	2	1061	1322	1			
Sma I	ccc/ggg	1	1	239	2	240	2143	1			
Sph I	gcatg/c	1	1	1976	1	1977	406	2			
Spl I	c/gtacg	1	1	2270	1	2271	112	2			
Stu I	agg/cct	1	1	2338	1	2339	44	2			
Tth111 I	gacn/nngtc	1	1	2217	1	2218	165	2			
Xma I	c/ccggg	1	1	239	2	240	2143	1			
Aat II	gacgt/c	2	1	140	3	141	856	2	997	1386	1
Acc I	gt/mkac	2	1	896	2	897	334	3	1231	1152	1
Ava II	g/gwcc	2	1	978	1	979	484	3	1463	920	2
Ban II	grgcy/c	2	1	79	2	80	2270	1	2350	33	3
Bbe I	ggcgc/c	2	1	474	3	475	1107	1	1582	801	2
Bgl II	a/gatct	2	1	427	2	428	402	3	830	1553	1
BsaA I	yac/gtr	2	1	1677	1	1678	618	2	2296	87	3
Cla I	at/cgat	2	1	1398	1	1399	978	2	2377	6	3
Dde I	c/tnag	2	1	1003	1	1004	695	2	1699	684	3
Drd I	gacnnnn/nngtc	2	1	1832	1	1833	64	3	1897	486	2

Eco47 III	agc/gct		2	1(	282)	2	283(	157)	3	440(	1943)	1							
EcoN I	cctnn/nnnagg		2	1(	36)	3	37(	1661)	1	1698(	685)	2							
Fsp I	tgc/gca		2	1(	466)	2	467(	1860)	1	2327(	56)	3							
Kas I	g/gcgcc		2	1(	474)	3	475(	1107)	1	1582(	801)	2							
Nar I	gg/cgcc		2	1(	474)	3	475(	1107)	1	1582(	801)	2							
Nru I	tcg/cga		2	1(	961)	2	962(	1165)	1	2127(	256)	3							
PaeR7 I	c/tcgag		2	1(	40)	3	41(	728)	2	769(	1614)	1							
Pf1M I	ccannnn/ntgg		2	1(	2204)	1	2205(	98)	2	2303(	80)	3							
Pst I	ctgca/g		2	1(	696)	2	697(	664)	3	1361(	1022)	1							
Xcm I	ccannnnn/nnnntgg		2	1(	1118)	1	1119(	347)	3	1466(	917)	2							
Xho I	c/tcgag		2	1(	40)	3	41(	728)	2	769(	1614)	1							
Xmn I	gaann/nnttc		2	1(	773)	2	774(	846)	1	1620(	763)	3							
ApaL I	g/tgcac		3	1(	49)	4	50(	812)	2	862(	576)	3	1438(	945)	1				
Bbs I	gaagac	2/6	3	1(	516)	2	517(	229)	4	746(	275)	3	1021(	1362)	1				
Bbv II	gaagac	2/6	3	1(	516)	2	517(	229)	4	746(	275)	3	1021(	1362)	1				
Hae I	wgg/ccw		3	1(	123)	3	124(	1886)	1	2010(	329)	2	2339(	44)	4				
HinC II	gty/rac		3	1(	606)	2	607(	156)	4	763(	468)	3	1231(	1152)	1				
Mae I	c/tag		3	1(	207)	3	208(	57)	4	265(	1794)	1	2059(	324)	2				
Nsp7524 I	r/catgy		3	1(	346)	4	347(	588)	2	935(	1042)	1	1977(	406)	3				
NspH I	rcatg/y		3	1(	346)	4	347(	588)	2	935(	1042)	1	1977(	406)	3				
Rma I	c/tag		3	1(	207)	3	208(	57)	4	265(	1794)	1	2059(	324)	2				
Sac II	ccgc/gg		3	1(	400)	4	401(	712)	1	1113(	651)	2	1764(	619)	3				
Alw I	ggatc	4/5	4	1(	839)	1	840(	144)	5	984(	429)	3	1413(	398)	4				
				1811(	572)	2													
Eag I	c/ggccg		4	1(	887)	2	888(	20)	5	908(	993)	1	1901(	179)	4				
				2080(	303)	3													
Nae I	gcc/ggc		4	1(	342)	3	343(	105)	5	448(	1137)	1	1585(	493)	2				
				2078(	305)	4													
NspB II	cmg/ckg		4	1(	400)	3	401(	313)	5	714(	399)	4	1113(	651)	1				
				1764(	619)	2													
Ple I	gagtc	4/5	4	1(	167)	5	168(	915)	1	1083(	652)	2	1735(	249)	4				
				1984(	399)	3													
SfaN I	gcatc	5/9	4	1(	211)	4	212(	1094)	1	1306(	269)	3	1575(	620)	2				
				2195(	188)	5													
Afl III	a/crygt		5	1(	775)	1	776(	159)	5	935(	744)	2	1679(	252)	4				
				1931(	21)	6	1952(	431)	3										
Aha II	gr/cgyc		5	1(	140)	6	141(	334)	4	475(	522)	2	997(	161)	5				
				1158(	424)	3	1582(	801)	1										
BsmA I	gtctc	1/5	5	1(	297)	3	298(	702)	2	1000(	81)	5	1081(	24)	6				
				1105(	105)	4	1210(	1173)	1										
Bsr I	actgg	1/-1	5	1(	121)	6	122(	381)	4	503(	546)	1	1049(	501)	3				
				1550(	289)	5	1839(	544)	2										
BstN I	cc/wgg		5	1(	566)	1	567(	408)	4	975(	501)	3	1476(	537)	2				
				2013(	120)	6	2133(	250)	5										
BstY I	r/gatcy		5	1(	427)	3	428(	402)	4	830(	9)	6	839(	573)	1				

				1412(399) 5		1811(572) 2									
EcoR II	/ccwgg		5	1(566) 1		567(408) 4		975(501) 3		1476(537) 2					
				2013(120) 6		2133(250) 5									
HgiA I	gwgcw/c		5	1(49) 5		50(812) 1		862(576) 3		1438(105) 4					
				1543(807) 2		2350(33) 6									
Hph I	ggtga	8/7	5	1(787) 1		788(29) 6		817(380) 4		1197(601) 2					
				1798(94) 5		1892(491) 3									
Sau96 I	g/gncc		5	1(200) 4		201(778) 1		979(474) 3		1453(10) 6					
				1463(168) 5		1631(752) 2									
Sty I	c/cwwgg		5	1(206) 3		207(1523) 1		1730(435) 2		2165(138) 4					
				2303(42) 5		2345(38) 6									
Tfi I	g/awtc		5	1(353) 3		354(919) 1		1273(246) 5		1519(150) 6					
				1669(349) 4		2018(365) 2									
Bbv I	gcagc	8/12	6	1(61) 6		62(381) 4		443(361) 5		804(550) 2					
				1354(36) 7		1390(386) 3		1776(607) 1							
Bcg I	cgannnnnntgc	12/10	6	1(364) 3		365(120) 5		485(39) 6		524(1065) 1					
				1589(519) 2		2108(252) 4		2360(23) 7							
Hae II	rgcgc/y		6	1(282) 3		283(157) 5		440(35) 7		475(88) 6					
				563(1019) 1		1582(186) 4		1768(615) 2							
Mae III	/gtnac		6	1(143) 5		144(117) 6		261(194) 4		455(342) 3					
				797(1002) 1		1799(110) 7		1909(474) 2							
Fok I	ggatg	9/13	7	1(210) 5		211(91) 8		302(319) 3		621(292) 4					
				913(392) 2		1305(210) 6		1515(732) 1		2247(136) 7					
Gdi II	yggccg	-5/-1	7	1(887) 1		888(20) 8		908(143) 6		1051(65) 7					
				1116(354) 3		1470(431) 2		1901(179) 5		2080(303) 4					
Hga I	gacgc	5/10	7	1(798) 1		799(74) 7		873(286) 3		1159(266) 4					
				1425(211) 6		1636(250) 5		1886(427) 2		2313(70) 8					
Ava I	c/ycgrg		8	1(40) 8		41(36) 9		77(163) 6		240(529) 2					
				769(375) 4		1144(42) 7		1186(594) 1		1780(187) 5					
				1967(416) 3											
Bcn I	ccs/gg		8	1(239) 4		240(1) 9		241(973) 1		1214(207) 5					
				1421(34) 8		1455(124) 6		1579(71) 7		1650(249) 3					
				1899(484) 2											
Bsp1286 I	gdgch/c		8	1(49) 7		50(30) 9		80(475) 3		555(307) 4					
				862(576) 2		1438(105) 5		1543(756) 1		2299(51) 6					
				2350(33) 8											
Eae I	y/ggccr		8	1(887) 1		888(20) 9		908(143) 5		1051(65) 8					
				1116(354) 3		1470(431) 2		1901(109) 6		2010(70) 7					
				2080(303) 4											
Nci I	cc/sgg		8	1(239) 4		240(1) 9		241(973) 1		1214(207) 5					
				1421(34) 8		1455(124) 6		1579(71) 7		1650(249) 3					
				1899(484) 2											
Ban I	g/gyrcc		9	1(32) 8		33(442) 4		475(66) 6		541(15) 9					
				556(33) 7		589(453) 3		1042(540) 2		1582(125) 5					

DNA Strider 1.0 ### Friday, July 24, 1992 7:47:28 AM

Nitrite redux. (P. stutzeri) -> [redacted]

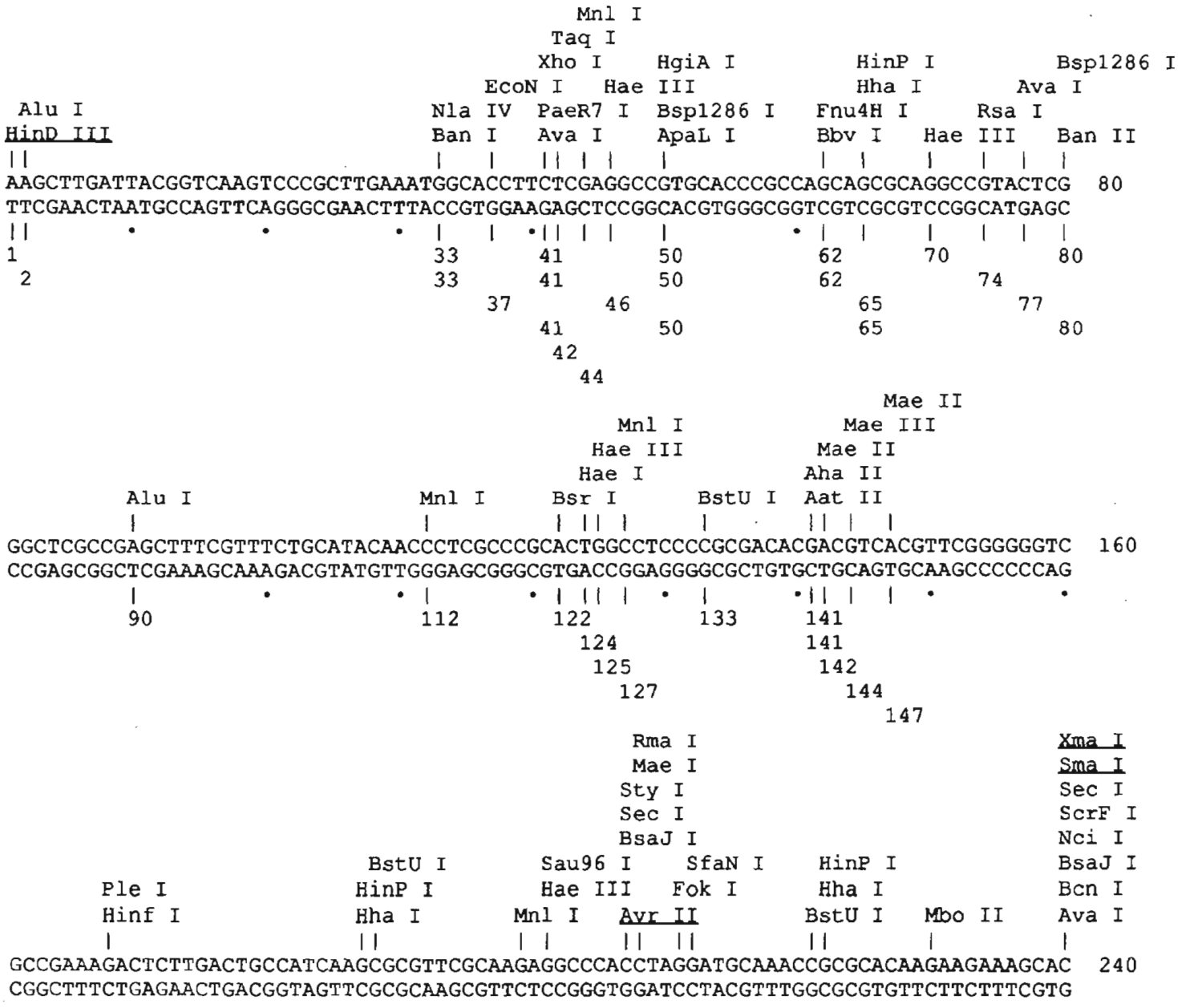
DNA sequence 2382 b.p. AAGCTTGATTAC ... GGCACCATCGAT linear

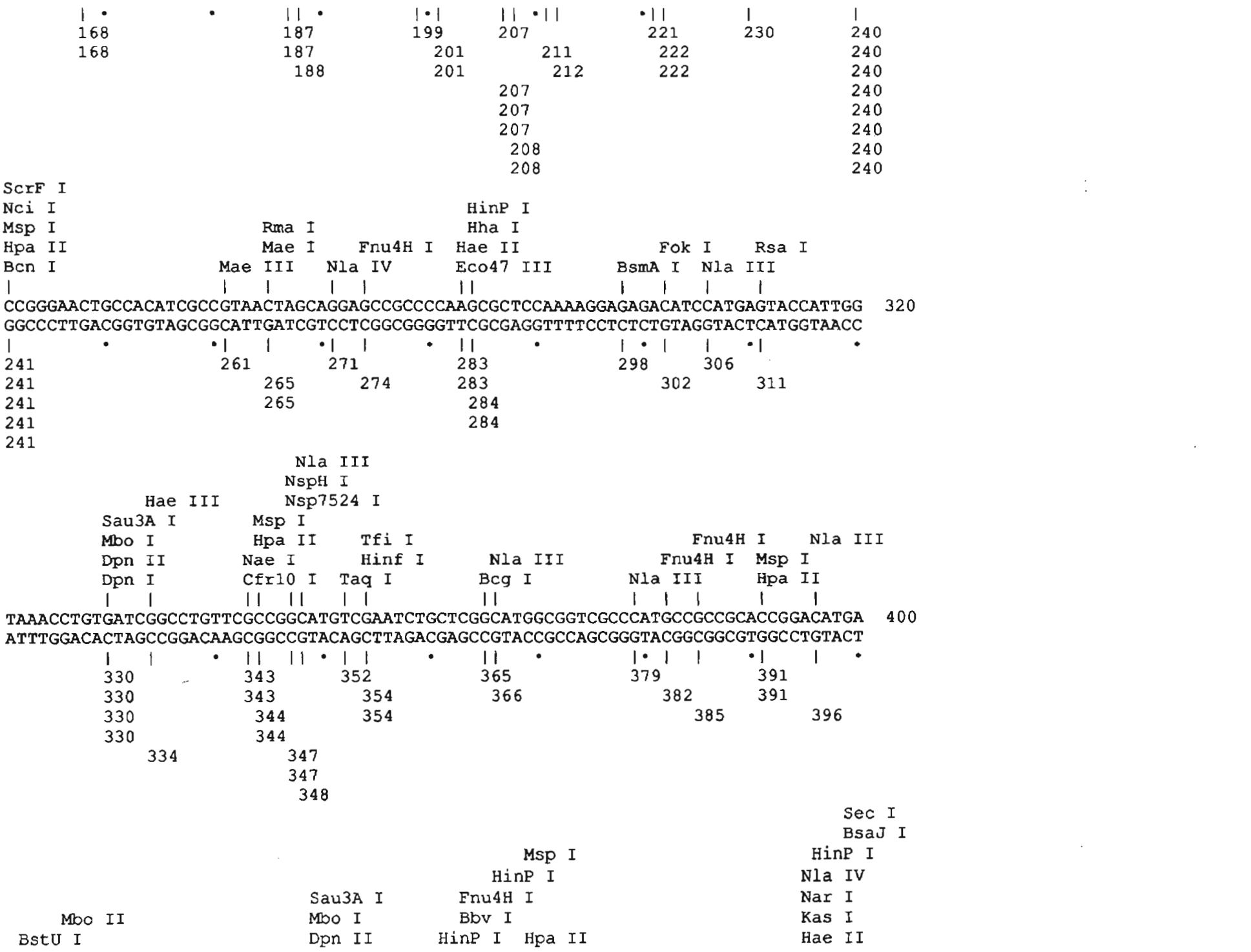
Afl II	c/ttaag	Ear I	ctcttc	1/4	Pac I	ttaat/taa	
Apa I	gggcc/c	EcoO109 I	rg/gnccy		PpuM I	rg/gwccy	
Ase I	at/taat	EcoR I	g/aattc		Rsr II	cg/gwccg	
BamH I	g/gatcc	EcoR V	gat/atc		Sfi I	ggccnnnn/nggcc	
Bcl I	t/gatca	Esp I	gc/tnagc		SnaB I	tac/gta	
BspH I	t/catga	Hpa I	gtt/aac		Spe I	a/ctagt	
BstB I	tt/cgaa	Mlu I	a/cgcgt		Ssp I	aat/att	
BstE II	g/gtnacc	Mse I	t/taa		Tth111 II	caarca	11/9
Bsu36 I	cc/tnagg	Nde I	ca/tatg		Xba I	t/ctaga	
Dra I	ttt/aaa	Not I	gc/ggccgc		Xca I	gta/tac	
Dra III	cacnnn/gtg	Nsi I	atgca/t				

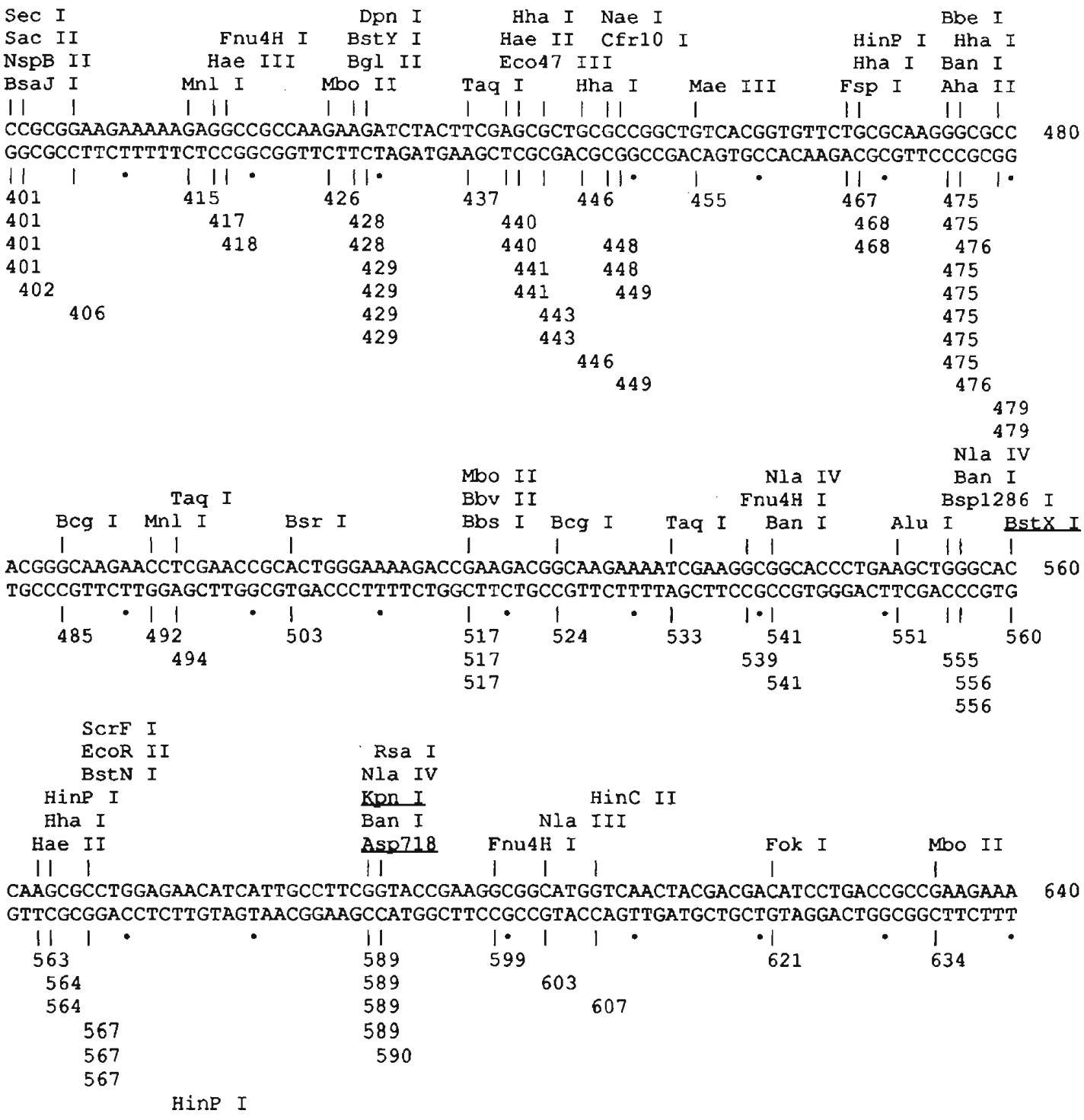
DNA Strider 1.0 ### Friday, July 24, 1992 7:48:42 AM

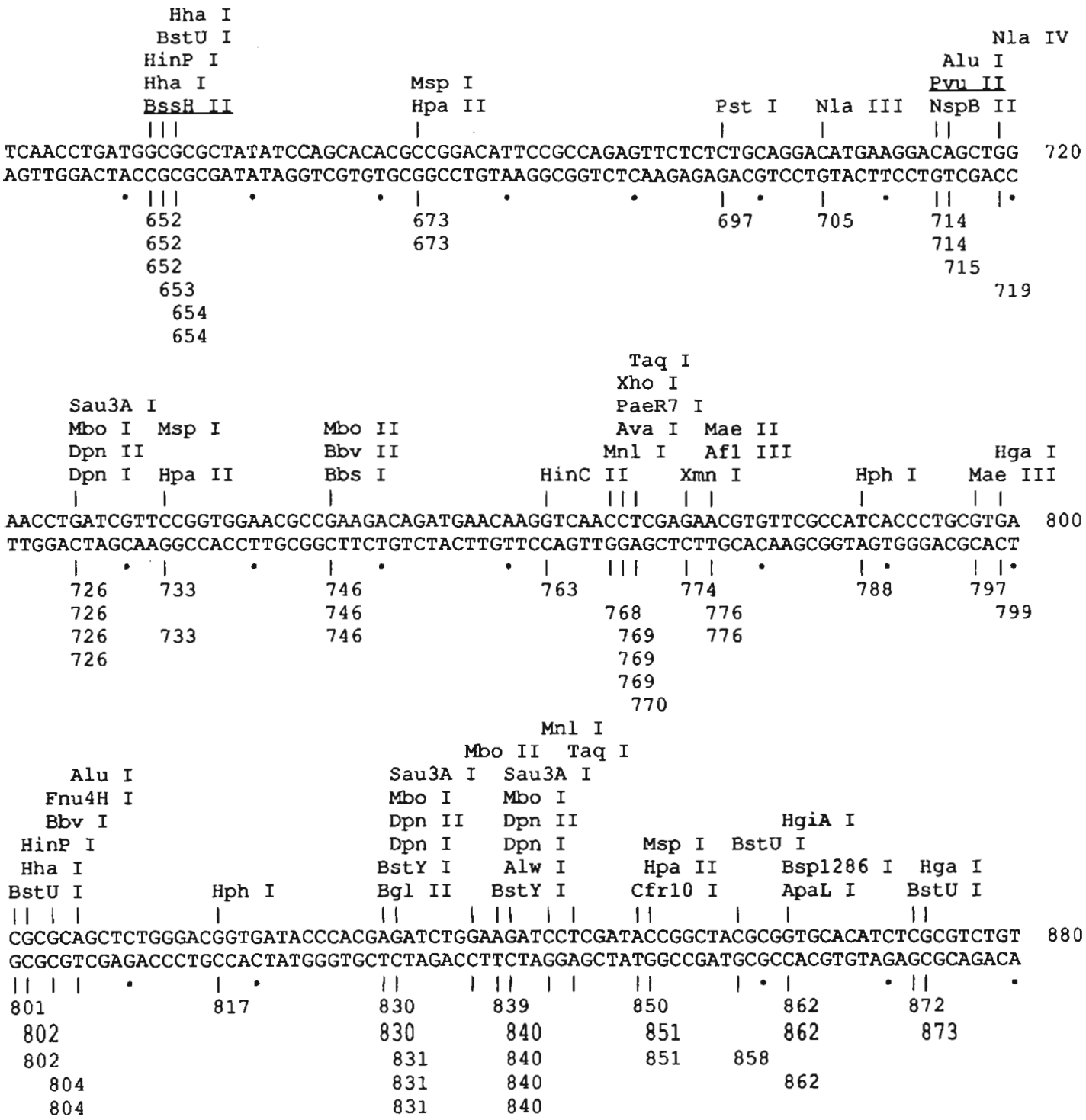
Nitrite redux. (P. stutzeri) -> **Restriction Map**

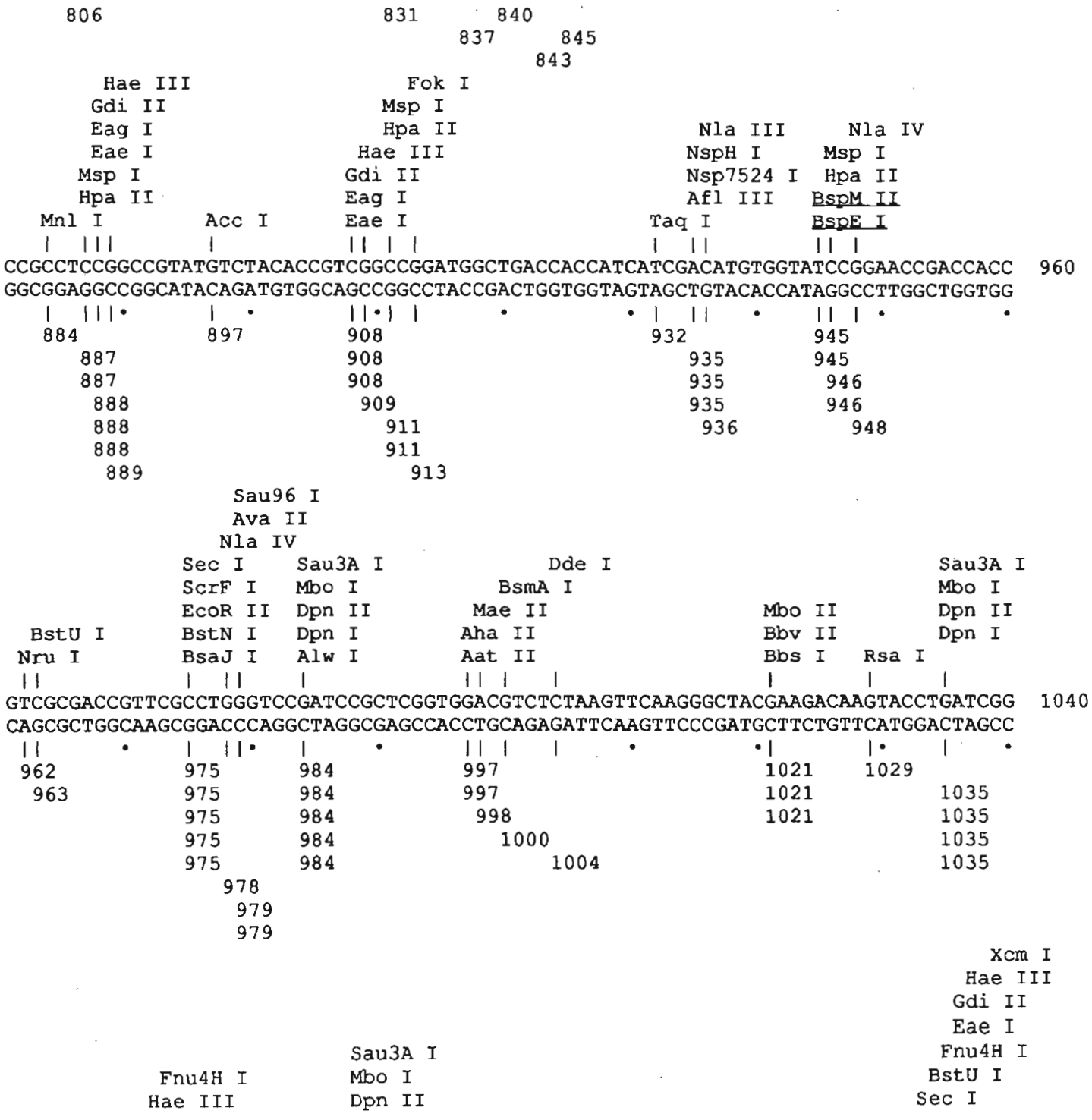
DNA sequence 2382 b.p. AAGCTTGATTAC ... GGCACCATCGAT linear

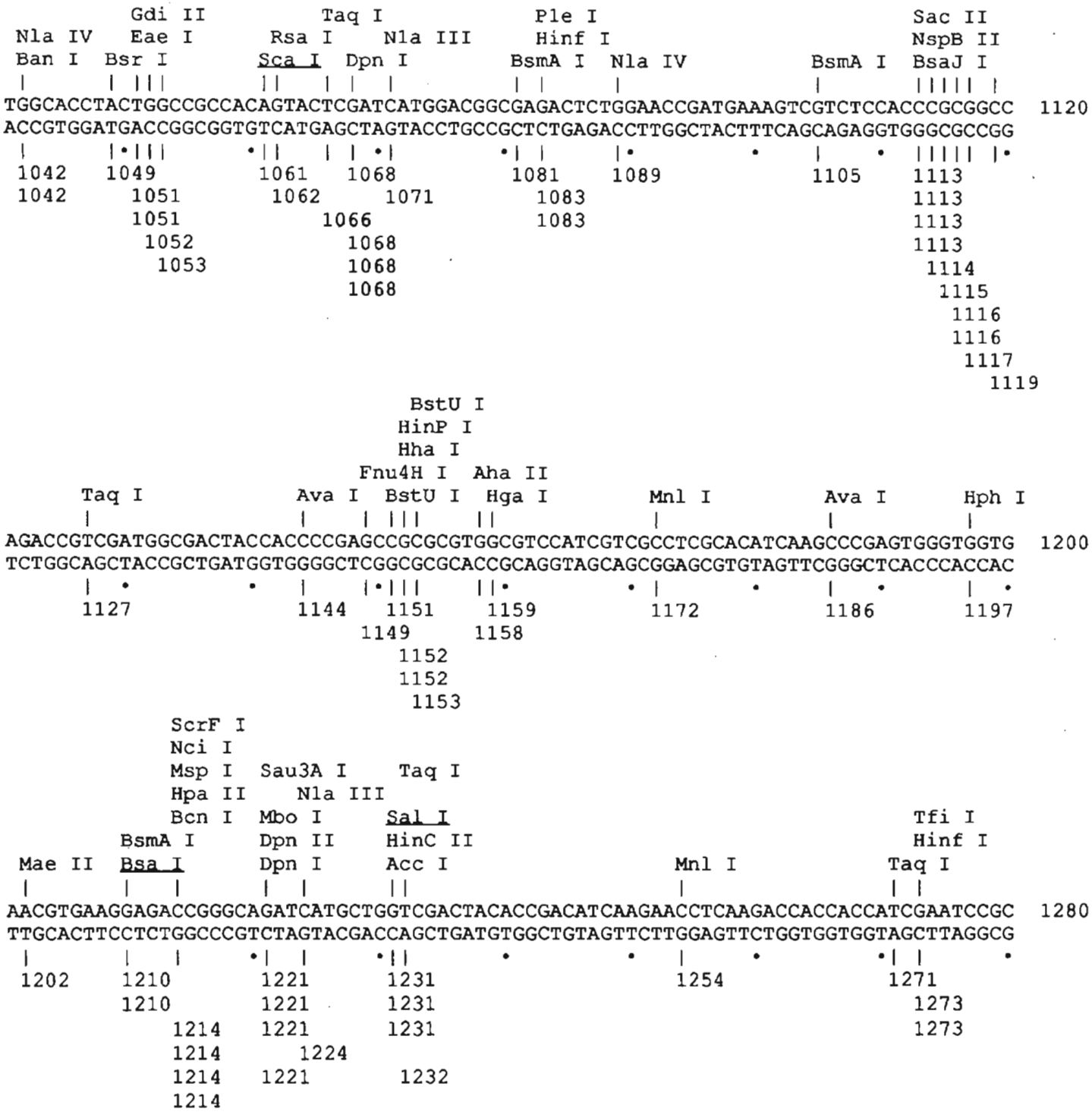


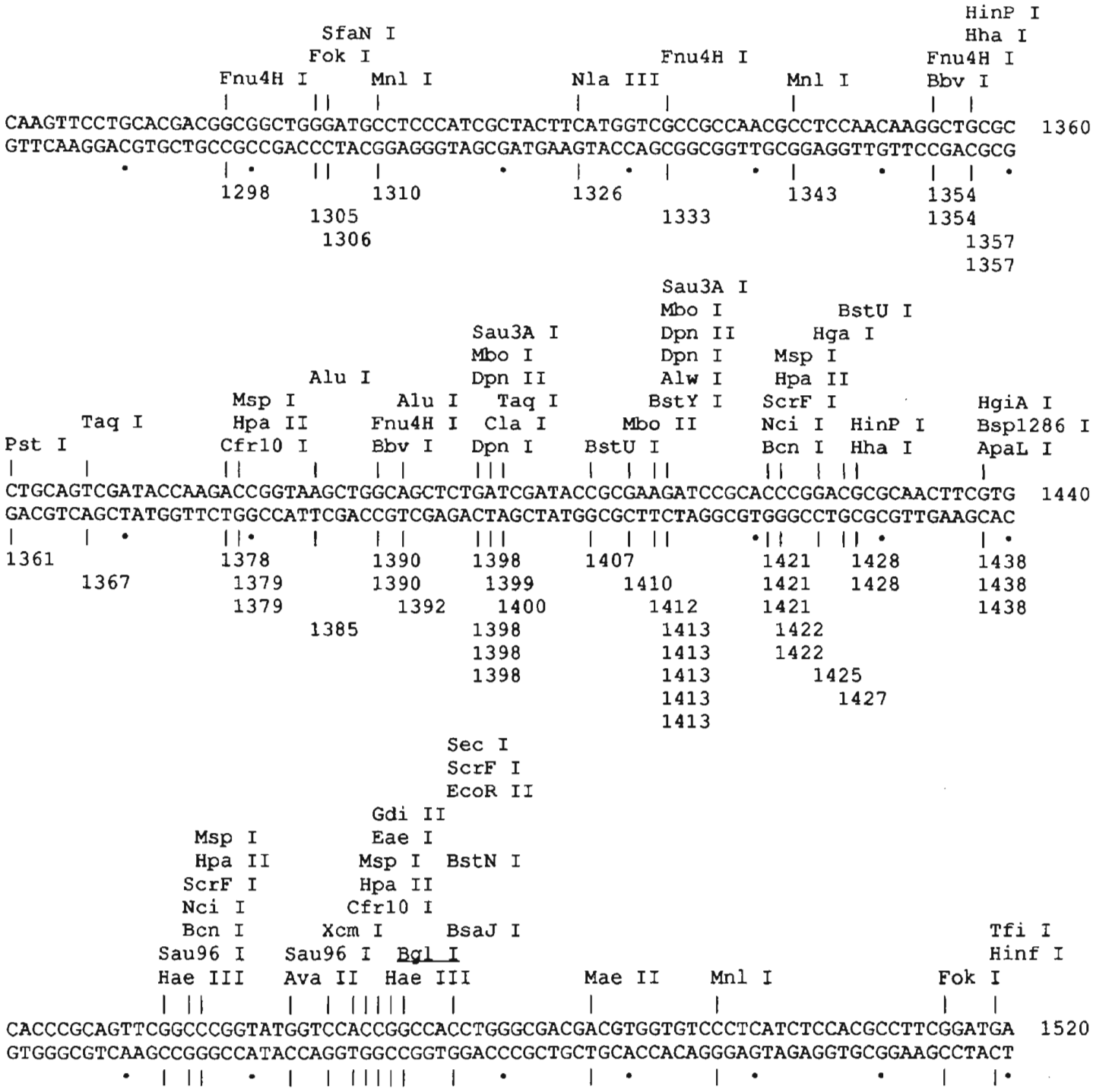

ScrF I
Nci I
Msp I
Hpa II
Bcn I
Rma I
Mae I
Mae III
Fnu4H I
Nla IV
HinP I
Hha I
Hae II
Eco47 III
Fok I
Rsa I
BsmA I
Nla III
Nla III
NspH I
Hae III
Nsp7524 I
Sau3A I
Msp I
Mbo I
Hpa II
Dpn II
Tfi I
Nae I
Hinf I
Dpn I
Cfr10 I
Taq I
Nla III
Bcg I
Fnu4H I
Nla III
Fnu4H I
Msp I
Hpa II
Sec I
BsaJ I
HinP I
Nla IV
Nar I
Kas I
Hae II
Msp I
HinP I
Sau3A I
Mbo I
Dpn II
Fnu4H I
Bbv I
HinP I
Hpa II
Mbo II
BstU I

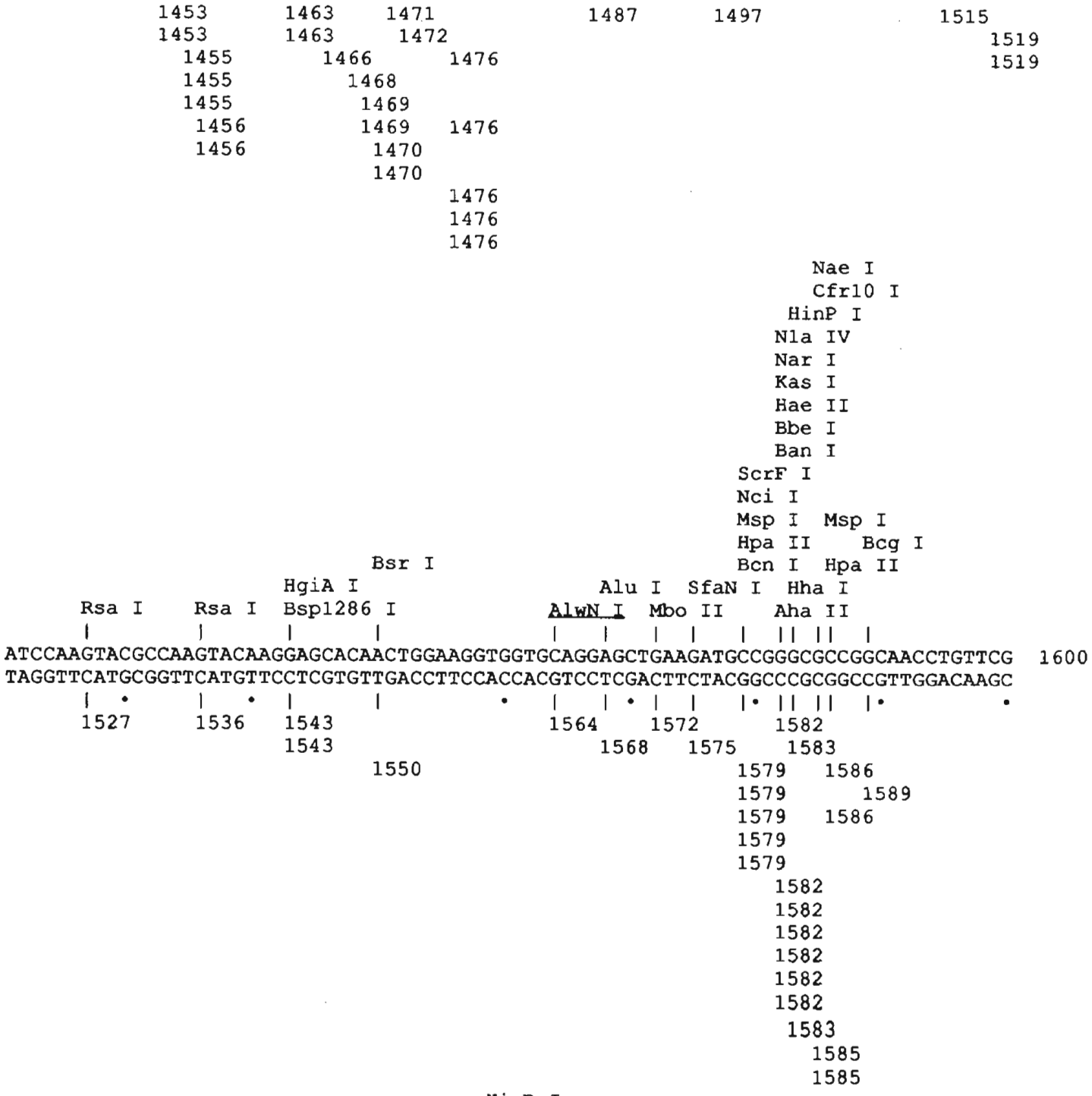

Sec I
Sac II
NspB II
BsaJ I
Fnu4H I
Hae III
Mnl I
Dpn I
BstY I
Bgl II
Mbo II
Hha I
Hae II
Eco47 III
Taq I
Hha I
Nae I
Cfr10 I
Mae III
HinP I
Hha I
Fsp I
Bbe I
Hha I
Ban I
Aha II
CCGCGGAAGAAAAAGAGGCCGCCAAGAAGATCTACTTCGAGCGCTGCGCCGGCTGTCACGGTGTTCTGCGCAAGGGCGCC 480
GGCGCCTTCTTTTTCTCCGGCGGTTCTTCTAGATGAAGCTCGCGACGCGGCCGACAGTGCCACAAGACGCGTTCCCGCGG
401 415 426 437 446 455 467 475
401 417 428 440 468 475
401 418 428 440 448 468 476
401 429 441 448 475
402 429 441 449 475
406 429 443 475
429 443 475
446 475
449 476
479
479
Nla IV
Ban I
Bsp1286 I
BstX I
Mbo II
Bbv II
Bbs I
Nla IV
Fnu4H I
Ban I
Taq I
Bcg I
Mnl I
Bsr I
Bcg I
Taq I
Alu I
ACGGGCAAGAACCTCGAACCGCACTGGGAAAAGACCGAAGACGGCAAGAAAATCGAAGGCGGCACCCTGAAGCTGGGCAC 560
TGCCCGTTCTTGGAGCTTGGCGTGACCCTTTTCTGGCTTCTGCCGTTCTTTTAGCTTCCGCCGTGGGACTTCGACCCGTG
485 492 503 517 524 533 541 551 560
494 517 539 555
517 541 556
556
ScrF I
EcoR II
BstN I
Rsa I
Nla IV
Kpn I
Ban I
Asp718
HinP I
Hha I
Hae II
HinC II
Nla III
Fnu4H I
Fok I
Mbo II
CAAGCGCCTGGAGAACATCATTGCCTTCGGTACCGAAGGCGGCATGGTCAACTACGACGACATCCTGACCGCCGAAGAAA 640
GTTCGCGGACCTCTTGTAGTAACGGAAGCCATGGCTTCCGCCGTACCAGTTGATGCTGCTGTAGGACTGGCGGCTTCTTT
563 589 599 621 634
564 589 603
564 589 607
567 589
567 590
567
HinP I

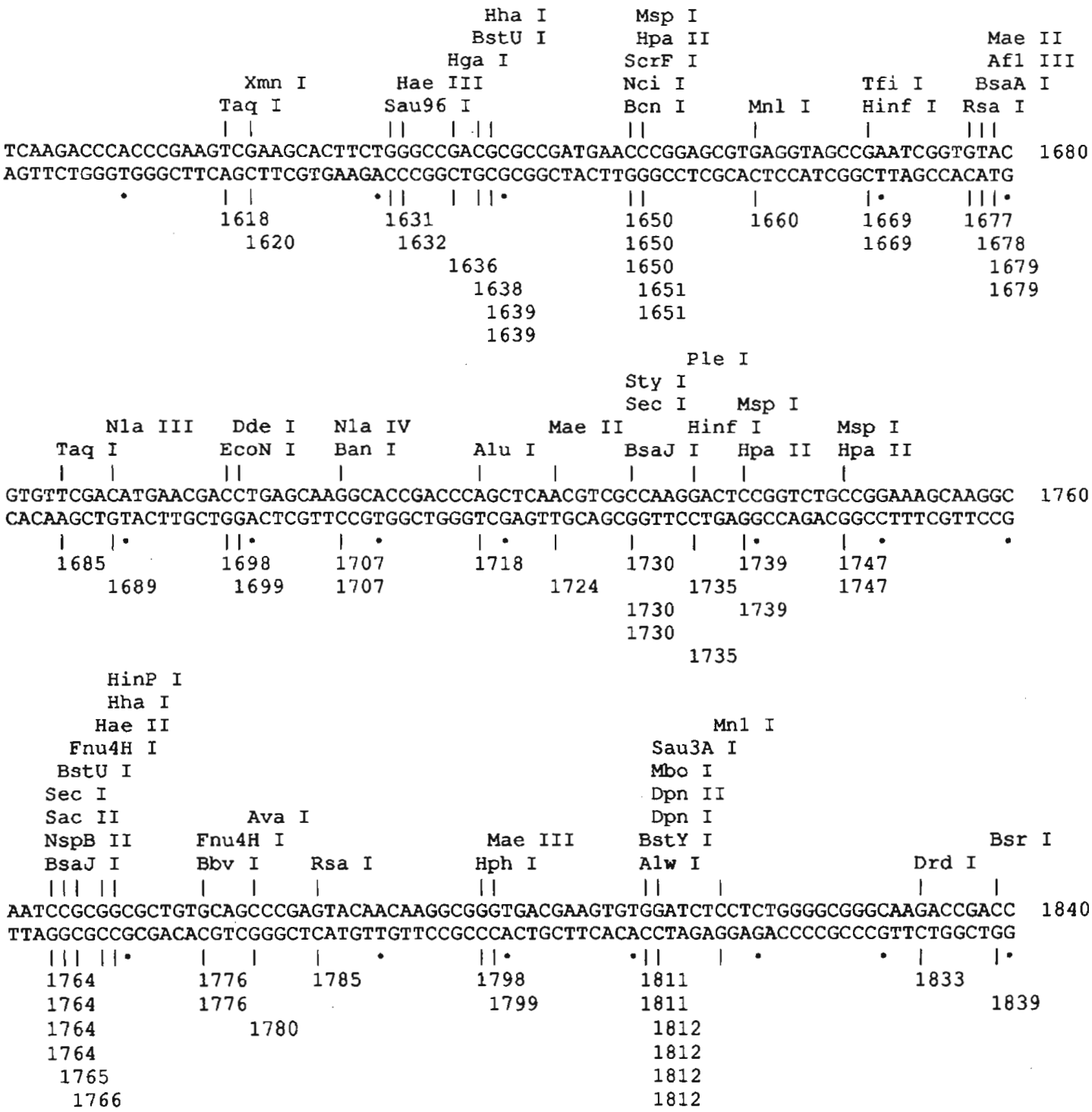

Hha I
BstU I
HinP I
Hha I
BssH II
Msp I
Hpa II
Pst I
Nla III
Nla IV
Alu I
Pvu II
NspB II
TCAACCTGATGGCGCGCTATATCCAGCACACGCCGGACATTCCGCCAGAGTTCTCTCTGCAGGACATGAAGGACAGCTGG 720
AGTTGGACTACCGCGCGATATAGGTCGTGTGCGGCCTGTAAGGCGGTCTCAAGAGAGACGTCCTGTACTTCCTGTCGACC
652
652
652
653
654
654
673
673
697
705
714
714
715
719
Taq I
Xho I
PaeR7 I
Sau3A I
Mbo I
Dpn II
Dpn I
Msp I
Hpa II
Mbo II
Bbv II
Bbs I
Ava I
Mae II
Mnl I
Afl III
HinC II
Xmn I
Hph I
Hga I
Mae III
AACCTGATCGTTCCGGTGGAACGCCGAAGACAGATGAACAAGGTCAACCTCGAGAACGTGTTCGCCATCACCCTGCGTGA 800
TTGGACTAGCAAGGCCACCTTGCGGCTTCTGTCTACTTGTTCCAGTTGGAGCTCTTGCACAAGCGGTAGTGGGACGCACT
726
726
726
726
733
733
746
746
746
763
768
769
769
769
770
774
776
776
788
797
799
Mnl I
Mbo II
Taq I
Alu I
Fnu4H I
Bbv I
HinP I
Hha I
BstU I
Hph I
Sau3A I
Mbo I
Dpn II
Dpn I
BstY I
Bgl II
Sau3A I
Mbo I
Dpn II
Dpn I
Alw I
BstY I
Msp I
Hpa II
Cfr10 I
HgiA I
BstU I
Bsp1286 I
ApaL I
Hga I
BstU I
CGCGCAGCTCTGGGACGGTGATACCCACGAGATCTGGAAGATCCTCGATACCGGCTACGCGGTGCACATCTCGCGTCTGT 880
GCGCGTCGAGACCCTGCCACTATGGGTGCTCTAGACCTTCTAGGAGCTATGGCCGATGCGCCACGTGTAGAGCGCAGACA
801
802
802
804
804
817
830
830
831
831
839
840
840
840
850
851
851
858
862
862
862
872
873

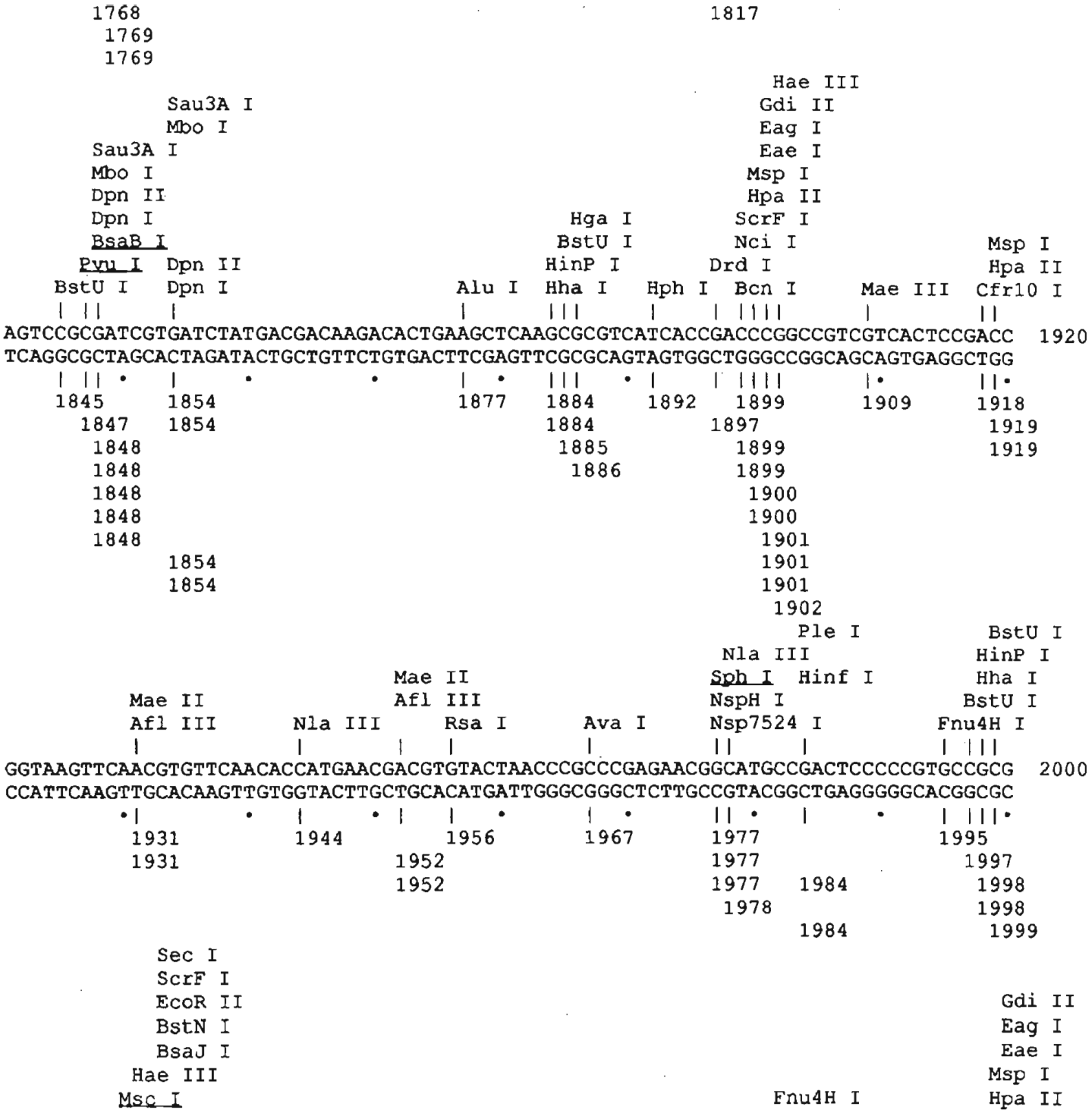

806
831
840
837
845
843
Hae III
Gdi II
Eag I
Eae I
Msp I
Hpa II
Fok I
Msp I
Hpa II
Hae III
Gdi II
Eag I
Eae I
Nla III
NspH I
Nsp7524 I
Afl III
Taq I
Nla IV
Msp I
Hpa II
BspM II
BspE I
Mnl I
Acc I
CCGCCTCCGGCCGTATGTCTACACCGTCGGCCGGATGGCTGACCACCATCATCGACATGTGGTATCCGGAACCGACCACC 960
GGCGGAGGCCGGCATACAGATGTGGCAGCCGGCCTACCGACTGGTGGTAGTAGCTGTACACCATAGGCCTTGGCTGGTGG
884
897
908
932
945
887
908
935
945
887
908
935
946
888
909
935
946
888
911
936
948
888
911
889
913
Sau96 I
Ava II
Nla IV
Sec I
ScrF I
EcoR II
BstU I
Nru I
BstN I
BsaJ I
Sau3A I
Mbo I
Dpn II
Dpn I
Alw I
Dde I
BsmA I
Mae II
Aha II
Aat II
Mbo II
Bbv II
Bbs I
Rsa I
Sau3A I
Mbo I
Dpn II
Dpn I
GTCGCGACCGTTCGCCTGGGTCCGATCCGCTCGGTGGACGTCTCTAAGTTCAAGGGCTACGAAGACAAGTACCTGATCGG 1040
CAGCGCTGGCAAGCGGACCCAGGCTAGGCGAGCCACCTGCAGAGATTCAAGTTCCCGATGCTTCTGTTCATGGACTAGCC
962
963
975
984
997
998
1000
1004
1021
1029
1035
978
979
Xcm I
Hae III
Gdi II
Eae I
Fnu4H I
BstU I
Sec I
Sau3A I
Mbo I
Dpn II
Fnu4H I
Hae III

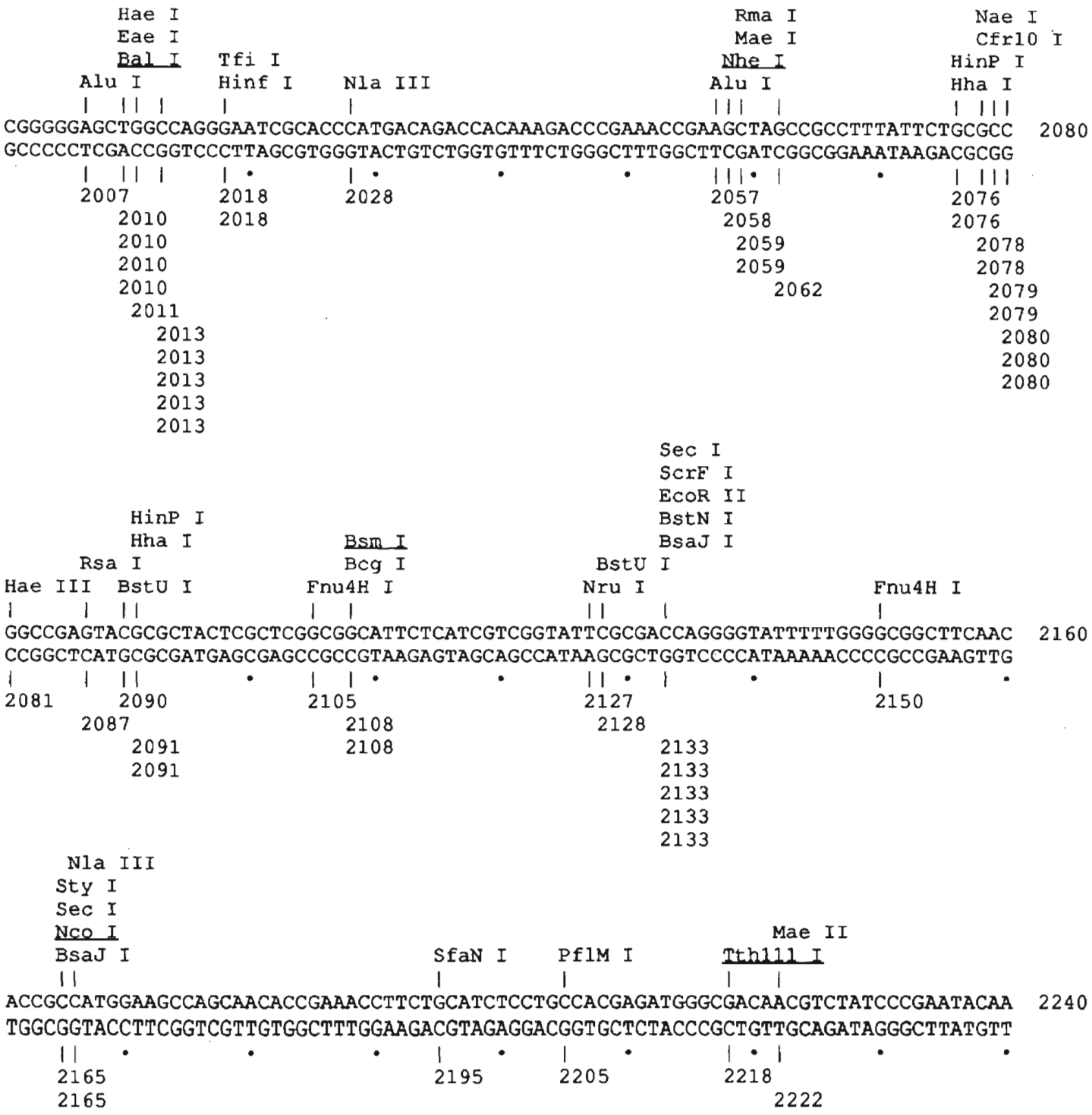

Gdi II
Eae I
Taq I
Rsa I
Nla III
Ple I
Hinf I
Sac II
NspB II
Nla IV
Ban I
Bsr I
Sca I
Dpn I
BsmA I
Nla IV
BsmA I
BsaJ I
TGGCACCTACTGGCCGCCACAGTACTCGATCATGGACGGCGAGACTCTGGAACCGATGAAAGTCGTCTCCACCCGCGGCC 1120
ACCGTGGATGACCGGCGGTGTCATGAGCTAGTACCTGCCGCTCTGAGACCTTGGCTACTTTCAGCAGAGGTGGGCGCCGG
1042 1049 1061 1068 1081 1089 1105 1113
1042 1051 1062 1071 1083 1113
1051 1066 1083 1113
1052 1068 1113
1053 1068 1114
1068 1115
1116
1116
1117
1119
BstU I
HinP I
Hha I
Fnu4H I
Aha II
Taq I
Ava I
BstU I
Hga I
Mnl I
Ava I
Hph I
AGACCGTCGATGGCGACTACCACCCCGAGCCGCGCGTGGCGTCCATCGTCGCCTCGCACATCAAGCCCGAGTGGGTGGTG 1200
TCTGGCAGCTACCGCTGATGGTGGGGCTCGGCGCGCACCGCAGGTAGCAGCGGAGCGTGTAGTTCGGGCTCACCCACCAC
1127 1144 1151 1159 1172 1186 1197
1149 1158
1152
1152
1153
ScrF I
Nci I
Msp I
Hpa II
Bcn I
Sau3A I
Nla III
Taq I
Mbo I
Dpn II
Dpn I
Sal I
HinC II
Acc I
Tfi I
Hinf I
BsmA I
Bsa I
Mae II
Mnl I
Taq I
AACGTGAAGGAGACCGGGCAGATCATGCTGGTCGACTACACCGACATCAAGAACCTCAAGACCACCACCATCGAATCCGC 1280
TTGCACTTCCTCTGGCCCGTCTAGTACGACCAGCTGATGTGGCTGTAGTTCTTGGAGTTCTGGTGGTGGTAGCTTAGGCG
1202 1210 1221 1231 1254 1271
1210 1221 1231 1273
1214 1221 1231 1273
1214 1224
1214 1221 1232
1214

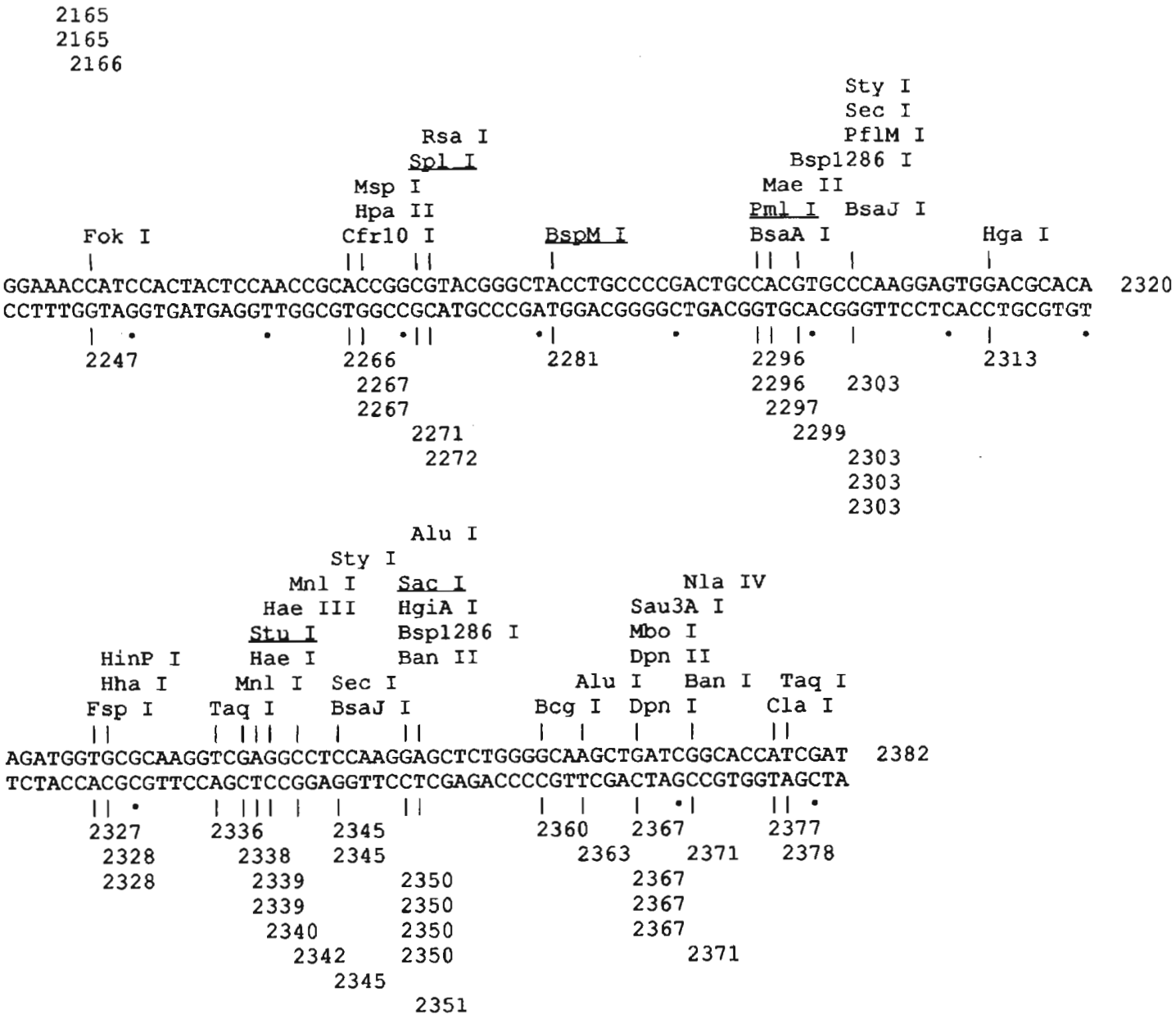

1214
HinP I
Hha I
SfaN I
Fok I
Fnu4H I
Fnu4H I
Fnu4H I
Mnl I
Nla III
Mnl I
Bbv I
CAAGTTCCTGCACGACGGCGGCTGGGATGCCTCCCATCGCTACTTCATGGTCGCCGCCAACGCCTCCAACAAGGCTGCGC 1360
GTTCAAGGACGTGCTGCCGCCGACCCTACGGAGGGTAGCGATGAAGTACCAGCGGCGGTTGCGGAGGTTGTTCCGACGCG
1298
1310
1326
1343
1354
1305
1333
1354
1306
1357
1357
Sau3A I
Mbo I
BstU I
Sau3A I
Dpn II
Hga I
Mbo I
Dpn I
Msp I
Alu I
Dpn II
Alw I
Hpa II
Msp I
Alu I
Taq I
BstY I
ScrF I
HgiA I
Taq I
Hpa II
Fnu4H I
Cla I
Mbo II
Nci I
HinP I
Bsp1286 I
Pst I
Cfr10 I
Bbv I
Dpn I
BstU I
Bcn I
Hha I
ApaL I
CTGCAGTCGATACCAAGACCGGTAAGCTGGCAGCTCTGATCGATACCGCGAAGATCCGCACCCGGACGCGCAACTTCGTG 1440
GACGTCAGCTATGGTTCTGGCCATTCGACCGTCGAGACTAGCTATGGCGCTTCTAGGCGTGGGCCTGCGCGTTGAAGCAC
1361
1378
1390
1398
1407
1421
1428
1438
1367
1379
1390
1399
1410
1421
1428
1438
1379
1392
1400
1412
1421
1438
1385
1398
1413
1422
1398
1413
1422
1398
1413
1425
1413
1427
1413
Sec I
ScrF I
EcoR II
Gdi II
Eae I
Msp I
Hpa II
Msp I
Hpa II
BstN I
ScrF I
Nci I
Cfr10 I
Bcn I
Xcm I
BsaJ I
Tfi I
Hinf I
Sau96 I
Hae III
Sau96 I
Bgl I
Ava II
Hae III
Mae II
Mnl I
Fok I
CACCCGCAGTTCGGCCCGGTATGGTCCACCGGCCACCTGGGCGACGACGTGGTGTCCCTCATCTCCACGCCTTCGGATGA 1520
GTGGGCGTCAAGCCGGGCCATACCAGGTGGCCGGTGGACCCGCTGCTGCACCACAGGGAGTAGAGGTGCGGAAGCCTACT


1453
1453
1455
1455
1455
1456
1456
1463
1463
1466
1468
1469
1469
1470
1470
1471
1472
1476
1476
1476
1476
1487
1497
1515
1519
1519
Nae I
Cfr10 I
HinP I
Nla IV
Nar I
Kas I
Hae II
Bbe I
Ban I
ScrF I
Nci I
Msp I
Msp I
Hpa II
Bcg I
Bcn I
Hpa II
Bsr I
HgiA I
Alu I
SfaN I
Hha I
Rsa I
Rsa I
Bsp1286 I
AlwN I
Mbo II
Aha II
ATCCAAGTACGCCAAGTACAAGGAGCACAACTGGAAGGTGGTGCAGGAGCTGAAGATGCCGGGCGCCGGCAACCTGTTCG 1600
TAGGTTCATGCGGTTCATGTTCCTCGTGTTGACCTTCCACCACGTCCTCGACTTCTACGGCCCGCGGCCGTTGGACAAGC
1527
1536
1543
1543
1550
1564
1568
1572
1575
1579
1579
1579
1579
1579
1582
1582
1582
1582
1582
1582
1582
1583
1583
1585
1585
1586
1586
1589
HinP I


Hha I
BstU I
Hga I
Msp I
Hpa II
ScrF I
Mae II
Afl III
Xmn I
Taq I
Hae III
Sau96 I
Nci I
Bcn I
Mnl I
Tfi I
Hinf I
BsaA I
Rsa I
TCAAGACCCACCCGAAGTCGAAGCACTTCTGGGCCGACGCGCCGATGAACCCGGAGCGTGAGGTAGCCGAATCGGTGTAC 1680
AGTTCTGGGTGGGCTTCAGCTTCGTGAAGACCCGGCTGCGCGGCTACTTGGGCCTCGCACTCCATCGGCTTAGCCACATG
1618
1631
1650
1660
1669
1677
1620
1632
1650
1669
1678
1636
1650
1679
1638
1651
1679
1639
1651
1639
Ple I
Sty I
Sec I
Msp I
Nla III
Dde I
Nla IV
Mae II
Hinf I
Msp I
Taq I
EcoN I
Ban I
Alu I
BsaJ I
Hpa II
Hpa II
GTGTTCGACATGAACGACCTGAGCAAGGCACCGACCCAGCTCAACGTCGCCAAGGACTCCGGTCTGCCGGAAAGCAAGGC 1760
CACAAGCTGTACTTGCTGGACTCGTTCCGTGGCTGGGTCGAGTTGCAGCGGTTCCTGAGGCCAGACGGCCTTTCGTTCCG
1685
1698
1707
1718
1730
1739
1747
1689
1699
1707
1724
1735
1747
1730
1739
1730
1735
HinP I
Hha I
Hae II
Mnl I
Fnu4H I
Sau3A I
BstU I
Mbo I
Sec I
Dpn II
Sac II
Dpn I
Ava I
NspB II
Fnu4H I
Mae III
BstY I
Bsr I
BsaJ I
Bbv I
Rsa I
Hph I
Alw I
Drd I
AATCCGCGGCGCTGTGCAGCCCGAGTACAACAAGGCGGGTGACGAAGTGTGGATCTCCTCTGGGGCGGGCAAGACCGACC 1840
TTAGGCGCCGCGACACGTCGGGCTCATGTTGTTCCGCCCACTGCTTCACACCTAGAGGAGACCCCGCCCGTTCTGGCTGG
1764
1776
1785
1798
1811
1833
1764
1776
1799
1811
1839
1764
1780
1812
1764
1812
1765
1812
1766
1812


1768
1769
1769
1817
Hae III
Sau3A I
Gdi II
Mbo I
Eag I
Sau3A I
Eae I
Mbo I
Msp I
Dpn II
Hpa II
Dpn I
Hga I
ScrF I
BsaB I
BstU I
Nci I
Msp I
Pvu I
Dpn II
HinP I
Drd I
Hpa II
BstU I
Dpn I
Alu I
Hha I
Hph I
Bcn I
Mae III
Cfr10 I
AGTCCGCGATCGTGATCTATGACGACAAGACACTGAAGCTCAAGCGCGTCATCACCGACCCGGCCGTCGTCACTCCGACC 1920
TCAGGCGCTAGCACTAGATACTGCTGTTCTGTGACTTCGAGTTCGCGCAGTAGTGGCTGGGCCGGCAGCAGTGAGGCTGG
1845
1854
1877
1884
1892
1899
1909
1918
1847
1854
1884
1897
1919
1848
1885
1899
1919
1848
1886
1899
1848
1900
1848
1900
1848
1901
1854
1901
1854
1901
1902
Ple I
BstU I
Nla III
HinP I
Mae II
Sph I
Hinf I
Hha I
Mae II
Afl III
NspH I
BstU I
Afl III
Nla III
Rsa I
Ava I
Nsp7524 I
Fnu4H I
GGTAAGTTCAACGTGTTCAACACCATGAACGACGTGTACTAACCCGCCCGAGAACGGCATGCCGACTCCCCCGTGCCGCG 2000
CCATTCAAGTTGCACAAGTTGTGGTACTTGCTGCACATGATTGGGCGGGCTCTTGCCGTACGGCTGAGGGGGCACGGCGC
1931
1944
1956
1967
1977
1995
1931
1952
1977
1997
1952
1977
1984
1998
1978
1998
1984
1999
Sec I
ScrF I
EcoR II
Gdi II
BstN I
Eag I
BsaJ I
Eae I
Hae III
Msp I
Msc I
Fnu4H I
Hpa II


Hae I
Eae I
Bal I
Tfi I
Alu I
Hinf I
Nla III
Rma I
Mae I
Nhe I
Alu I
Nae I
Cfr10 I
HinP I
Hha I
CGGGGGAGCTGGCCAGGGAATCGCACCCATGACAGACCACAAAGACCCGAAACCGAAGCTAGCCGCCTTTATTCTGCGCC 2080
GCCCCCTCGACCGGTCCCTTAGCGTGGGTACTGTCTGGTGTTTCTGGGCTTTGGCTTCGATCGGCGGAAATAAGACGCGG
2007
2018
2028
2057
2076
2010
2018
2058
2076
2010
2059
2078
2010
2059
2078
2010
2062
2079
2011
2079
2013
2080
2013
2080
2013
2080
2013
2013
Sec I
ScrF I
EcoR II
BstN I
BsaJ I
HinP I
Hha I
Bsm I
Bcg I
Rsa I
BstU I
Hae III
BstU I
Fnu4H I
Nru I
Fnu4H I
GGCCGAGTACGCGCTACTCGCTCGGCGGCATTCTCATCGTCGGTATTCGCGACCAGGGGTATTTTTGGGGCGGCTTCAAC 2160
CCGGCTCATGCGCGATGAGCGAGCCGCCGTAAGAGTAGCAGCCATAAGCGCTGGTCCCCATAAAAACCCCGCCGAAGTTG
2081
2090
2105
2127
2150
2087
2108
2128
2091
2108
2133
2091
2133
2133
2133
2133
Nla III
Sty I
Sec I
Nco I
BsaJ I
SfaN I
PflM I
Mae II
Tth111 I
ACCGCCATGGAAGCCAGCAACACCGAAACCTTCTGCATCTCCTGCCACGAGATGGGCGACAACGTCTATCCCGAATACAA 2240
TGGCGGTACCTTCGGTCGTTGTGGCTTTGGAAGACGTAGAGGACGGTGCTCTACCCGCTGTTGCAGATAGGGCTTATGTT
2165
2195
2205
2218
2165
2222


2165
2165
2166
Sty I
Sec I
PflM I
Rsa I
Spl I
Bsp1286 I
Msp I
Mae II
Hpa II
Pml I
BsaJ I
Fok I
Cfr10 I
BspM I
BsaA I
Hga I
GGAAACCATCCACTACTCCAACCGCACCGGCGTACGGGCTACCTGCCCCGACTGCCACGTGCCCAAGGAGTGGACGCACA 2320
CCTTTGGTAGGTGATGAGGTTGGCGTGGCCGCATGCCCGATGGACGGGGCTGACGGTGCACGGGTTCCTCACCTGCGTGT
2247
2266
2267
2267
2271
2272
2281
2296
2296
2297
2299
2303
2303
2303
2313
Alu I
Sty I
Mnl I
Sac I
Nla IV
Hae III
HgiA I
Sau3A I
Stu I
Bsp1286 I
Mbo I
HinP I
Hae I
Ban II
Dpn II
Hha I
Mnl I
Sec I
Alu I
Ban I
Taq I
Fsp I
Taq I
BsaJ I
Bcg I
Dpn I
Cla I
AGATGGTGCGCAAGGTCGAGGCCTCCAAGGAGCTCTGGGGCAAGCTGATCGGCACCATCGAT 2382
TCTACCACGCGTTCCAGCTCCGGAGGTTCCTCGAGACCCCGTTCGACTAGCCGTGGTAGCTA
2327
2328
2328
2336
2338
2339
2339
2340
2342
2345
2345
2345
2350
2350
2350
2351
2360
2363
2367
2367
2367
2367
2371
2371
2377
2378