DOE/ER/13546--4

DE92 018530

46-4

Progress Report
Summary

In prior support periods we identified, cloned and sequenced three genes involved in the regulation of nif gene expression in R. capsulatus. These were called nifR1, nifR2 and nifR4; they turn out to be homologue of the ntrC, ntrB and ntrA genes of enterobacteria. We subsequently found that mutations in an additional gene, nifR5, render R. capsulatus nif genes constitutive with respect to ammonia. The nifR5 gene was shown to be similar to glnB of enteric bacteria, encoding the regulatory protein PII, and furthering the intersection of the glutamine synthetase adenylylation cascade with the control of nif gene transcription. In pursuit of the mechanism of O_2 control of nif gene expression, we constructed and analyzed the topology of a small plasmid in R. capsulatus as a function of O_2 concentration. We also cloned and obtained partial sequence data for two genes encoding the B subunit of DNA gyrase. The nucleotide sequence of the rpoB gene encoding RNA polymerase was nearly completed. A method for isolation of genes expressed differentially, developed for cyanobacteria, was applied successfully to R. capsulatus. Several genes that depend on nifR4 for their transcription were isolated. A transcription start site for a nifA gene was identified and the promoter sequence was analyzed. A physical map of the Rhodobacter capsulatus SB1003 chromosome was prepared, based on pulsed-field electrophoresis of XbaI and AseI fragments and hybridization with a gridded cosmid library, using a device that permits 864 cosmids to be hybridized at one time with a labeled chromosomal fragment.

Personnel

During the past several years, this project was carried out by Ms. Lise Sveen and Ms. Su Zheng, graduate students; Ms. Lori Scappino, technician, and Dr. Francisco Romero and Mikle Fonstein. Romero had a fellowship from the Juan March Foundation. He has returned to Spain and now works for a pharmaceutical company. Ms. Scappino took over the rpoB gene sequencing, left incomplete when Ben Abella went off to college. Dr. Fonstein is an experienced molecular biologist from Moscow who has taken over the mapping program.

Detailed Report

I. The R. capsulatus nifA gene promoter

This part of our program has been disappointing. Based on Kranz' ability to make lacZ fusions to nifR1, nifR2 and nifR4, I thought it would be easy to use the same method (in vivo transposition of Mud lac elements into plasmids containing the target gene) to create nifA fusions. Ms. Sveen succeeded in cloning both copies of Rc nifA into a suitable vector, but when these were used as targets for Mud lac hops, none could be isolated. Then Ms. Zheng joined the group and she began in vitro constructions to insert a promoterless lacZ cassette into nifA. While this was in progress, we received a copy of Hübner, Willison, Vignais and Bickle (1991) which described the construction and use of nifA::lacZ fusions, so we decided not to continue. However, many important questions remain, so we might resume the constructions.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

We also proposed to use S1 nuclease protection and primer extension experiments to define the nifA gene transcription start site. The nifAI gene is downstream by a considerable distance (~750 bp) from nifS; the nifAII gene starts only 250 bp from the C-terminal of nifR4. The two genes are regulated similarly but not identically, according to Hübner et al: nifAI expression is higher than nifAII in a wildtype background and the residual expression of nifAI is higher in the nifR1 mutant background. We found weak and variable signals in S1 nuclease protection experiments on nifAII. With a nifAI probe, there were some irreproducible bands but also one band that showed repeatedly with different RNA preparations, present with RNA from induced cells and absent with RNA from ammonia-grown cells, located at -112 with respect to the start of translation of nifAI. This sequence is shown in Figure 1. The entire region upstream of the RNA end is GC rich. It could not be a start site for the major RNA polymerase (consistent with the lac fusion results) and it does not contain the sequences at -12 and -24 required by the σ^{54}-RNA polymerase (also consistent with lac fusion results). If this is a real start site for RNA polymerase, there must be another unknown σ factor, as suggested by Hübner et al, as well as an upstream binding site for an activating protein.

We attempted to find similarities in the sequence flanking nifAI and the sequences flanking nifAII, nifR4, nifH and any other genes of Rc for which the transcription start sites have been mapped. There were no similarities between the nifAI sequence (750 bp) and the nifAII sequence as far as the latter was determined (200 bp). Short stretches of conserved sequences can be found almost anywhere because Rc DNA is rich in GC and short stretches of G_nC_m abound. The only possibly meaningful similarity found is in the region flanking the pufQ gene, which encodes a protein necessary for assembly of photochemical reaction centers. The operon of which pufQ is the first gene is expressed only under anaerobic conditions but it is not repressed by ammonia and therefore not likely to be regulated by the NIFR1/NIFR2 system. In fact, the puf genes are transcribed in NifR1 mutants.

II. Oxygen control of DNA supercoiling and nifA expression

In order to study the regulation of DNA supercoiling by oxygen, Ms. Sveen constructed a small plasmid, pUCA9, that could be mobilized to transfer into Rc by conjugation. Electrophoresis of plasmid DNA in the presence of quinacrine, comparing plasmids from cells grown anaerobically or with vigorous aeration, showed no difference in the linking number in the two cases. We are forced to conclude that oxygen does not reduce the global degree of DNA supercoiling in Rc. Of course, there could be an effect on local supercoiling, so we intend to insert the nifA::lacZ fusions into pUCA9 in both orientations to explore this possibility.

Oxygen control of nifA expression was not completely determined by Hübner et al. Their experiments compare nifA::lacZ activities in O_2+NH_4^+ with anaerobic, NH_4^+ and anaerobic, N_2; there is no experiment with O_2 under nitrogen repressing conditions. However, in ammonia medium, O_2 reduces nifAI expression five-fold and nifAII expression two-fold. We measured nifA mRNA levels directly by dot blot. Our results are that ammonia reduces nifA (AI + AII) expression about 10-fold under anaerobic conditions and that novobiocin at 100 μg/ml had no effect on the level of nifA mRNA. Moreover, the level of nifA mRNA

was reduced at least 10-fold in the nifR4 mutant LJ1. In the latter mutant, the level of nifR4 mRNA was also reduced several-fold. For all of these blots, the control was a pufQ probe, which detected the same level of pufQ transcript in all conditions in both the wild type and the NifR4 mutant.

III. Purification of RNA polymerase from Rc

This has turned out to be a formidable task. Dr. Romero tried a wide range of growth conditions, methods for cell breakage, and purification steps. He was unsuccessful in obtaining even partially purified enzyme with good activity. We knew from the sequencing presented below that the core subunit β is a little larger than the E. coli β subunit, so he atttempted to purify on the basis of the β subunit detected on protein gels. This did not succeed either.

At the same time, Ms. Sveen was trying to express the nifR4 gene in E. coli. A very large number of attempts failed here too, including the use of the T7 promoter/T7 RNA polymerase system. Eventually, she used the pATH vector with a trp promoter and a leader of trpE protein; this worked well and produced a great deal of hybrid protein. We intend to extend this system to make enough NifR4 protein for antibody production.

IV. The rpoBC operon

We described in previous reports the cloning, mapping and partial sequencing of the rpoBC operon, encoding the largest two subunits of the RNA polymerase core. This was done by Ben Abella when he was in high school and then continued by him during college breaks. The project was taken over by Lori Scappino and her results are shown in the Figure 2 comparison of the Rc and E. coli rpoB sequences. As we show there, the Rc gene is about 300 bp longer than the E. coli gene. The nucleotide sequences are about 60% similar. Usually, where they differ the Rc sequence has a G or C where E. coli has A or T, in keeping with the very high GC content of Rc DNA. We cannot make the amino acid comparison until a few hundred bp remaining of Rc rpoB are determined, to establish the correct reading frame.

V. The DNA gyrase gene(s) of Rc

When R. Kranz was still here, we tried to isolate Rc mutants resistant to coumermycin or novobiocin, in order to get a handle on the structural genes for DNA gyrase. All the mutants Kranz found were transport mutants; they all had drug-sensitive gyrase. We dropped the idea until W.M. Huang asked us for Rc DNA in order to attempt to amplify the gyrB gene using PCR primers from highly conserved regions of bacterial gyrB genes that she discovered. We sent the DNA and soon after she returned to us two cloned PCR fragments with slightly different sequences. Ms. Sveen used these to isolate, from a LORIST 6 cosmid library, two different gyrB genes. Sequencing these genes and studies of their transcription are in progress. We showed long ago that the level of the gyrase subunit proteins increases in cells treated with novobiocin; as the nif gene products disappear, the gyr gene products are induced.

VI. The physical map of the Rhodobacter capsulatus chromosome

The program to create a high resolution physical map of the Rc chromosome envisions the following steps: construction of a low resolution map using pulsed field electrophoresis (PFE); construction of a representative cosmid library; grouping cosmids in sets by hybridization with fragments eluted from the PFE gels; analysis of restriction patterns of the cosmids in each group to generate a unique set of overlapping cosmids; final mapping of the cosmid set to generate the chromosome map. We have nearly finished the first three parts of this program.

Among 28 restriction endonucleases checked for rare cutting, the following were useful (Figure 3): XbaI, 15 fragments; AseI, 25; SpeI, 3; ScaI, 15 and DraI, 20-25. XbaI and AseI were selected for the low resolution mapping. Figure 4 shows one step of this procedure, in which a single XbaI fragment is used as probe to identify the fragments in AseI and AseI/XbaI digests that overlap. The resulting map is also shown for this region of the Rc chromosome in Fig. 4.

The cosmid library was constructed in the vector LORIST 6, which is stable and accommodates a large insert size. Sau3a fragments of Rc chromosomal DNA were size-fractionated, dephosphorylated and cloned into cosmid arms. 1728 individual clones were picked and grown in the wells of microtiter plates. A device was constructed to facilitate the transfer of 864 colonies from the wells to 140 mm petri dishes containing nylon filters (Figure 5). After growth, the colonies can be lysed in situ and the DNA probed with individual eluted bands from the PFE gels. An example of this procedure is shown in the autoradiogram of Figure 6. The autoradiogram contains 864 possible sites, of which 37 are positive in this test with XbaI band (10) as probe. The autoradiogram is read on a tablet with a position-detecting pen coupled to a Macintosh, in a manner analogous to reading DNA sequencing gels. The positions are stored in the Macintosh. Two autoradiograms can be compared, using software written here by W. Buikema, and colonies containing cosmids identified by two probes can be identified rapidly. These cosmids provide the physical linkage between the two fragments used as probes. In this way, the physical map obtained as shown in Figure 5 can be extended. Our current map of the Rc chromosome is shown in Figure 7. At present there is one very large linkage group and two small ones. Each of the smaller ones could be a plasmid; alternatively they could be linked to the larger chromosome. Further probing of the cosmid library will determine which of these possibilities is correct.

We have also begun to locate cloned genes on the physical map. Figure 7 shows the assignments made to date. Of significance among these early results are the finding that the various clusters of nif genes are widely distributed on the large chromosome, consistent with the first genetic mapping by Willison et al.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

AACTGCCGTGCGGTGAAACGGCGTCGAGGCCGGGGATCTGCGCGCA
TGGTGCCGCTCGCCCGCCCGTCGCCCCGAACCTGATTGCCCGAACGT
GAATGAAGGAGACCCAGATGACCCCCGAAAGCCCCACGCTTGCGG
CGTTGACGAAACTGTCCTCGGCCGAGGAAATCTTCGCCTTCCTGGGC
GTCGAGCCGATCCGCGAGGTGCTGAACTCGTCCCGGCTGCACATCAT
GACAGCTTCGCGCCTATCTGCGCGAGACCGACATGACCGGCCTGAC
CGAAGACGGGATCTTTGAGCGCCGCGATGCGCTGCTGCGGGCGCAG
GCCGATTTCGTCGCCTCGACGCCGCTCAAGGAAAAGGTCTTCAAGG
TCTTCGAAACCGAGGCCGCGAAACGCAAAGCCGCTTCGTCGGCCTTGA
AACGCTGAAAGTGATCAAATCCTGACCGCCTGCAATCTGCCAGAT
TTGGTCCCCCGGGGCCGGTTCGCCGCATAATTGCGCAATCGACCCGG
TGACCTTGCAAAAATGAACCGAACTTGCCTGCCGGACGCGTCGGA
AGACGTGCCTTTTTTCGCCCATGCCCTATGGTCAGTGCAGAAGTCCC
GACTGCAGCGGCTGCCGCCCGAGTCCCTTGCCCCGGGTCATCCGCCG -110
-108 CGGAAATTCCTTCCTGTGACGGGGGTGACTGCATGACCGATCAACA
-12 ATCCCGCCCCGCCTCGCCGCGGCGTCAGCTCGACGAGTCGATCGCCG
-15 ACCGCCTTGCCTTGGatg

Figure 1. Nucleotide sequence upstream of the nifA1 gene of Rhodobacter capsulatus. This sequence was determined in Chicago; it agrees exactly with the sequence published by W. Klipp. The atg marks the beginning of the open reading frame. Three S1 nuclease protection experiments, using RNA from cells induced for nitrogenase under anaerobic conditions, revealed the major band of protected DNA at position -110, the indicated C residue. The sequence 5' to this possible start site is very GC rich and could not be recognized by the major sigma-70 RNA polymerase. On the other hand, the sequences at -12 and -24, underlined, do not conform to the Klebsiella consensus for those regions, which is TGGYRYR at -24 and TGCA at -12. Much further upstream there are several TGG sequences, but they are not followed by appropriately spaced TGCAs. All this to reenforce the conclusion that nifA in Rhodobacter is transcribed by yet another form of RNA polymerase.

Figure 2.

The following is a comparison of the *E.coli rpoB* nucleotide sequence to an incomplete *Rhodobacter capsulatus rpoB* nucleotide sequence. The comparisons were made via FastA and were carried out with these restrictions: 1) FastA limits the size of the sequences that it can handle to less than 4000 nucleotides. The *E. coli* sequence was divided into two segments; a seqment that ranged from 1 to 2500 and the second segment that overlapped the first and covered nucleotides 2001 to 4059. The *R. capsulatus* sequence was divided into four overlapping segments. The first, nucleotides 1 to 1000, second, 1 to 2500, third, 2001 to 4284 and fourth, 2871 to 4284, were compared to the *E. coli* "library" in turn. The results were joined together to show a complete comparison of the two sequences. The peptide sequence could not be used because the nucleotide sequence is not yet complete. According to the FastA results, the *R. capsulatus rpoB* nucleotide sequence identified with the *E.coli* sequence with these percentages:

1)	1-1000	59.0%
2)	1-2500	61.5%
3)	2001-4284	65.0%
4)	2871-4284	66.0%

KEY: EC: *E. coli*
RC: *R. capsulatus*

```
                                                                  10
EC                                                            ATGGGTCGACT

                  10         20         30         40         50
RC      AAGCTTACGTTGG-TCAGAAGCGTATCCGCCGCTACTACGGCAAAATCCG--CG--AA
                :   ::  :   :  :: :   ::  ::       ::::      ::  ::
EC   TGTCAGCGAGCTGAGGAACCCTATGGTTTACTCCTATACCGAGAAAAAACGTATTCG-TAA
           20         30         40         50         60         70
           60         70         80         90        100
RC   G--TTCTGG-AAATGCCGAACCTCATCGAGG-TTCAGAAGGCCTCCT--ACGACC-TGTT-
     :  :: ::: ::: :      :   :   :: :  :      ::::: :    ::   ::
EC   GGATTTTGGTAAACG--TCCACAAGTTCTGGATGTACCTTATCTCCTTTCTATCCAGCTTG
           80         90        100        110        120        130
        110        120        130        140        150        160
RC   CCTGCG-CTC-GGGCGATGCCGATC-GGCCGCTGGACGGC--GAGGGCATTCAGGGCGTCT
      :: ::  :: :     : ::: :  :   ::: : :::   : ::  :  : :  :  :
EC   ACT-CGTTTCAGAAATTTATCGAGCAAGATCCTGAAGGGCAGTATGGTCTGGAAGCTGCTT
            140        150        160        170        180        190

            170        180        190        200        210        220
RC   TCCAGTCGGTCTTCCCGATCAAGGACTTCAATGACAACGCGATCCTCGAATTCGTGAAAT
     :::  :: :: :::::::  ::  :: ::  :  ::  :    ::  ::: ::: :  :
EC   TCCGTTCCGTATTCCCGATTCAGAGCTACAGCGGTAATTCCGAGCTGCAATACGTCAGCT
            200        210        220        230        240        250
```

```
          230       240       250       260       270       280
RC   ACGAGCTCGAAAAGCCGAAATACGACGTCGACGAATGCATGAGCCGCGACCTGACCTATT
     ::   :: :   : :::   :  :::::: : :::::    : ::: : : :::::::::
EC   ACCGCCTTGGCGAACCGGTGTTTGACGTCCAGGAATGTCAAATCCGTGGCGTGACCTATT
            260       270       280       290       300       310

          290       300       310       320       330       340
RC   CGGCGCCGCTCAAGGTGACGCTGCGTCTGATCGTGTTCGATGTCGATGAAACCACCGGGG
     : :: :::::    :: :  :::::::::   ::: :: :::  :    :: : ::::
EC   CCGCACCGCTGCGCGTTAAACTGCGTCTG---GTGATCTATG-AGCGCGAAGCGCCGGAA
            320       330           340       350       360

          350       360       370       380       390       400
RC   CGCGGTCGGTCAAGGACATCAAGGAACAGGACGTCTACATGGGCGACATGCCCCTGATGA
      ::   : :: :: X:::: :: ::::: :: :::::::::::::: :: :: :: ::::
EC   GGC--ACCGTAAAAGACATTAAAGAACAAGAAGTCTACATGGGCGAAATTCCGCTCATGA
        370       380       390       400       410       420

          410       420       430       440       450       460
RC   CGCCGAACGGGACCTTCATCGTCAACGGCACCGAACGCGTGATCGTGTCGCAGATGCACC
     :    ::::: :::::  :  ::::::: :: :: :: :: ::::: :: ::: ::::::
EC   CAGACAACGGTACCTTTGTTATCAACGGTACTGAGCGTGTTATCGTTTCCCAGCTGCACC
        430       440       450       460       470       480

          470       480       490       500       510       520
RC   GCTCGCCCGGTGTGTTCTTCGACCATGACCGCGGCAAGACCCATTCCTCGGGCAAGCTTC
     :    :: :: :: ::::: :::   :::   :: :: ::::: :: ::::: ::  : :
EC   GTAGTCCGGGCGTCTTCTTTGACTCCGACAAAGGTAAAACCCACTCTTCGGGTAAAGTGC
        490       500       510       520       530       540

          530       540       550       560       570       580
RC   TGTTCGCCTGCCGGATCATTCCCTATCGCGGCTCGTGGCTGGATTTCGAATTCGACGCCA
     :::    :   :: ::::: :: :: :: :: :: :::::::: ::::::::::  : :
EC   TGTATAACGCGCGTATCATCCCTTACCGTGGTTCCTGGCTGGACTTCGAATTCGATCCGA
        550       560       570       580       590       600

          590       600       610       620       630       640
RC   AGGATCTGGTGTTCGCGCGCATCGACCGTCGCCGCAAGCTGCCGGTGACGACGCTGCTTA
     ::::     ::::::  :: :::::::::::::: :: ::::: : ::: :   : ::
EC   AGGACAACCTGTTCGTACGTATCGACCGTCGCCGTAAACTGCCTGCGACCATCATTCTGC
        610       620       630       640       650       660

          650       660       670       680       690
RC   ATGCGCTCGGCCTCGATCA-GGAAGGCATCATGGA--TGCCTATTACAACACCGTCACCT
       :: :X :  ::  : ::  ::    ::: : ::  ::  : ::  :: :   :  : :
EC   GCGCCCT-GAACTACACCACAGAGCAGATCCTCGACCTG-TTCTTTGAAAAAGTTATCTT
        670       680       690       700       710       720
```

```
   700        710        720        730        740        750
RC   TCCGCGCCGAGGGCAAACGCGGCTGGGTGTCGAAATTCTTCCCCGAGCGGGTGCGCG--G
     :      ::: :    : : :: ::      :  : :: :  : :: :: ::  :::: :  :
EC   TGAAATCCGTGATAACAAGCTGC----AGATGGAACTGGTGCCGGAACGCCTGCGTGGTG
          730        740            750        760        770

     760        770        780        790        800        810
RC   CACCCGTCCGACCTATGACCTGATCGACGCCGCTTCGGGCGAAGTGATCCTGAAGGCGGG
      : :::  :  : : ::::   :::::    :::   ::  :::::  : :  :    ::
EC   AAACCG--CATCTTTTGAC---ATCGA---AGCTAACGGTAAAGTGTACGTAGAAAAAGG
    780          790           800           810        820        830

     820        830        840        850        860        870
RC   CGAAAAGGTCACGCCGCGGATGGTGAAGAAATGGCTTGATGAAGCCAAGGTGACCGAACT
     :        ::::  ::::      :      :   ::: ::  ::: : : ::   : ::::
EC   CCGCCGTATCACTGCGCGCCACATTCGCCA---GCTGGAAAAAGACGACGT---CAAACT
            840        850        860           870           880

     880           890        900        910        920        930
RC   GCTC---TTGCCGTTCGACCACATCCTCGGCCGCTACTGCGCCAAGGACATCATCAACGA
     : ::    : ::: : ::  :::::   ::           :: :: :::   ::  : ::
EC   GATCGAAGTCCCGGTTGAGTACATCGCAGGTAAAGTGGTTGCTAAAGACTATATTGATGA
        890        900        910        920        930        940

        940        950        960        970        980        990
RC   GCGCACCGGCTACATCTACATCGAAGCGGGCGACGAGCTGA-CGGCGGAATACGACCGTG
     :   ::::::                          ::::::: : ::: :   ::: : ::
EC   GTCTACCGGC-------------------------GAGCTGATCTGCGCA--GCGAACATG
        950                                 960         970

        1000        1010        1020      1030      1040      1050
RC   ACGGCCA-ACTGG-TCGGC-GGTTCGGTGAAGACGCTTCTGGACAACGGCATCACCGAGA
       :   :  :::: :: :: :: :  :   :: :   ::::: ::   :: : : ::: :
EC   GAGCTGAGCCTGGATCTGCTGGCTAAGCTGAG-CCAGTCTGGTCACAAGCGT-ATCGAAA
    980        990      1000      1010        1020        1030

            1060        1070      1080      1090      1100
RC   TCCCCTGTCTCGACATCGA-CAACATCAACGTCGGCCCCTACATCCGCAACACGATGGCG
       : :::: ::  :: ::: :   ::::    :::::: :: :::    : ::  :
EC   --CGCTGT-TCACCAACGATCTGGATCA----CGGCCCATATATCTCTGAAACCTTACGT
       1040        1050      1060          1070      1080

  1110      1120        1130      1140      1150      1160
RC   GCGGACAAGAACATGGGCCG-CGACGCGCGCTGATGGACATCTACCGCGTCATGCGTCCG
     :  :::   :  :  : ::: :   X::: ::: : :: :::::::: : ::::: ::
EC   GTCGACCCAACTAACGACCGTCTGAGCGCACTGGTAGAAATCTACCGCATGATGCGCCCT
  1090      1100      1110      1120      1130      1140
```

```
   1170      1180      1190      1200      1210      1220
RC   GGGGAACCGCCGACCGTCGAAGCCGCAAGCTCGATGTTCGACACGCTTTTCTTCGATTCG
     :: :: ::::::::    ::::: ::        ::::::: :  :: ::::::
EC   GGCGAGCCGCCGACTCGTGAAGCAGCTGAAAGCCTGTTCGAGAACCTGTTCTTCTCCGAA
  1150      1160      1170      1180      1190      1200

   1230      1240      1250      1260      1270      1280
RC   GAACGCTATGACCTCTCGGCCGTGGGCCGGGTGAAGATGAACATGCGTCTGGATCTGGAT
     :: :: :::::: : :: :: :: :: ::  ::::: : :::     ::::   :  ::
EC   GACCGTTATGACTTGTCTGCGGTTGGTCGTATGAAGTTCAACCGTTCTCTGCTGCGCGAA
  1210      1220      1230      1240      1250      1260

   1290      1300      1310      1320      1330      1340
RC   GCGCCGGACACCCAGCGCACCCTGCGCCGCGCCGACATCATCGCCTGCATCAAGGGTCTG
      :      ::        : : :::: ::   : ::::::::: :     :: ::    ::
EC   GAAATCGAAGGTTCCGGTATCCTGAGCAAAGACGACATCATTGATGTTATGAAAAAGCTC
  1270      1280      1290      1300      1310      1320

   1350      1360      1370      1380      1390      1400
RC   GTGGAGCTGCGCGACGGCAAGGGCGAGATCGACGACATCGACCACCTCGGCAACCGCCGC
      : ::  : ::  :::: :: :::::  :::: :: :::::::::::::::::::: ::
EC   ATCGATATCCGTAACGGTAAAGGCGAAGTCGATGATATCGACCACCTCGGCAACCGTCGT
  1330      1340      1350      1360      1370      1380

   1410      1420      1430      1440      1450      1460
RC   GTGCGCTCGGTCGGCGAGCTGATGGAAAACCAGTATCGCGTCGGCCTTCTGCGCATGGAA
      : :: :: :: :::::  ::  ::::::::::  ::::: :::::  : ::  : ::
EC   ATCCGTTCCGTTGGCGAAATGGCGGAAAACCAGTTCCGCGTTGGCCTGGTACGTGTAGAG
   390       1400      1410      1420      1430      1440

   1470      1480      1490      1500      1510      1520
RC   CGGGCGATCAAGGAACGCATGTCCTCGGTCGAGATCGACACGGTGATGCCGCAGGACCTG
     :: ::: : :: :: ::  ::::   :: :::  : :: ::  ::::::: :::::  ::
EC   CGTGCGGTGAAAGAGCGTCTGTCTCTGGGCGATCTGGATACCCTGATGCCACAGGATATG
  1450      1460      1470      1480      1490      1500

   1530      1540        1550      1560      1570      1580
RC   ATCAACGCGAAACCG--GCGCGGGGCGGTGCGCGAATTCTTCGGCTCCTCGCAGCTGTCG
     :::::::: :: ::X     ::  :: :::   :: :::::::: :::   ::::::::
EC   ATCAACGCCAAGCCGATTTCCGCAGCAGTGAAAGAGTTCTTCGGTTCCAGCCAGCTGTCT
  1510      1520      1530      1540      1550      1560

     1590      1600      1610      1620      1630      1640
RC   CTGTTCATGGACCAGACCAACCCGCTGTCGGAAGTCACGCACAAGCG-CGTCTTTCGGCG
     : ::: :::::::::: :::::::::::: ::  : ::::::: :: ::: : :: ::
EC   CAGTTTATGGACCAGAACAACCCGCTGTCTGAGATTACGCACAAACGTCGTATCTCCGCA
  1570      1580      1590      1600      1610      1620
```

```
      1650      1660          1670      1680      1690      1700
RC    CTCGGGCCGGGCGGTCTGAC---GGAACGCGCGGGCTTCGAAGTGCGCGACGTTCACCCG
      ::::: :: :::::::::::    ::::: :: ::::::::::: :: ::::: ::::::
EC    CTCGGCCCAGGCGGTCTGACCCGTGAACGTGCAGGCTTCGAAGTTCGAGACGTACACCCG
  1630      1640      1650      1660      1670      1680

          1710      1720      1730      1740      1750      1760
RC    ACCCACTACGGCCGGATGTGCCCCATTGAGACGCCCGAAGGTCAGAACATCGGTCTGATC
      :: :::::::: ::  : :: :: :: :: :: :: ::::::: ::::::::::::::::
EC    ACTCACTACGGTCGCGTATGTCCAATCGAAACCCCTGAAGGTCCGAACATCGGTCTGATC
  1690      1700      1710      1720      1730      1740

          1770      1780      1790      1800      1810      1820
RC    AACTCGTTGGCCACTTTCGCGCGCGTCAACAAATACGGCTTCATCGAGACCCCCTATCGC
      :::::  :: ::   : ::: :     ::: :::::::::: : ::::: :: :::::
EC    AACTCTCTGTCCGTGTACGCACAGACTAACGAATACGGCTTCCTTGAGACTCCGTATCGT
  1750      1760      1770      1780      1790      1800

          1830      1840      1850      1860      1870      1880
RC    AAGGTCCACGATGGCCAGGTGACCGACGAAGTCGTCTACATGTCGGCCACCGAAGAGATG
      :: ::   ::: ::    :: :: :::::: :   :::: :::: :: : ::::::
EC    AAAGTGACCGACGGTGTTGTAACTGACGAAATTCACTACCTGTCTGCTATCGAAGAAGGC
  1810      1820      1830      1840      1850      1860

          1890      1900      1910      1920      1930
RC    CGCCACACGATCGCGCAGGCGAACGCCACGCTCGACGAGAACGGCAAGCTTCGT-GA--A
        : ::   ::::: ::::::::: :::   : :: ::  : ::: : :::::: ::  :
EC    AACTACGTTATCGCCCAGGCGAACTCCAACTTGGATGAAGAAGGCCA-CTTCGTAGAAGA
  1870      1880      1890      1900      1910      1920

  1940      1950      1960      1970      1980      1990
RC    CC--GCATCTCGCGTCCTCCGACCCG--TCCAGGCCGCGCCGCACATTCCCATCGCCTCA
      ::  : : :: ::     :  :  ::  :::::   :  : :: :    :::      :
EC    CCTGGTAACTTGCCGTAGCAAAGGCGAATCCAGCTTGTTCAGC-CGCGACCAGGTTGACT
    1930      1940      1950      1960      1970      1980

      2000      2010      2020      2030      2040      2050
RC    ACCCGCGCTCGGCAAGCCGTGGCGCCTGCACTCCGACGTGTCG--CCGAAACACGCTGCG
      ::  :  :    : : ::   ::   ::   :::: ::   ::  ::  : : :: : :
EC    ACATGGACGTATCCACCC--AGCAGGTGGTATCCGTCGGTGCGTCCCTGATCCCGTTCC-
    1990      2000      2010      2020      2030      2040

          2060      2070      2080      2090      2100      2110
RC    TTTCGGTCGCCGCGGCGCTGCATCCCTTCCTTGA--CAAACGACGACGCAACGCCGCCTC
        : :  : :   : :::  :: :: : : ::::      :::  : ::::::     ::
EC    --TGGAACACGATGACGC--CAACCGTGCATTGATGGGTGCGAACATGCAACG-----TC
        2050      2060      2070      2080      2090
```

```
          2120      2130        2140          2150      2160
RC   ATCGCGGTTCGAACATGCACAC--GTAGGCC---GGTT-CCGCTTCTGCGCAACGAGGCG
     :   :: : :::   ::: : :   :: :::   ::::    ::  :  : :::: :  :
EC   AGGCCGTTCCGACTCTGCGCGCTGATAAGCCGCTGGTTGGTACTGGTATGGAACGTGCTG
      2100      2110      2120      2130      2140      2150

      2170      2180      2190      2200      2210      2220
RC   CGTTCGTCGGCACCGGCATGGAAGCCACGGTGGCGCGCGATTCGGGGGCCGCGATCATGG
         ::: : : ::::  :: :: :  ::::   ::   :   :: ::   :: :::  :
EC   TTGCCGTTGACTCCGG--TGTAA-CTGCGGT--AGCTAAACGTGGTGGTGTCGTTCA--G
      2160      2170         2180        2190      2200

      2230      2240      2250      2260      2270      2280
RC   CCCGCCGCGGCGATCGTGGATCAGGTCGATGCGAGCCGGATCGTTATCCGCGCCAATGAA
       ::  :  ::  ::  : :::  ::  ::         :  :::: :   : : : ::
EC   TACGTGGATGC-TTCCCGTATC--GT-TAT--------CAAAGTTAACGAAGACGA-GAT
   2210      2220        2230                2240      2250

      2290      2300      2310      2320      2330      2340
RC   GACATCGGCATGGGCACGCGGGCGTCGACATCTATCGTCTGCGCAAGTTCAAACGCTCGA
     :  ::: :   ::  : :: ::  :::::::::   :::  ::: : ::  :: :: :
EC   G-TATCCG---GGTGAAGCAGGTATCGACATCTACAACCTGACCAAATACACCCGTTCTA
       2260        2270      2280      2290      2300      2310

      2350      2360      2370      2380      2390      2400
RC   ACCAGTCCTCGACGATCAACCAGCGTCCGATCGTCAAGGTGGGCGACATCGTTGCCAAGG
     :::::  : :    :::::::::   :::   ::    :::: ::    ::::     :
EC   ACCAGAACACCTGTATCAACCAGATGCCGTGTGTGTCTCTGGGTGAACCGGTTGAACGTG
          2320      2330      2340      2350      2360      2370

      2410      2420      2430      2440      2450      2460
RC   GTCAGGTCGTGGCCGACGGTCCCTCGACCGATCAGGGCGAACTGGCGCTGGGGCGCAACG
     :  : ::  :::: :::::::: :: ::::: :  :: :::::::::: :: :  :::
EC   GCGACGTGCTGGCAGACGGTCCGTCCACCGACCTCGGTGAACTGGCGCTTGGTCAGAACA
          2380      2390      2400      2410      2420      2430

      2470      2480      2490      2500      2510      2520
RC   TGGTCGTCGCCTTCATGCCCTGGAACGGCTACAACTACGAAGACTCCATCCTGATTTCGG
     ::  ::: :: :::::::: ::::: :: ::::::: :::::::::::::  : :: :
EC   TGCGCGTAGCGTTCATGCCGTGGAATGGTTACAACTTCGAAGACTCCATCCTCGTATCCG
          2440      2450      2460      2470      2480      2490

        2530      2540      2550
RC   AACGCATCCACCGCGATGACGTGTTCAC
     :  :  :    :  :: :::   :::::
EC   AGCGTGTTGTTCAGGAAGACCGTTTCAC
          2500      2510
```

```
         2560      2570      2580      2590      2600      2610
RC   CTCGATCCATATCGAGGAATACGAAGTGGCGGCGCGCGACACGAAACTCGGGCCGGAGGA
     : : ::::: ::  :::::   :     : : : :: ::::: :: :: :: ::::: ::
EC   CACCATCCACATTCAGGAACTGGCGTGTGTGTCCCGTGACACCAAGCTGGGTCCGGAAGA
  2520      2530      2540      2550      2560      2570

        2620      2630      2640      2650      2660      2670
RC   GATCACCCGCGATATCCCGAACGTCGGCGAGGAAGCGCTGCGCAACCTCGACGAGGCGGG
     :::::::   :: :::::::::: :: :: :  :::::   ::: :: :: ::  : ::
EC   GATCACCGCTGACATCCCGAACGTGGGTGAAGCTGCGCTCTCCAAACTGGATGAATCCGG
  2580      2590      2600      2610      2620      2630

        2680      2690      2700        2710      2720      2730
RC   CATCGTCTATATCGGCGCCGAAGTCGGTCCG--GGCGATATCCTTGTGGGCAAGATCACG
      ::::: :: :: :: :: :::::  : :::  ::::: :: :: :: :: ::: : :::
EC   TATCGTTTACATTGGTGCGGAAGT--GACCGGTGGCGACATTCTGGTTGGTAAGGTAACG
  2640      2650      2660        2670      2680      2690

            2740      2750      2760      2770      2780      2790
RC   CCG-AAGGCGAAAGCCCGATGACGCCGGAAGAAAAACTTCTGCGCGCCATCTTCGGTGAA
     ::: :::: ::::  v^; :::: :: :::::::::::::::: :: :::::::::::
EC   CCGAAAGGTGAAACTCAGCTGACCCCAGAAGAAAAACTGCTGCGTGCGATCTTCGGTGAG
    2700      2710      2720      2730      2740      2750

            2800      2810      2820      2830      2840      2850
RC   AAGGCCTCGGACGTGCGGGATACCTCGCTGCGTCTGCCGCCGGGCGCCTATGGCACGATC
     :: ::::: :::::    ::  : :: :::::  : ::    :: :  :  :: ::: :
EC   AAAGCCTCTGACGTTAAAGACTCTTCTCTGCGCGTACCAAACGGTGTATCCGGTACGGTT
    2760      2770      2780      2790      2800      2810

            2860      2870      2880      2890      2900      2910
RC   GTCGAAGTCCGCGTCTTCAACCGTCATGGTGTGGACAAGGACGAGCGCGCGCTGCAGATC
      :::: :: :  ::::: :  ::  :::: :: :: :: ::: : :: :::::: : :::
EC   ATCGACGTTCAGGTCTTTACTCGCGATGGCGTAGAAAAAGACAAACGTGCGCTGGAAATC
    2820      2830      2840      2850      2860      2870

            2920      2930      2940      2950      2960
RC   GAGCGTGAGGAGATCGAACGTCTCGCCCGCGACCGCGACGACGAGATGGCGATT--CGAA
     ::       : :: :: :::             :::     :: ::  ::  :::   ::::
EC   GAAGAAATGCAGCTCAAACAGGCGAAGAAAGACCTGTCTGAAGAACTGCAGATCCTCGAA
    2880      2890      2900      2910      2920      2930

RC 2970      2980      2990      3000      3010      3020      3030
     CAACATCTATGCGC-TCTGAAAACGCTGATCCTTGGCAAGGTGACGGTCAAAGGGCCGAA
                :::  ::::   :  :  :::: :
EC   ----------GCGGGTCTGTTCAGCCGTATCCGT---------------------------
```

```
RC   GAACACC

EC   -------

       3040      3050      3060      3070      3080      3090
RC   TCGATCAAGGCCGGGTCGGAAATCACCGAAGAGCTGCTGGCGTCGCTGTCGCGTGGCCAG
      :  :   ::             :        :::  ::: :     ::: :::: :  :
EC   GCTGTGCTGGTAGCCGGTGGCGTTGAAGCTGAGAAGCTCGACAAACTGCCGCGCGATCGC
            970       980       990      1000      1010      1020

      3100      3110      3120      3130      3140      3150
RC   TGGTGGCAGCTGGCGCTGGAAAACGAAGCCGATGCCAAGGAAGTGGAAGCGCTGAACGAC
     :::  : ::::::  :::  : ::::::   :    ::  :  :::::  ::::   ::
EC   TGGCTGGAGCTGGGCCTGACAGACGAAGAGAAACAAAATCAGCTGGAACAGCTGGCTGAG
           1030      1040      1050      1060      1070      1080

      3160      3170      3180      3190      3200      3210
RC   CAGTTCGACGCGCAGAAACGCGCCCTTGAGCACCGCTTCGACGACAAGGTGGAGAAAGTC
     ::::  ::::  : ::::: ::   : ::: :     :::: :  ::       ::: ::
EC   CAGTATGACGAACTGAAACACGAGTTCGAGAAGAAACTCGAAGCGAAACGCCGCAAAATC
           1090      1100      1110      1120      1130      1140

      3220      3230      3240      3250      3260      3270
RC   CGTCAGGGCGACGATCTGCCCCCCGGCGTGATGAAGATGGTCAAGGTCTTCGTCGCGGTG
       :::::::::::::: : :: ::::: ::::::: :: ::::: :   : :::::
EC   ACCCAGGGCGACGATCTGGCACCGGGCGTGCTGAAGATTGTTAAGGTATATCTGGCGGTT
           1150      1160      1170      1180      1190      1200

      3280      3290      3300      3310      3320      3330
RC   AAGCGCAAGCTGCAAGCGGGCGACAAGATGGCCGGTCGTCACGGCAACAAGGGTGTGGTC
     :: :::    : ::  : :: ::::::::::: :::::::::: :::::::::::  :
EC   AAACGCCGTATCCAGCCTGGTGACAAGATGGCAGGTCGTCACGGTAACAAGGGTGTAATT
           1210      1220      1230      1240      1250      1260

      3340      3350      3360      3370      3380      3390
RC   TCGAAAGTC-GTTCGATCGAAGACATGCCCTTCCTGGCCGATGGCACCCCGGTCGACCTC
     :: :: ::    X:::::::: ::::: : :   :   : :: :: ::::: ::: ::
EC   TCTAAGATCAACCCGATCGAAGATATGCCTTACGATGAAAACGGTACGCCGGTAGACATC
           1270      1280      1290      1300      1310      1320

      3400      3410      3420      3430      3440      3450
RC   GTTCTGAACCCGCTCGGCGTGCCGTCGCGGATGAACGTCGGGCAGATCCTGGAAACCCAC
     :: ::::::::::: ::::: ::::: :: :::::: :::: :::::::: :::::::::
EC   GTACTGAACCCGCTGGGCGTACCGTCTCGTATGAACATCGGTCAGATCCTCGAAACCCAC
           1330      1340      1350      1360      1370      1380
```

```
        3460      3470      3480      3490      3500          3510
RC   ATGGGCTGGGCCAGCCGTGGTCTCGGTATCCGGATCGACGAGGCGCTGAA--AGC--CTA
      ::::   :::        ::: ::::   :  :::: :::    ::::::  :::  : :
EC   CTGGGTATGGCTGCGAAAGGTATCGGCGACAAGATCAACGCCATGCTGAAACAGCAGCAA
          1390      1400      1410      1420      1430      1440

                3520      3530      3540      3550      3560
RC   ----TCGTCCGACCG-GCGACCTGACCCCG-GTGCGCGAGGAAATGCGCCACGGCTATGG
         :::    :: : ::::  : : :: : :::::  : ::  :: ::   : :  : :
EC   GAAGTCGCGAAACTGCGCGAATTCATCCAGCGTGCG-TACGATCTGGGCGCTGACGTTCG
          1450      1460      1470      1480      1490      1500

        3570          3580      3590      3600      3610      3620
RC   CGACGA----GACCTACGAGGACGTCTTCGCCGAGATGGAGAACGAGCGGTTCATCGAAT
       :  :    :::::  ::: :.  :::  : : :::: ::::      :::   ::  :
EC   TCAGAAAGTTGACCT--GAGTAC--CTT--CAGCGATGAAGAA------GTTATGCGTCT
            1510      1520          1530            1540

            3630      3640      3650      3660      3670      3680
RC   GCGCCGAACATGTGCGCGGCGGTGTGCCGATCGCCACCCCGGTCTTTGACGGGGCCAAGG
     : :: ::: :  :::::   ::: :::: ::::: :: ::::: :: ::::: :: :: :
EC   G-GCTGAAAACCTGCGCAAAGGTATGCCAATCGCAACGCCGGTGTTCGACGGTGCGAAAG
    1550      1560      1570      1580      1590      1600

            3690      3700      3710      3720      3730      3740
RC   AAGCCGATGTGAACGATGCGCTGATCCGCGCGGGCTTCAACAAGTCGGGCCAATCGATCG
     :::: ::  : :: ::   :::::  :  :  : : :    :  :: :: ::       :
EC   AAGCAGAAATTAAAGAGCTGCTGAAACTTGGCGACCTGCCGACTTCCGGTCAGATCCGCC
    1610      1620      1630      1640      1650      1660

            3750      3760      3770      3780      3790      3800
RC   TCTTCGACGGCCGCACCGGCGATCAATTCGCCCGTCCGGTCACCGTGGGCGTGAAGTACT
     : : ::: :: ::::: :: :: :: ::::  :::::::: ::::: ::    : ::::
EC   TGTACGATGGTCGCACTGGTGAACAGTTCGAGCGTCCGGTAACCGTTGGTTACATGTACA
    1670      1680      1690      1700      1710      1720

            3810      3820      3830      3840      3850      3860
RC   TCCTCAAGCTGCACCACCTTGTCGATGACAAGATGCACGCCCGTTCGACCCCGCCGTATT
     : :: :: ::: ::::::: ::::: :::::::::::::: ::::: :::::  : ::
EC   TGCTGAAACTGAACCACCTGGTCGACGACAAGATGCACGCGCGTTCCACCGGTTCTTACA
    1730      1740      1750      1760      1770      1780

            3870      3880      3890      3900      3910
RC   CGCTCGTCAC-CAGCA-ACGCTGGGCGGCAAGGCGCAGTTCGGCGGTCAGCGTCTGGGGG
       :: :: :X :::::  ::::::: :: ::::: :::::::: ::::::::: : ::::
EC   GCCTGGTTACTCAGCAGCCGCTGGGTGGTAAGGCACAGTTCGGTGGTCAGCGTTTCGGGG
    1790      1800      1810      1820      1830      1840
```

```
   3920       3930       3940       3950       3960       3970
RC   AGATGGAGGTCTGGGCTCTGGAAGCTTACGGCGCCGCCTATACCCTGCAGGAGATGCTGA
     ::::::: :: ::::: :::::::: :::::::: :: :: :::::::::: ::::: :
EC   AGATGGAAGTGTGGGCGCTGGAAGCATACGGCGCAGCATACACCCTGCAGGAAATGCTCA
    1850       1860       1870       1880       1890       1900

   3980       3990       4000       4010       4020       4030
RC   CGGTGAAGTCGGACGACGTCGCGGGCCGGACCAAGGTCTACGAGAGCATCGTCAAGGGCG
     : :: ::::: :: :::::    :: :: :::::: : ::  : : ::::::  : :::
EC   CCGTTAAGTCTGATGACGTGAACGGTCGTACCAAGATGTATAAAAACATCGTGGACGGCA
    1910       1920       1930       1940       1950       1960

   4040       4050       4060        4070       4080       4090
RC   AGGACAACTTCGAGGCCGGCGTGC--GAATCGTTCAACGTGCTCGTCAAGGAAGTGCGGG
     :  :  :  : ::: : ::: :::  ::::: ::::::::  :  : :: ::  : ::
EC   ACCATCAGATGGAGCCGGGCATGCCAGAATCCTTCAACGTATTGTTGAAAGAGATTCGTT
    1970       1980       1990       2000       2010       2020

      4100       4110       4120       4130       4140       4150
RC   C-CTCCGGCTCAACATGGAACTCCTGGATTCGGAGGAAGAGTGAGCGCGAAACTGCACAC
     : ::  :  ::::::: :::::            ::::::  :::
EC   CGCTGGGTATCAACATCGAACT------------GGAAGA-CGAG
    2030       2040                   2050

RC   CCCCCTCTGACCTCACGCAAGGAAGATCTGATGAACCAGGAAATCACCAACAACAACCCG
     TTCAACCCGCTGGCGCACGAAGGCCTTTGACGAGATCAAGATCTCGCTTGCGAGCCCTGA
     ACGGATCC *4284
```

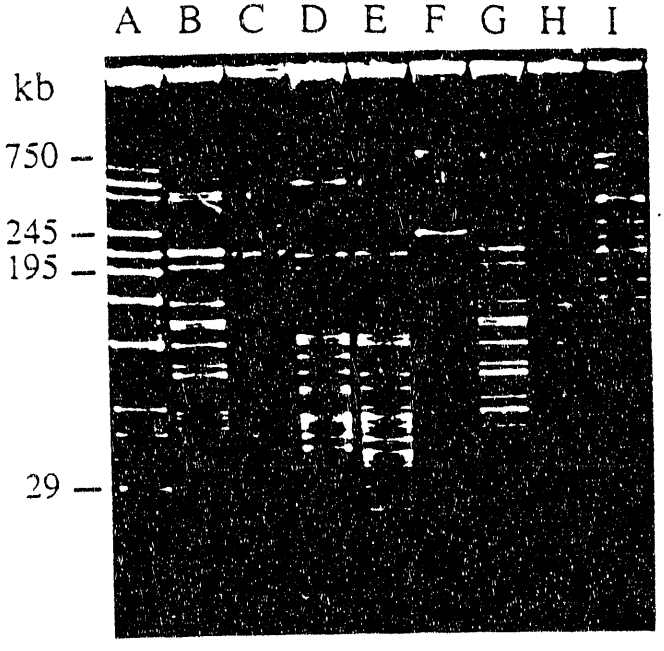
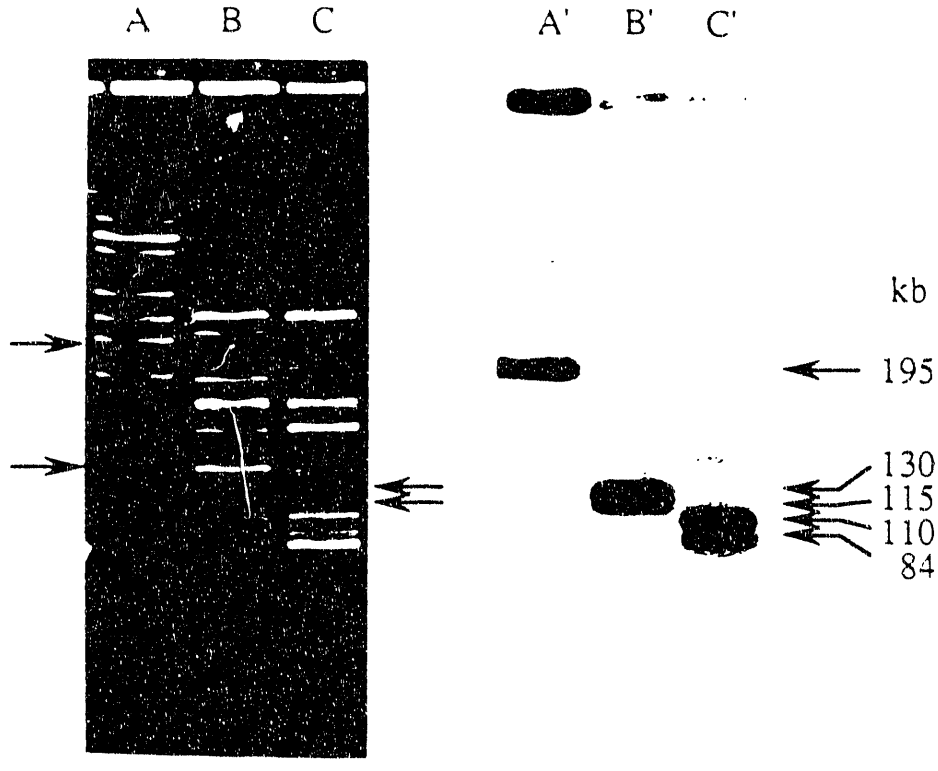


Figure 3. Pulsed field electrophoresis of restriction fragments of chromosomal DNA from Rhodobacter capsulatus SB1003. Conditions of electrophoresis: 1% agarose, switching time 5-25 sec, 10 V/cm in the BioRad CHEF apparatus, run for 20 hours. A=XbaI, B=AseI, C=Xba/Ase, D=DraI, E=Dra/Ase, F=SpeI, G=Spe/Ase, H=Xba/Spe, and I=ScaI. Sizes in kb at the left are based on standards of lambda DNA multimers, not shown.



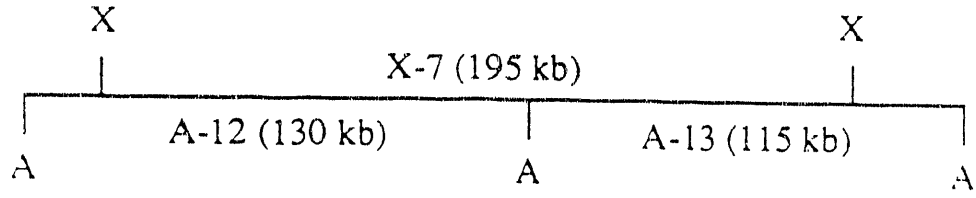


Figure 4. Determining the physical map by hybridization with eluted bands of DNA from the pulsed field gels. Lanes A,B, and C contain DNA digested with XbaI, AseI, and the double digest, respectively. Band 7 (195 kb) was eluted from a gel like that shown in A, labeled and used to probe the blot A'B'C'. The bands identified are, in lane A', Xba-7 itself, in lane B', Ase-12 and-13 (the lower single arrow on the left) and in lane C' the two bands indicated by the two lowest arrows in both parts of the figure. A physical map consistent with these results is shown below.

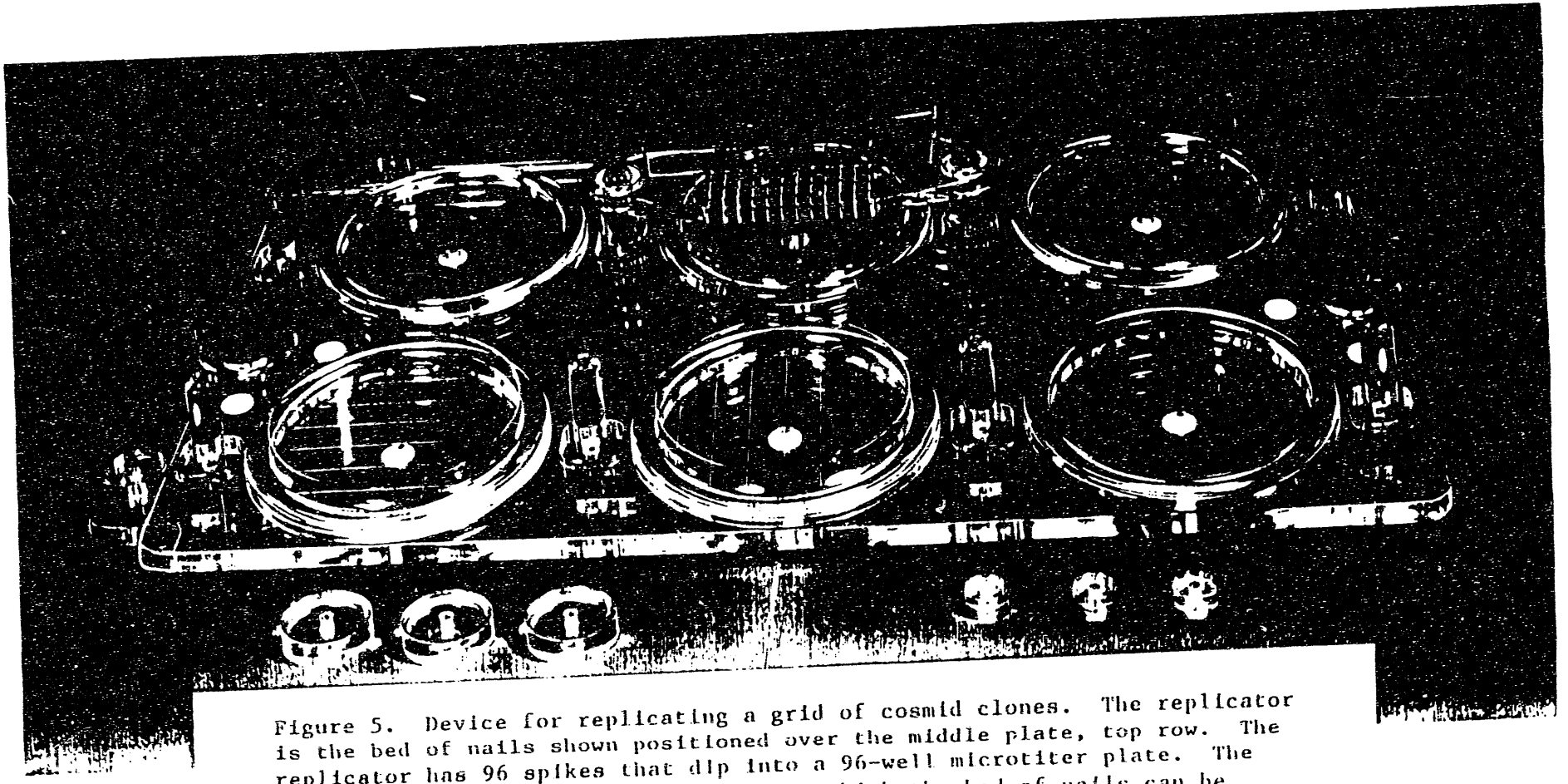


Figure 5. Device for replicating a grid of cosmid clones. The replicator is the bed of nails shown positioned over the middle plate, top row. The replicator has 96 spikes that dip into a 96-well microtiter plate. The lower plastic plate contains poles over which the bed of nails can be slipped reproducibly. These poles are fixed in position. The upper sheet of plastic, which holds the petri dishes, is precisely located with respect to the lower sheet by the differently sized spacer discs shown below. There are three east/west positions and three north/south positions, making nine possibilities. Thus the bed of nails can be used to transfer nine different 96-well sets of cosmids in a reproducible pattern, for a total of 864 cosmids per plate. The cosmid colonies are actually grown on nylon filters on the plates, then lysed and the DNA baked on. These filters can then be probed with gel-purified fragments or cloned genes.

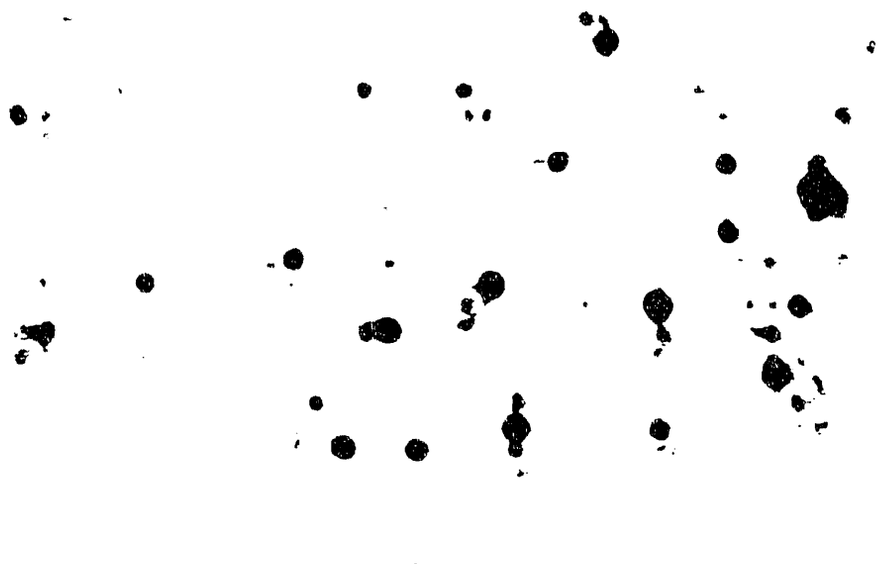


Figure 6. A typical result obtained using the device shown in Figure 5. The probe was Xba-10 (about 60 kb). Approximately 20 cosmids react strongly, another 17 weakly. Since this fragment represents 1-1.5% of the chromosome, we would expect to find 8-12 cosmids hybridizing strongly if the library is completely representative. The excess observed may mean that there are repeated sequences in the probe. Alternatively, Xba-10 may be overrepresented in the library.

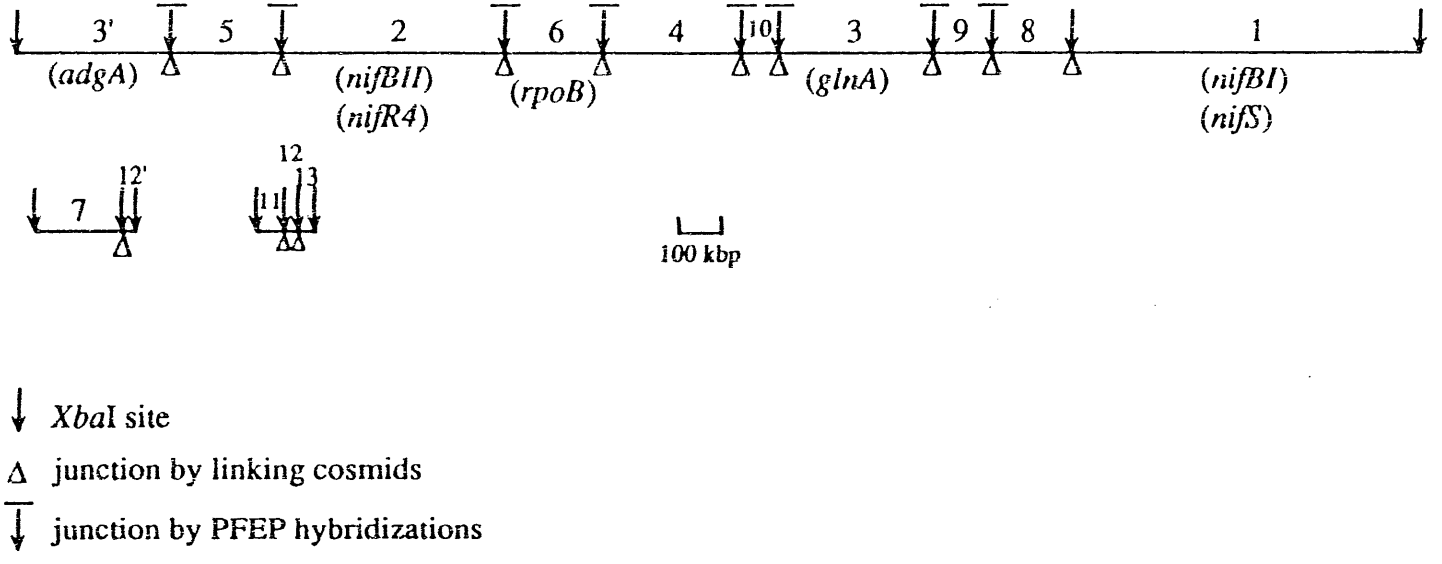


Figure 7. Physical map of the chromosome of Rhodobacter capsulatus SB 1003. At present there is one large linkage group and two smaller ones, which could be megaplasmids or could be joined to the large one. Note that the larger map linkage is determined both by the AseI fragment overlaps and by linking cosmids. Two of the XbaI fragments (3 and 12) are doublets. The genes assigned by hybridization to blots of the PFE gels are adgA (ammonia-dependent growth), nifr4 and nifBII, which are linked to the nifHDK operon, rpoB (RNA polymerase), glnA (glutamine synthetase, linked to glnB), and nifS (linked to most of the other nif structural genes) and nifBI, linked to nifAI.

END

DATE FILMED

9/14/92