The nucleotide sequences of the heterologous region between the genomes of *Bacillus* phages M2 and Nf that indicate the two phages are originally identical

Yukiko Mizukami, Takao Sekiya* and Hideo Hirokawa+

Life Science Institute, Sophia University, 7–1 Kioi-cho, Chiyoda-ku, Tokyo 102 and **National Cancer Center, Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104, Japan*

Received 27 November 1985; revised version received 7 January 1986

It has been suggested that the heterologous population of *Bacillus* phage M2 is derived from an original clone, which is identical with phage Nf, by the deletion on a particular region of the genome. We have determined the nucleotide sequence of this region of M2 subclones and Nf genomes. The results clearly indicate that the homologous recombination through the short direct repeated sequence is the main cause of the varied deletions found in the genomes of M2 subclones.

(Bacillus *phage*) *Nucleotide sequence* *DNA rearrangement* *Homologous recombination*

1. INTRODUCTION

There have been precedent cases where strains of an identical phage isolated independently in different laboratories contained minor rearrangements of the nucleotide sequence in their genomes [1]. The case of *Bacillus* phages M2 [2] and Nf [3] has been suggested to be similar to those precedent cases. Both phages, like phage ϕ29, have a unique genome structure which is a linear double-stranded DNA with the terminal proteins covalently linked to the 5′-ends [4,5]. In considering our previous results on the formation of heteroduplex molecules between the genomes of M2 and Nf and the comparison between the physical maps of the two phages [6], it has been shown first that the genome of M2 original stock has heterogeneity of length in a particular region, and secondly that each genome of the subclones of M2 original stock is the same as that of Nf except in the region described above. Furthermore, the genome of subclone H6, which has the largest molecular size among the genomes of M2 subclones, is identical with that of Nf as shown in the results. Thus, it was suggested that the heterologous population of M2 is derived from an original clone, which is identical to phage Nf, by the deletion on a particular region of the genome. To investigate further the relationship between M2 and Nf, we have determined the complete nucleotide sequences of the heterologous region between M2 and Nf genomes and revealed the deletion end-points and their flanking sequences in the genomes of M2 subclones.

2. MATERIALS AND METHODS

Phage Nf was obtained from Dr Takagi. The subclones of phage M2, H1, H2, H3, H5 and H6 are the single plaque isolates of our laboratory [6]. The phage DNA and restriction fragments were prepared according to [7]. The nucleotide sequence was determined by the Maxam-Gilbert chemical method [8].

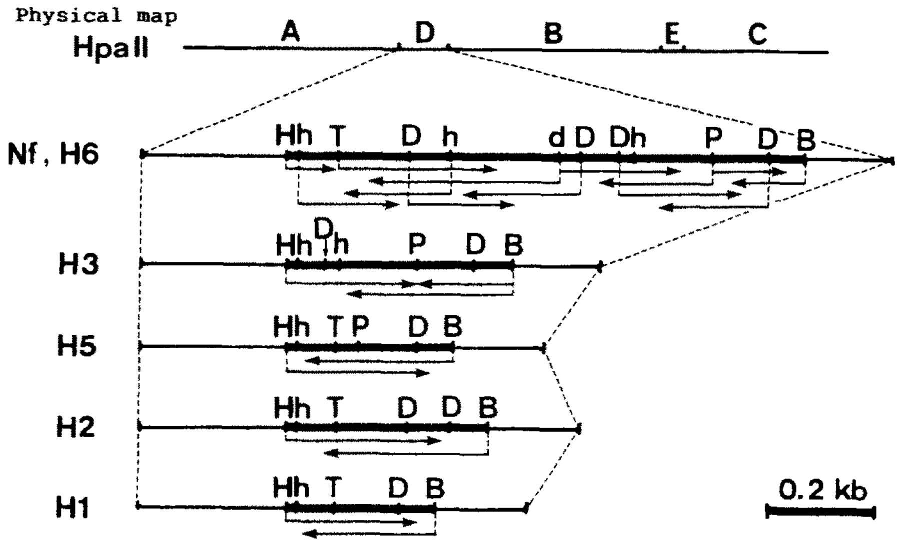
3. RESULTS AND DISCUSSION

As described previously, the heterologous region among the genomes of M2 subclones and Nf has

+ To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

a.



b.

```
HinfI       ******20           ORF       40                  H360                 80
5'-GATTCATCTTCTGGAGGTGCGCAAGCGTAATGTATTCATTCACTGCATATGCAAACTCTGACATTGTCGCCTTTCATTTA
                                M  Y  F  T  A  Y  A  N  S  D  I  V  A  F  H  L  L
                 100                     120      H1  H5          140                 160
TTGAAATTTCTATCGAGCGAAAACAACATAGAAATTTCTTATGCAGATGAAGACACAATACCAGAATACGTTTCAATTCG
L  K  F  L  S  S  E  N  N  I  E  I  S  Y  A  D  E  D  T  I  P  E  Y  V  S  I  R
                 180                     200   H2                 220                 240
TGACCTAAAAGCTGGAGATAAAACAACAATTGATTTATACCCATTAGTTGCGTGGAAAGTGATTGCTCAGGAGGACATAA
 D  L  K  A  G  D  K  T  T  I  D  L  Y  P  L  V  A  W  K  V  I  A  Q  E  D  I  T
                 260                     280                      300                 320
CAACTGGGGACAGAGTATCCGTTGGAAAGAATGGGCAAGTAAAAAAGACAACTGACTTGCGCACAGCATTCGGGTACGCT
  T  G  D  R  V  S  V  G  K  N  G  Q  V  K  K  T  T  D  L  R  T  A  F  G  Y  A
                 340                     360                      380                 400
GTTAGCCTTGCAAAAGCTGGACAACTTGTTACAGTCGCTATTTCAACGGTATTTGACACGATTATCACGCCAGATGATTT
V  S  L  A  K  A  G  Q  L  V  T  V  A  I  S  T  V  F  D  T  I  I  T  P  D  D  L
                 420                     440                      460                 480
AAGGGAATGTTGACGATGATGTAAAAGCATTCTTGAAGTCAAATGCAACAGACGCAAATAAAGCAAAGTTGCGTGATCTA
 R  E  C
                 500                     520                      540                 560
CTCGTATCCAATCTTGATGTGAAAGCTTTTCTCAATGGTTCAACAAGTGAAGACAATAAGATAAATCTTAGAAATTTACT

                 580                     600H3                    620                 640
TGTTAGCAACCCTGCAATCCTAGCGTTTCTTAATGCAAACCCTGACACCGATACTCAGACGACATTAAGAACAATGATTG

                 660                     680                      700                 720
GCGCTGGTACACCATATACACTACCAGCCGCCACTACCACAACACTAGGTGGTGTGAAACGAATACCAGCATTTGGAAAC

                 740                     760                      780 H5     H2       800
TCAACTGCTACTGATGTTGCAACACTGGTAAAAGATTTCAACAACCTACTTGCGGCGATGAGAACAGCTGGTTACATTCT

      H1         820                     840                      860                 880
TATAGAGGTGAAATTTCATGGCATATGTGCCTTTATCTGGTTCGGACGTTCGGATTTTTCGGCGTTGCGTTCTCCAATG
                                                                                BclI
                 900                     920                      940                 960
ATTATAAAAATACTAAGGTGGTTCACAAGTGCGGACGCTCAATATTCGTATTTTAACGCTAAACCACGTGTGCATGTGAT

CA-3'
```

Fig.1. (a) DNA sequencing strategy of the HB region with the labeled restriction enzyme cleavage sites. Horizontal arrows indicate the direction and the extent of sequence determination. B, *Bcl*I; D, *Dde*I; d, *Hin*dIII; H, *Hin*fI; h, *Hha*I; P, *Pvu*II; T, *Taq*I. (b) Complete nucleotide sequence of the HB region of the Nf genome. The corresponding sequences of the short direct repeats found on the deletion end-points of M2 subclones H1, H3 and H5 are indicated in boxes. The deduced amino acid sequence is presented under the open reading frame. * Potential ribosome binding sequence.

been mapped within the region between the cleavage sites of *Hin*fI and *Bcl*I on the *Hpa*II-D fragment [6]. Therefore, the nucleotide sequence of this region (the HB region) was determined. The sequencing strategy and the resulting nucleotide sequences are shown in fig.1.

By comparing the nucleotide sequences among the DNAs of M2 subclones and Nf, it was seen that the nucleotide sequence of the HB region of subclone H6 DNA was identical with that of Nf DNA. The DNA of the other subclones of M2, H1, H2, H3 and H5 had deletions of 681, 585, 544 and 650 bp, respectively. Moreover, these deletions had different end-points from each other. The flanking sequence of the novel joint in the HB region of each subclone DNA was the same as the corresponding sequence of Nf DNA. These results clearly indicate that the various deletions have actually occurred in the genome of M2 original clone which is identical with Nf. In other words, the two phages are originally identical. This result is supported by the previous study that has shown the identity of the terminal repeated sequences of the genomes between both phages, M2Y (one of the M2 subclones isolated by Yoshikawa) and Nf [9].

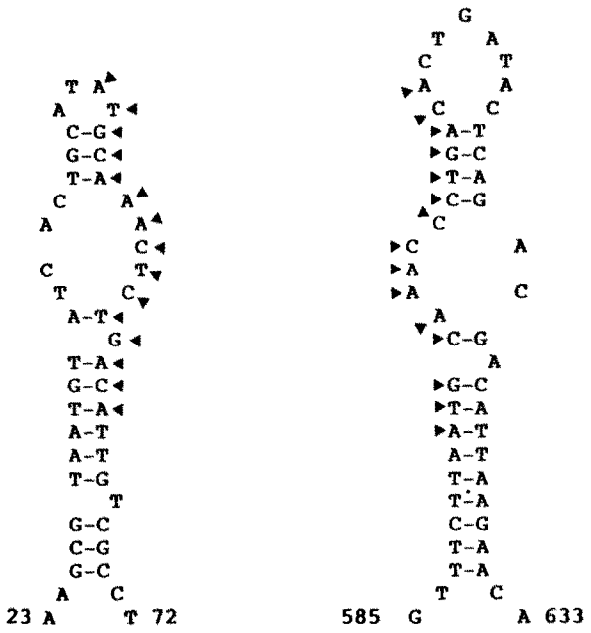


Fig.3. Potential hairpin structures containing the short direct repeated sequence which exists at the corresponding region of the deletion end-points of the H3 genome in the Nf genome. The solid triangles indicate the short homologous sequence.

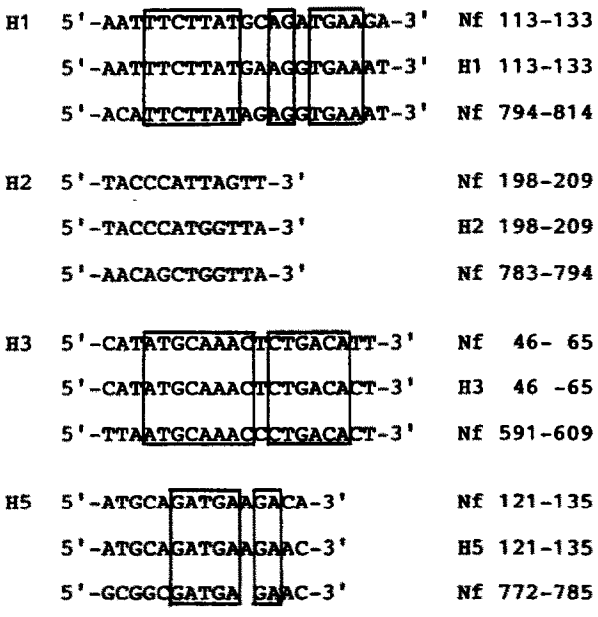


Fig.2. Nucleotide sequences of the novel joints of M2 subclone DNAs. The short homologous sequences are indicated in boxes.

At the nucleotide sequence of Nf DNA corresponding to the deletion end-points of the genomes of M2 subclones, but not of subclone H2, there exist distinct short direct repeated sequences of 8–16 nucleotides long (fig.2). This suggests that the homologous recombinations through the repeated sequences are mainly the cause of the varied deletions found in the genomes of M2 subclones. Interestingly, in the Nf genome, we found a potential hairpin structure containing the short direct repeated sequence at every corresponding region of the deletion end-point of M2 genome. Fig.3 shows the possible hairpin structures at the corresponding region of the deletion end-points of subclone H3 genome. These structures may possibly function in recognizing the homology between the two short repeats and/or in forming the mismatch pairing between them [10].

In addition, an open reading frame (ORF) of 381 nucleotides long was found in the HB region of the Nf genome where the nucleotide sequences were deleted in the genomes of M2 subclones, except in subclone H6 (fig.1b). This ORF has a

reasonable *Bacillus* ribosome-binding site (ΔG = −14.7 kcal) preceding the initiation codon ATG [11], and is able to code for a protein of 14.7 kDa. Our previous electron-microscopic observation has shown that the deletions affect the phage morphology, namely, the presence of head projections in the phage head [6]. Furthermore, the molecular size of the projection was indicated to be 24 kDa by electrophoretic analysis, and the 24 kDa protein was synthesized in minicells which were infected by M2 subclones as well as by Nf (not shown). Therefore, it is implied that the product of the above ORF is not the protein of the projection itself, but a factor which probably participates in the assembly of the head projections.

ACKNOWLEDGEMENTS

The support of Drs H. Saito and S. Nishimura in this investigation is gratefully acknowledged. We are grateful to Dr M. Hasegawa for his help in computer analysis of the sequence data. We also thank Dr T. Mizukami for helpful discussions and critical reading of this manuscript.

REFERENCES

[1] Studier, F.W. (1979) Virology 95, 70–84.
[2] Ikeda, Y., Saito, H., Miura, K., Takagi, J. and Aoki, H. (1965) J. Gen. Appl. Microbiol. 11, 181–190.
[3] Shimizu, N., Miura, K. and Aoki, H. (1970) J. Biochem. 68, 277–286.
[4] Hirokawa, H. (1972) Proc. Natl. Acad. Sci. USA 69, 1555–1559.
[5] Takagi, J., Hirokawa, H. and Saito, H. (1978) J. Gen. Appl. Microbiol. 24, 135–137.
[6] Mizukami, Y. and Hirokawa, H. (1985) J. Gen. Appl. Microbiol. 31, in press.
[7] Kawamura, F. and Ito, J. (1977) Virology 83, 233–245.
[8] Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[9] Yoshikawa, H. and Ito, J. (1981) Proc. Natl. Acad. Sci. USA 78, 2596–2600.
[10] Albertini, A.M., Hofer, M., Calos, M.P. and Miller, J.F. (1982) Cell 29, 319–328.
[11] Moran, C.P. jr, Lang, N., LeGrice, S.F.J., Lee, G., Stephens, M., Sonenshein, A.L., Pero, J. and Losick, R. (1982) Mol. Gen. Genet. 186, 339–346.