Spontaneous recombination between homologous prophage regions causes large-scale inversions within the Escherichia coli O157:H7 chromosome

Iguchi, Atsushi
Iyoda, Sunao
Terajima, Jun
Watanabe, Haruo
Osawa, Ro

(Citation)
Gene,372:199-207

(Issue Date)
2006-05

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90000322

Spontaneous recombination between homologous prophage regions causes large-scale inversions within the *Escherichia coli* O157:H7 chromosome

Atsushi Iguchi[1], Sunao Iyoda[2]. Jun Terajima[2], Haruo Watanabe[2], and Ro Osawa[1*]

Department of Bioscience, Graduate School of Science and Technology, Kobe University, Kobe, Hyogo 657-8501,[1] and Department of Bacteriology, National Institute of Infectious Diseases, Shinjuku-ku, Tokyo, 162-8640,[2] Japan

Abbreviations: PFGE, pulsed-field gel electrophoresis; Stx, Shiga toxin; RFLPs, restriction fragment length polymorphisms; HIA, heart infusion agar; Dam, deoxyadenosine methyltransferase; OI, O-islands

・ Corresponding author. +81 78 803 5804: fax: +81 78 803 5804.
E-mail address: osawa@ans.kobe-u.ac.jp (Ro Osawa)

[1]Present address: Department of Bioscience, Graduate School of Science and Technology, Kobe University, 1-1 Rokko-dai, Nada-ku, Kobe, Hyogo, 657-8501, Japan.

Abstract

It is known that *Xba*I-digested chromosomal DNAs of strains of *Escherichia coli* O157:H7 exhibit a wide variety in pulsed pulsed-field gel electrophoresis (PFGE) fragment patterns, which is used for epidemiological surveillance of this important pathogen. The variety in the restriction enzyme-digestion patterns suggests a wide genomic diversity, however, only a few studies have been conducted to investigate involvement of large-scale chromosomal rearrangements in development of the diversity. In this study, through rounds of subculturing *E. coli* O157:H7 strain EDL933, naturally occurring genome variation in the isolated derivatives was investigated. By comparing the PFGE patterns among clonally related derivatives, we found five types of large-scale inversions taking place within the chromosome. The five inversions found were across the replication axis and ranged from 250-kb to 1.4-Mb long, and all the corresponding recombination sites were associated with prophages or phage-like regions. Four inversions out of the five were resulted from recombination between pairs of lambda-like prophages disturbing the symmetry of the origin and terminus of the replication axis. These observations indicate that those prophage regions represent some of the hot spots for intrachromosomal recombination within the *E. coli* O157:H7 chromosome, where recombination between the prophage regions results not only in the large chromosomal inversions but might also in generation of chimeric phages.

1. Introduction

Escherichia coli O157:H7 is an important food-borne pathogen that causes diarrhea, hemorrhagic colitis and hemolytic-uremic syndrome (Griffin, 1995, Kaper et al., 1998 and O'Brien et al., 1984). *E. coli* O157:H7 was first recognized as a pathogen in 1982 (Riely et al., 1983) and has been implicated in several outbreaks throughout the world (Kaper et al., 2004, Karch et al., 1999 and Watanabe et al., 1996.). In recent years, molecular based techniques, such as amplified polymorphic DNA analysis (Cebula et al., 1995 and Wang et al., 2002) and pulsed-field gel electrophoresis (PFGE) (Gouveia et al., 1998 and Izumiya et al, 1997; Johnson et al., 1995) have been shown to be useful methods for discriminating among isolates of pathogens in sporadic or multiple outbreaks. Among these techniques, PFGE is currently considered to be one of the most reliable typing procedures. This method is a well-established and highly effective epidemiological tool for the molecular analysis of large fragments generated by restriction-enzyme digestion of genomic DNA. *E. coli* O157:H7 isolates demonstrated varied PFGE patterns following *Xba*I digestion, which enable epidemiological surveillance of this important human pathogen. The variability of restriction enzyme-digestion patterns of *E. coli* O157:H7 genomes suggested that the presence of wide genomic diversity among the strains.

In previous studies, researchers attempted to decipher the genetic events underlying strain variation in *E. coli* O157:H7, as evidenced by the PFGE pattern. For instance, a minor variation in PFGE fragment patterns in clonally related *E. coli* O157:H7 strains has been reported (Akiba et al., 1999 and Murase et al., 1999). Murase et al. (1999) showed that PFGE fragment patterns of *E. coli* O157:H7 isolates were changed by the loss of single fragments during maintenance or subculturing, which was possibly due to the curing of Shiga toxin (Stx)-converting bacteriophages, while Akiba et al. (1999) reported that genetic differences detected by PFGE, accompanied by the loss of a large plasmid, have been observed between the inoculated and recovered isolates from experimentally infected cattle. However, other genetic events underlying the PFGE differences between clonally related strains have not been defined.

We recently reported that the changing of a few bands on the PFGE patterns

during cultivation of the *E. coli* O157:H7 strains, even between the same clones, has been observed (Iguchi et al., 2002). Genomic diversification occurs among cells of the same clone in culture and presumably also in nature. In order to characterize such spontaneous genomic rearrangements resulted in the difference in PFGE patterns, we initiated a comparative analysis of the genomes between two derivatives, E0T-1 and E50T-11, belonging to the same lineage of *E. coli* O157:H7 strains, EDL933 (Perna et al., 2001). In addition, we investigated a total of 17 clones isolated from a series of subcultures of EDL 933 that indicated different PFGE patterns.

In this study, we used combined PFGE and Southern hybridization to determine restriction fragment length polymorphisms (RFLPs) among the derivatives, to assign these RFLPs to particular regions of their genome, and to identify the genetic events that could be correlated with these genome rearrangements.

2. Materials and Methods

2.1. Strains

The strains used in this study are listed in Table 1. All clones isolated from repeated subculturing were derivatives of the *E. coli* O157:H7 strain, EDL933 (ATCC43895) (Perna et al., 2001), a human isolate from an outbreak in Michigan in 1982. Two strains, E0T-1 and E50T-11, were isolated in previous study (Iguchi et al., 2002). E0T-1, isolated from a storage culture of EDL933 maintained in our laboratory, was the initial strain used in the repeated subculturing test, and E50T-11, isolated after 50 subcultures, showed the greatest changes in the PFGE fragment patterns (Iguchi et al., 2002). All isolates were stored at – 80^{O} C in 10 % skim milk.

2. 2. Media and culture conditions

The strains were continually subcultured on a slant medium of heart infusion agar (HIA) (Nissui) at a rate of twice per week (approximately 3 days intervals) at 37 ^{O} C, and the subculturing was repeated 50 times (for 25 weeks). After every 10 subcultures, a loopful of cultures was streaked onto a separate HIA plate and

incubated at 37 ^{O} C for 18 h. After incubation, 18 well-isolated colonies were randomly selected and subjected to the PFGE analysis.

2. 3. PFGE methods

Genomic DNAs embedded in agarose gels were prepared as previously described (Davis et al., 2003) with some modifications. Briefly, plugs were prepared from bacterial suspensions (200 μl) to which proteinase K (20 μl at 10 mg/ml) was added. SeaKem gold agarose (FMC BioProducts) was added (1.0 % with 1 % sodium dodecyl sulfate in 200 μl of TE buffer [100 mM Tris and 100 mM EDTA]), and the mixture was immediately dispensed into plug molds. The plugs were lysed in ES buffer (0.5 M EDTA, pH 9.0, 1 % sodium dodecyl sulfate) with proteinase K at 55 ^{O} C for 1 h and then washed at 50 ^{O} C in TE buffer (four times for 20 min each time). Plug slices were washed with the appropriate restriction enzyme buffer and then incubated for 4 h at 37 ^{O} C with 30 U of *Xba*I (Roche Biochemicals). PFGE was performed with the Bio-Rad CHEF Mapper or Bio-Rad CHEF DRII electrophoresis system using the following parameters: separation on a 1 % agarose gel in 0.5x Tris-Borate-EDTA at 14^{O} C. A lambda ladder PFG marker (New England Biolabs) was used as a molecular size marker. The gels were run at a voltage gradient of 6.0 V/cm for 22 h at an initial switch time of 0.1 s and a final switch time of 38 s. After PFGE, the gels were stained with ethidium bromide and photographed with UV transillumination.

2. 4. Predicted restriction fragments from genome sequence

The sequenced EDL933 genome (GenBank accession no. ***AE005174***) was used to search for the *Xba*I recognition sites (TCTAGA). PFGE of the EDL933 strain following *Xba*I digestion was performed under the conditions described above.

2. 5. PCR

Probes for the Southern analysis (described below) were PCR fragments corresponding to part of the five prophages, CP-933M, CP-933O, CP-933R, CP-933P and CP-933U, and two phage-like elements, O islands (OI) 43 and 48

(Fig. 1), amplified from EDL933 DNA as a template using specific primer pairs (Table 2). For PCR, TaKaRa ExTaq polymerase (Takara Biomedical) was used according to the manufacturer's instructions in 100-μl reaction mixtures with 20 ng of genomic DNA as a template. A three-step PCR was used: 94 ^{O} C for 20 sec, 58 ^{O} C for 20 sec, and 72 ^{O} C for 30 sec (25 cycles). All the reactions were performed using a GeneAmp model 9700 thermal cycler (Perkin Elmer Applied Biosystems). Six additional primer pairs designed to examine the chromosomal rearrangements were also listed in Table 2. The conditions for PCR were the same as above.

2. 6. Southern hybridization

DNA fragments separated in PFGE gels were transferred to Hybond N+ nylon membranes (Amersham Biosciences) using the capillary method. Hybridization probes were synthesized with PCR, as described above. Probes targeted for the both ends of the five prophage elements, flanking regions of the *Xba*I site on the CP-933O and neighboring regions of the two phage-like elements, as shown in Fig. 1. Probe labeling, hybridization, washing and detection were performed with on AlkPhosDirect labeling kit and CDP-star detection substrate (Amersham Biosciences) according to the manufacturer's instructions. The signals were detected by exposure to X-ray film.

3. Results

3. 1. PFGE variation through repeated subculturing

An isolated derivative of EDL933, E50T-11, showed the PFGE pattern M. 50 additional rounds of subculturing E50T-11 were performed to investigate if any further changes in the PFGE patterns might be seen in its derivatives. The 18 randomly selected colonies did not show any variation in PFGE fragment pattern for the first 10 subcultures, but did start to show variation after 20 subcultures (Table 3). The fragment patterns shown by the strain at the 50th subculturing differed from the pattern M by at most six fragments (Table 3). A total of 11 patterns (M-1 to 11) were observed through the repeated subculturing. Meanwhile, the initial strain, E0T-1 showing pattern H, was continually

subcultured by the same way, and 6 patterns (H-1 to 6) were observed (Table 3). The fragment patterns shown by the strain through the repeated subculturing differed from the pattern H by at most four fragments (Table 3). We suspected that there might be genomic rearrangements among the EDL933 derivatives, which revealed multiple PFGE patterns through repeated subculturing.

3. 2. XbaI cleavage of EDL933 genomic DNA

The *E. coli* O157:H7 EDL933 genome has 41 *Xba*I restriction sites. As the enzyme deoxyadenosine methyltransferase (Dam) acts on the adenine residue of GATC quartets, 11 of 41 restriction sites would be methylated and undigested by *Xba*I. For example, the *Xba*I restriction site (positions 1,341,923 to 1,341,928) located on the lysogenic Stx2-converting bacteriophage BP-933W was methylated, so BP-933W region did not separate on the PFGE gels (data not shown). The genome sequence data indicated that 30 fragments would be expected from complete *Xba*I digestion of the EDL933 genome. Four of the predicted fragments, smaller than 10-kb, could not be detected on the gel using the standard electrophoresis conditions, and the remaining 26 fragments, which ranged from 39 to 612-kb long, appeared to be almost the same size in the actual PFGE fragment pattern of sequenced EDL933 strain.

3. 3. Characterization of genomic rearrangements between E0T-1 and E50T-11

The PFGE patterns of E0T-1 (pattern H) and E50T-11 (pattern M) were compared. In the pattern M, three fragments, (612, 43 and 39-kb), that had been seen in the pattern H disappeared, and two of new fragments (approximately 410 and 240-kb) were appeared, as shown in Fig. 2. To trace the destination of these fragments, we performed Southern hybridization analysis by using the specific probes targeted at the both ends of the prophages and phage-like regions (Fig. 1) considered as high-plasticity zones in *E. coli* O157:H7 genomes (Ohnishi et al., 2002 and Wick et al., 2005).

First, we traced the largest 612-kb fragment seen in the pattern H. The 612-kb fragment contained some prophages and phage-like regions, such as CP-933M, BP-933W, OI 43 and OI 48 (Fig. 1). The subsequent Southern

hybridization analysis revealed that both the 933MA and 933MB-targeted probes hybridized to the 612-kb fragment, while in pattern M, these probes hybridized to different fragments, 240 and 410-kb fragments respectively (Fig. 2). We then analyzed the 39-kb fragment seen in the pattern H. The CP-933U region has two *Xba*I sites, which yielded the 39-kb fragment (Fig. 1). The Southern hybridization analysis revealed that 933UA and 933UB-targeted probes hybridized to the 410 and 240-kb fragments respectively in the pattern M (Fig. 2). These results indicated that the new fragments observed in pattern M were constructed from the 612 and 39-kb fragments by a reciprocal crossover between CP-933M and CP-933U. Our analysis confirmed that the inversion resulting from recombination between CP-933M and CP-933U was occurred as shown in Fig. 3A. This type of inversion was designated as type 1 inversion. A 1.4-Mb segment, which constituted about 25% of the chromosome, was inverted across the replication axis between the two lambda-like prophages.

The 43-kb fragment, seen in the pattern H (Fig. 2) contained a part of the CP-933Y region (Fig. 3B). The position of the 43-kb fragment was out of the inverted segment described above, at position 3.5-Mb. To trace the 43-kb fragment, we performed PCR, Southern hybridization and sequencing analysis. Primer pairs targeted at the both ends of CP-933Y region yielded the expected PCR products, when E0T-1 was used as a template (date not shown). While, the primer pair targeted at the 933YA yielded the expected PCR product, but no product was amplified with 933YB-targeted primer pair, when E50T-11 was used as a template (date not shown). We designed additional primer pairs targeted 216U, -V, -W, and –X regions at 5 to 15-kb intervals (Fig. 3B), and tested isolates E0T-1 and E50T-11 with them. Primer pairs targeted 216V, -W, and –X regions did not yield amplicons from E50T-11, and primer pair 216U-F and 933YA-R yielded an approximately 8-kb amplicon. This indicated possible deletion of sequences at the start of CP-933Y region on the E50T-11 genome. We therefore characterized this deletion by sequencing the region encompassing the junction. As a result, this deletion spanned 52-kb and the left boundary was located exactly at the right terminus of the IS629 on the CP-933Y region. The deletion region included two *Xba*I sites, so the neighboring three fragments, 43, 5 and

216-kb, containing the deletion region were changed into an approximately 212-kb fragment on the pattern M (Fig. 3B). The 216-kb fragment seen in the pattern H and the bonded fragment in pattern M were similar in size, so our PFGE methods could not detect the change of a few kilobases on the PFGE fragment patterns (Fig 2).

3. 4. Inversions attributable to prophages and phage-like regions

To determine whether inversions occurred in other derivatives, additional 17 derivatives were isolated from a series of subcultures of EDL933, and probed for the prophages and phage-like regions. The PFGE and Southern hybridization analyses revealed that 3 types of inversion took place via recombination between pairs of prophages; CP-933M and CP-933U (type 1 inversion) (Fig. 4A), CP-933O and CP-933R (type 2 inversion) (Fig. 4B) and CP-933O and CP-933P (type 3 inversion) (Fig. 4C). All of these prophages were lambda-like phages, and the pairs of prophages were integrated in sites equidistant from the terminus of replication (*ter*) region (Fig. 5). The occurrence of type 1 inversion was also observed in an isolate (E20T-9; pattern H3) that was obtained from a separate set of subculturing experiment.

Meanwhile, the additional subculturing of E50T-11 resulted in the reversible rearrangement of type 1 inversion as seen M20T-17 (pattern M6)(Fig. 4E). Furthermore, we found a type of inversion between a pair of phage-like regions OI 43 and 48 (type 4 inversion) (Fig. 4D). These are identical islands encoding tellurite resistance, integrase, urease, phage proteins and an adhesin (Perna et al., 2001). Although the two islands were located on the same replichore of the chromosome and were inserted in the same orientation in the EDL933 genome, the type 1 inversion brought OI 48 to the opposite replichore as shown in Fig. 3A, and the additional subculturing of E50T-11 resulted in the further inversion between two islands across the replication axis. The occurrence of type 4 inversion was also observed in two derivatives showing pattern M10 and M11 (date not shown). In addition, we found that there was an inversion in E0T-1, compared with the sequenced EDL933 strain ATCC700927, between CP-933O and CP-933R (type 5 inversion), as shown in Fig. 4F. In this study, we confirmed

5 types of inversions resulted from recombination between pairs of prophages or phage-like regions, which ranged from 250-kb to 1.4-Mb long (Fig. 5).

4. DISCUSSION

Previous studies demonstrated some characteristics underlying the differences in genetic structure among the sequenced *E. coli* O157:H7 strains, EDL933 and Sakai, and other strains of this pathogen (Kudva et al., 2002, Ohnishi et al., 2002 and Wick et al., 2005). Kudva et al. (2002) reported that the genomic differences between *E. coli* O157:H7 strains are caused by discrete insertions or deletions of segments of DNA in the genome rather than point mutations in the *Xba*I sites. They also reported that the majority of genomic differences among *E. coli* O157:H7 strains occurred in O island sequences, suggesting that phage-associated events were associated with the diversity. Ohnishi et al. (2002) and Wick et al. (2005) analyzed the genomic structure of *E. coli* O157:H7 strains by whole-genome PCR scanning or DNA array and revealed that a high level variation occurred in the prophages among the *E. coli* O157:H7 chromosome, while the chromosomal DNA was conserved highly between the strains. These results confirmed that prophages are the most dynamic genetic elements in this *E. coli* O157:H7 lineage. Nevertheless, these studies were performed on genetically unrelated strains, and little has been described for detailed rearrangements in generating genomic diversity among the *E. coli* O157:H7 lineage.

In this context, we initiated rounds of subculturing strain EDL933 to obtain its clonal derivatives, and spontaneous mutants exhibiting different PFGE patterns after *Xba*I digestion were isolated. Subsequent comparisons of the genomes of derivatives provided useful data about the structural rearrangement, in which the intensively repeated subculturing allowed the large-scale inversions of *E. coli* O157:H7 chromosome, resulting in a remarkable diversification of PFGE patterns. All the inversion sites detected in the present study were all associated with prophages and phage-like regions. 18 regions of prophages or phage-like have been identified in EDL933 (Perna et al., 2001). Among them, 12 were lambda-like phages showing high similarities to each other. All these lambda-like

phages contain various types of deletions and/or insertions of IS elements, indicating that most of the prophages are defective. An Stx2-converting prophage, BP-933W, was only found to produce infectious phage particles, while the other phage fragments are mainly of incomplete or cryptic prophages (Plunkett et al., 1999). All lambda-like prophages of EDL933 showed the same orientation with respect to direction of chromosomal replication (Perna et al., 2001). Furthermore, 10 lambda-like prophages located on the *ter* side of the chromosome (Perna et al., 2001). The polarity in both the integration-points and orientation of prophages might enable the multiple large-scale inversions across the replication axis in the *E. coli* O157:H7 genomes. As a result, the regions close to *ter*, including many prophages, were a patchwork of inversion-refractory and inversion-tolerant region, as shown in Fig. 6. Such genetic events will lead to the observed genomic diversity in the lineages of the *E. coli* O157:H7 strains.

Replication of the chromosome begins at the origin of DNA replication (*oriC*) and proceeds bi-directionally to the *ter* region, on the opposite side of the circular chromosome. Theoretically, *oriC* and *ter* should be exactly opposite, which leads to chromosome balance with replication in both directions reaching *ter* at about the same time (Hill and Harnish, 1981). According to the EDL933 genome sequences, the sizes of the replichores were 2,713 and 2,815-kb. The genome balance was calculated by dividing the size of each replichore and the off-balance value, which was one-half the difference between the replichore sizes, and was about 50-kb. The off-balance value of all derivatives with one or more inversions ranged from 33 to 83-kb, except for M20T-15 in which type 4 inversion had occurred. These results indicated that most inversions did not disrupt the genome balance. Eisen et al. (2000) reported that chromosomal inversions around the origin and termination of replication are usually symmetrical, thus retaining a geographical chromosome balance. However, a striking unbalance was detected in M20T-15 in which the inversion had occurred between OI 43 and 48 regions, and the off-balance value was 416-kb (approximately 7.5% of the whole genome). Sequence analysis of EDL933 has shown that two OIs, designated OI 43 and 48 of 87,547-bp compose 92 open reading frames, including integrase, phage, tellurite resistance, urease and

adhesin genes (Taylor et al., 2002). In EDL933, the two islands are identical, with not even a single base difference between them, and being about 0.5-Mb apart at positions 1.05 and 1.54-Mb, respectively (Fig. 5). The translocation of OI 48 to the opposite side of the replichore by the type 1 inversion would permit further rearrangement. From this evidence, the insertion of a large block of high similarity fragments might allow the persistent inversion, thereby causing the chromosome unbalance.

Over the past few years, the complete nucleotide sequences for many bacterial species have been published, and the occurrence of large DNA inversions has been documented for the same or closely related species of bacteria, such as *Legionella pneumophila* (Cazalet et al., 2004), *Yersinia pestis* (Deng et al., 2002), *Shigella flexneri* (Wei et al., 2003), *Xylella fastidiosa* (Van Sluys et al, 2003) and *Streptococcus pyogenes* (Nakagawa et al., 2003). Most of these inversions apparently resulted from recombination between pairs of various homologous regions, especially IS elements, rRNA operons and prophages. Similar to what we have found in this study, a case of apparently prophage-mediated multiple inversions is reported elsewhere (Van Sluys et al., 2003) for a Gram-negative plant pathogen *Xylella fastidiosa*. These observations indicate that prophage regions represent important plasticity regions in the chromosome where recombination between homologous prophages can occur and result not only in large chromosomal rearrangement, but also in the generation of chimeric phages.

In the present study, we have shown that prophage regions represent some of the hot spots causing large-scale inversions within the *E. coli* O157:H7 chromosome, and the recombination occurs in the pairwise fashion between responsible pairs of prophage regions. Nakagawa et al. (2003) and Deng et al. (2002) reported chromosomal inversions in the genomes of *Streptococcus pyogenes* and *Yersinia pestis,* respectively, claiming that the inversion occurred as a result of homologous recombination between high similarity fragments in the bacterial genome. The above prophage regions in question are thus likely to be similar.

Although we were not able to describe molecular mechanisms involved in the

genomic inversions, one possible mechanism is that RecA is involved in the recombination between high similarity fragments of prophages since it is a highly conserved bacterial protein that facilitates DNA rearrangement via homologous recombination and often mediates homologous recombination (Karlin et al., 1995). Alternatively, a phage-encoded recombinase (e.g., an integrase) may be involved mechanisms of recombination between prophage regions. A control experiment using mutants in the absence of these recombinase activities may account for the mechanisms of large chromosomal inversions.

Changes in PFGE patterns of chromosomal restriction digests have sometimes been assumed to be due to spontaneous genomic rearrangements or recombinations of mobile elements (Akiba et al., 1999 and Murase et al., 1999). Our results shown that prophage-associated inversion is one of the important mechanisms by which *E. coli* O157:H7 maintain their genomic plasticity or diversity, and a large-scale inversion changed the size of the restriction fragments, resulting in four band differences in the PFGE patterns. In epidemiological studies, PFGE is useful for comparing *E coli* O157:H7 isolates within an outbreak that spans a narrow temporal window but may not provide stable clonal profiles for a longer-term due to their rather dynamic chromosomal rearrangements.

In this study, the multiple spontaneous inversions on EDL933 chromosome were observed during in a series of repeated subculturing on the HIA culture, in which there could be no specific stress. If the *E. coli* O157:H7 strains are subcultured under different media or incubation conditions, the frequency of occurrences of such chromosomal rearrangements may vary according to conditions. Furthermore, physical (e.g., heat or osmotic shock, UV, antibiotics) or nutritional stress may influence with such rearrangement. To explore these possibilities, we are investigating in prospective studies.

Acknowledgements

This work was supported by the health science research grant, 2005, from the Ministry of Health, Labor and Welfare in Japan. We are grateful to K. Yoshida of the Department of Bio-functional Chemistry, Faculty of Agriculture, Kobe

University for his valuable comments on an earlier draft of this paper.

References

Akiba, M., Sameshima, T., Nakazawa, M., 1999. The shift of genetic subtypes of *Escherichia coli* O157:H7 isolates from cattle. Epidemiol. Infect. 122, 343-346.

Cazalet, C., Rusniok, C., Bruggemann, H., Zidane, N., Magnier, A., Ma, L., Tichit, M., Jarraud, S., Bouchier, C., Vandenesch, F., Kunst, F., Etienne, J., Glaser, P., Buchrieser. C., 2004. Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. Nat. Genet. 36, 1165-1173.

Cebula, T.A., Payne, W.L., Feng, P., 1995. Simultaneous identification of strains of *Escherichia coli* serotype O157:H7 and their Shiga-like toxin type by mismatch amplification mutation assay-multiplex PCR. J. Clin. Microbiol. 33, 248-250.

Davis, M.A., Hancock, D.D., Besser, T.E., Call, D.R., 2003. Evaluation of pulsed-field gel electrophoresis as a tool for determining the degree of genetic relatedness between strains of *Escherichia coli* O157:H7. J. Clin. Microbiol. 41, 1843-1849.

Deng, W., et al, 2002. Genome sequence of *Yersinia pestis* KIM. J. Bacteriol. 184, 4601-4611.

Eisen, J.A., Heidelberg, J.F., White, O., Salzberg, S., 2000 Evidence for symmetric chromosomal inversions around the replication origin in bacteria. *Genome Biol.* 1, pp. 0011.1–0011.9.

Gouveia, S., Proctor, M.E., Lee, M.S., Luchansky, J.B., Kaspar, C.W., 1998. Genomic comparisons and Shiga toxin production among *Escherichia coli* O157:H7 isolates from a day care center outbreak and sporadic cases in southeastern Wisconsin. J. Clin. Microbiol. 36, 727-733.

Griffin, P.M. 1995. *Escherichia coli* O157:H7 and other enterohemorrhagic *Escherichia coli*. In: Blaser, M.J., Smith, P.D., Ravdin, J.I., Greenberg, H.B., Guerrant, R.L. (Eds.), Infections of the Gastrointestinal Tract. Raven Press, New York, pp. 739-762.

Hill, C.W., Harnish, B.W., 1981. Inversions between ribosomal RNA genes of

Escherichia coli. Proc. Natl. Acad. Sci. USA 78, 769-7072.

Iguchi, A., Osawa, R., Kawano, J., Shimizu, A., Terajima, J., Watanabe, H, 2002. Effects of repeated subculturing and prolonged storage at room temperature of enterohemorrhagic *Escherichia coli* O157:H7 on pulsed-field gel electrophoresis profiles. J. Clin. Microbiol. 40, 3079-3081.

Izumiya, H., Terajima, J., Wada, A., Inagaki, Y., Ito, K., Tamura, K., Watanabe, H., 1997. Molecular typing of enterohemorrhagic *Escherichia coli* O157:H7 isolates in Japan by using pulsed-field gel electrophoresis. J. Clin. Microbiol. 35,1675-1680.

Johnson, J.M., Weagant, S.D., Jinneman, K.C., Bryant, J.L., 1995. Use of pulsed-field gel electrophoresis for epidemiological study of *Escherichia coli* O157:H7 during a food-borne outbreak. Appl. Environ. Microbiol. 61, 2806-2808.

Kaper, J.B., O'Brien, A.D., 1998. *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli* strains. ASM Press, Washington, D.C.

Kaper, J.B., Nataro, J.P., Mobley, H.L.T., 2004. Pathogenic *Escherichia coli*. Nat. Rev. Microbiol. 2, 123-140.

Karch, H., Bielaszewska, M., Bitzan, M., Schmidt, H., 1999. Epidemiology and diagnosis of Shiga toxin-producing *Escherichia coli* infections. Diagn. Microbiol. Infect. Dis. 34, 229-243.

Karlin, S., Weinstock, G.M., Brendel, V., 1995. Bacterial classifications derived from RecA protein sequence comparisons, J. Bacteriol. 177**,** 6881-6893.

Kudva, I.T., Evans, P.S., Perna, N.T., Barrett, T.J., Ausubel, F.M., Blattner, F.R., Calderwood, S.B., 2002. Strains of *Escherichia coli* O157:H7 differ primarily by insertions or deletions, not single-nucleotide polymorphisms. J. Bacteriol.184, 1873-1879.

Murase, T., Yamai, S., Watanabe, H., 1999. Changes in pulsed-field gel electrophoresis patterns in clinical isolates of enterohemorrhagic *Escherichia coli* O157:H7 associated with loss of Shiga toxin genes. Curr. Microbiol. 38, 48-50.

Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y., Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T., Hayashi, H.,

Hattori, M., Hamada, S., 2003. Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. Genome Res. 13, 1042-1055.

O'Brien, A.D., Newland, J.W., Miller, S.F., Holmes, R.K., Smith, H.W., Formal, S.B., 1984. Shiga-like toxin-converting phages from Escherichia coli strains that cause hemorrhagic colitis or infantile diarrhea. Science 226, 694-696

Ohnishi, M., Terajima, J., Kurokawa, K., Nakayama, K., Murata, T. Tamura, K., Ogura, Y., Watanabe, H., Hayashi. T., 2002. Genomic diversity of enterohemorrhagic Es*cherichia coli* O157 revealed by whole genome PCR scanning. Proc. Natl. Acad. Sci. USA 99, 17043-17048.

Perna, N.T., et al., 2001. Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7. Nature 409, 529-533.

Plunkett, G., III, Rose, D.J., Durfee, T.J., Blattner, F.R., 1999. Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. J. Bacteriol. 181, 1767-1778.

Preston, M.A., Johnson, W., Khakhria, R., Borczyk, A., 2000. Epidemiologic subtyping of *Escherichia coli* O157 strains isolated in Ontario by phage typing and pulsed-field gel electrophoresis. J. Clin. Microbiol. 38, 2366-2368.

Riley, L.W., Remis, R.S., Helgerson, S.D., McGee, H.B., Wells, J.G., Davis, B.R., Hebert, R.J., Olcott, E.S., Johnson, L.M., Hargrett, N.T., Blake, P.A., Cohen, M.L., 1983. Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. N. Engl. J. Med. 308, 681-685.

Sonntag, A.K., Prager, R., Bielaszewska, M., Zhang, W., Fruth, A., Tschape, H., Karch, H., 2004. Phenotypic and genotypic analyses of enterohemorrhagic *Escherichia coli* O145 strains from patients in Germany. J. Clin. Microbiol. 42, 954-962.

Taylor, D.E., Rooker, M., Keelan, M., Ng, L.K., Martin, I., Perna, N.T., Burland, N.T., Blattner, F.R., 2002. Genomic variability of O islands encoding tellurite resistance in enterohemorrhagic *Escherichia coli* O157:H7 isolates. J. Bacteriol. 184, 4690-4698.

Van Sluys, M.A., et al., 2003. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. J. Bacteriol. 185, 1018-1026.

Vautor, E., Abadie, G., Guibert, J.M., Huard, C., Pepin, M., 2003. Genotyping of *Staphylococcus aureus* isolated from various sites on farms with dairy sheep using pulsed-field gel electrophoresis. Vet. Microbiol. 96, 69-79.

Wang, G., Clark, C.G., Rodgers, F.G., 2002. Detection in *Escherichia coli* of the genes encoding the major virulence factors, the genes defining the O157:H7 serotype, and components of the type 2 Shiga toxin family by multiplex PCR. J. Clin. Microbiol. 40, 3613-3619.

Watanabe, H., Wada, A., Inagaki, Y., Itoh, K., Tamura, K., 1996. Outbreaks of enterohaemorrhagic *Escherichia coli* O157:H7 infection by two different genotypes-strains in Japan. Lancet 348, 831-832.

Wei, J., et al., 2003. Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. Infect Immun. 71, 2775-2786.

Wick, L.M., Qi, W., Lacher, D.W., Whittam, T.S., 2005. Evolution of Genomic Content in the Stepwise Emergence of *Escherichia coli* O157:H7. J. Bacteriol. 187, 1783-1791.

FIGURE LEGENDS

Fig.1. Representative locations of prophages and phage-like regions on the EDL933 genome. Open boxes show the distribution of the prophages and phage-like regions from positions 1.0 to 2.8-Mb on the EDL933 chromosome. Filled triangles in the boxes indicate integrase genes. Open triangles indicate the *Xba*I sites and the asterisk indicates the location of the terminus. Black bars show the locations of DNA probes used for Southern hybridization analyses. The size of each block is proportional to the size of the corresponding island described by Perna et al. (2001).

Fig. 2. PFGE and Southern hybridization analyses of EOT-1 (pattern H) and E50T-11 (pattern M) restricted with *Xba*I, hybridized with the 933M and 933U-targeted probes. The probes listed at the bottom of the lanes. Lane Mk, Lambda Ladder PFG Marker (New England Biolabs). Arrows indicate the estimated sizes of the fragments detected by hybridization. Arrowheads indicate the fragments correlated with the deletion associated by IS629.

Fig.3 Diagrammatic representations of genomic rearrangements between E0T-1 and E50T-11. (A) The prophage-associated inversion (type 1 inversion). (B) The IS629-associated deletion. Filled triangles in the boxes indicate integrase genes. Open triangles indicate the *Xba*I sites and the asterisk indicates the location of the terminus region. Black bars show the locations of DNA probes used for Southern hybridization analyses and PCR tests. Gray bars show the *Xba*I fragments changed the size by genomic rearrangements.

Fig. 4. PFGE and Southern hybridization analyses with specific probes designed from resident prophages and phage-like regions. (A) Analysis of E0T-1 (pattern

H) and E20T-9 (pattern H3), hybridized with the 933M and 933U-targeted probes. (B) Analysis of E0T-1 (pattern H) and E30T-12 (pattern H5), hybridized with the 933O and 933R-targeted probes. (C) Analysis of E50T-11 (pattern M) and E20T-12 (pattern M4), hybridized with the 933O and 933P-targeted probes. (D) Analysis of E50T-11 (pattern M) and E20T-15 (pattern M3), hybridized with the OI 43 and OI 48-targeted probes. (E) Analysis of E50T-11 (pattern H) and E20T-17 (pattern M6), hybridized with the 933M and 933U-targeted probes. (F) Analysis of ATCC700927 (pattern EDL) and E0T-1 (pattern H), hybridized with the 933O and 933R-targeted probes. The probes listed at the bottom of the lanes. Arrows indicate the estimated sizes of the fragments hybridized with the indicated probes.

Fig. 5. Location and order of the prophages and phage-like regions on the EDL933 chromosome. Gray bars indicate the 5 types of inverted segments.

Fig. 6. Diverse genome structures among representative EDL933 derivatives. Open triangles indicate the *Xba*I sites and the asterisks indicate the location of the terminus of replication. Inverted segments are shown as gray bar at the bottom and new fragments yielded by inversions indicate AA to JJ. The sizes (in kilobases) of the fragments based on the EDL933 genome (Perna et al., 2001) are indicated at the bottom.

TABLE 1. Bacterial strains used in this study

EDL933 derivatives	Relevant characteristics	PFGE fragment pattern	reference or origin[a]
E0T-1	Initial strain of repeated subculturing	H	Iguchi et al. (2002)
(E0T-1 derivatives)			
E50T-11	Isolated from a 50-time-subculture	M	Iguchi et al. (2002)
H20T-9	Isolated from a 20-time-subculture	H3	This work
H30T-12	Isolated from a 30-time-subculture	H5	This work
(E50T-11 derivatives)			
M20T-12	Isolated from an additional 20-time-subculture	M4	This work
M20T-15	Isolated from an additional 20-time-subculture	M3	This work
M20T-17	Isolated from an additional 20-time-subculture	M6	This work
ATCC 700927	Sequenced strain of EDL933	EDL	ATCC

[a]ATCC, American Type Culture collection

TABLE 2. Primers used in this study

Primer [a]	Sequence (5'-3')	length (bp)
OI43A-F	TCGCGCTGGATAAAGAGAAA	383
OI43A-R	CGCATTTTATCTGCCTGTACG	
OI43B-F	CAAACCTCATGGTGTGCAA	425
OI43B-R	TACCGTTCTTGAAACGTTGC	
933MA-F	TGTAACGCTCTTTCTGGCTTT	406
933MA-R	GTAGACACAGGATTACCGGAA	
933MB-F	TAGAGGCCGATGTAAAGCGTT	496
933MB-R	ATTTAGGGAAGCCCGTCTGAT	
OI48A-F	TCCAGGCGGTACGCTTTTTT	471
OI48A-R	GTATATTGCACACACCGGGT	
OI48B-F	TGGAACGGGCGCTAATTTA	398
OI48B-R	TGCGAACACCGCTTTAAGAT	
933OA-F	TAACGATCATGGCAAAGAGGC	462
933OA-R	TGGATTACAACTGGCGTGCTA	
933OC-F	TGTTACGCCATCACTACGATG	300
933OC-R	TCGACATCTGGTGGTGGA	
933OD-F	AAAGGTATTGGCCGGGATT	343
933OD-R	TGAACGCATAACCCAACGTA	
933OB-F	TTATATTCGCTCGGCGCAG	409
933OB-R	TCAAGGTTCCGCAGTTAGTTC	
933RB-F	AAGCCAAATCGCAACTGGAA	382
933RB-R	TCAGAAGAGTATGCGTCATGC	
933RA-F	CAACCAAACCATTTGAAGGAG	506
933RA-R	GCGAATACCAGCACGTTTTA	
933PB-F	CTGCATTATTTCCCGACTCA	393
933PB-R	TATCAATGCTTGTGGCCAGA	
933PA-F	TGTGAVTGGTATTCTGCAACC	459
933PA-R	AACGCGCCTGTTTTTTGCT	
933UB-F	CCGATACATCATGCTCTCTGA	473
933UB-R	TGTAATATTGCGGTTGACGG	
933UA-F	TGGGTAAGCTTTCAGGCGT	613
933UA-R	CATTGCCGACCTCAAGTAAA	
933YA-F	CAATGCTCAACAACATACCGA	343
933YA-R	CCTTTGTTGCAAGTGATGCT	
933YB-F	TGGTGTCAAGTTTCCCCAAA	466
933YB-R	AAATACCGTGGCGTGGCAAA	
216U-F	CCAGTTCTTCGTTGCTGTTGA	354
216U-R	TGCCGAACAAAGAAGTGATG	
216V-F	AGAACAACAACTGACGATGCG	393
216V-R	GCGGTATTGGTCAGCTTTCC	
216W-F	GAAAATGCTTGATGAGTGGG	340
216W-R	ATTTATCCGGGGCAAGTGTT	
216X-F	ATCTAGTTACAGGGACGCCAG	464
216X-R	CGCAACGCTGAGTTACAACAT	

a. Primers were used for preparation of part of the probes for Southern and PCR analyses. The primers are named systematically as follows. The head of the names indicate the targeted regions as indicated in Fig. 1 and 3B. ″-F″ and ″-R″ indicate the sense and anti-sense primers, respectively.

TABLE 3. Occurrence of different PFGE fragment patterns in two strains of EDL933 derivatives during a 50-time period of repeated subculturing

Strain and PFGE fragment pattern	No. of fragment differences from pattern H	No. of fragment differences from pattern M	No. of colonies (n=18) showing various PFGE patterns after number of subculturing				
			10	20	30	40	50
E0T-1							
H	Original pattern		18	11	13	6	3
H1	2			1			
H2	3			4	4	12	14
H3	4			1			
H4	1			1			
H5	4				1		
H6	1						1
E50T-11							
M	6	Original pattern	18	10	8	6	
M1	6	2		1	1	1	
M2	9	3		2			
M3	6	4		2			
M4	10	4		1	2	9	12
M5	8	2		1	5		2
M6	2	4		1			
M7	7	1			1		
M8	8	2			1	1	2
M9	8	2				1	
M10	6	4					1
M11	8	6					1

Fig. 1

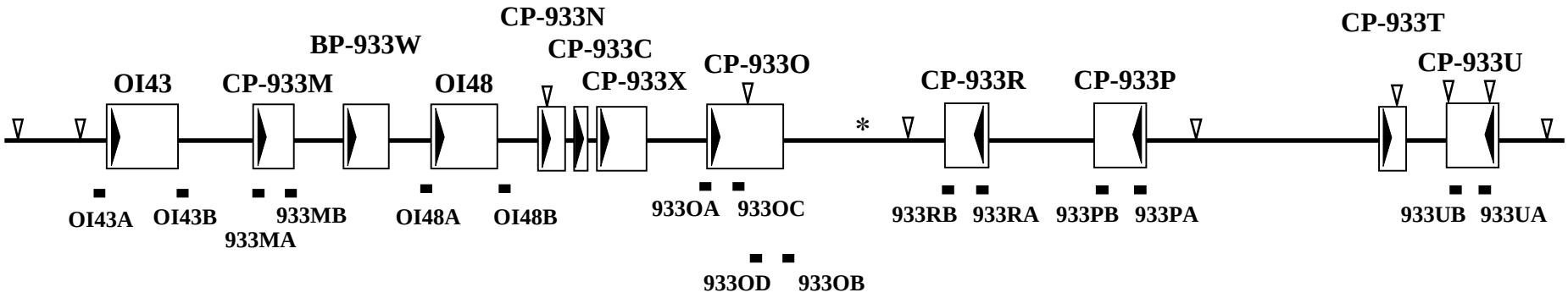

CP-933N
BP-933W
CP-933C
CP-933T
OI43
CP-933M
OI48
CP-933X
CP-933O
CP-933R
CP-933P
CP-933U
*
OI43A
OI43B
933MA
933MB
OI48A
OI48B
933OA
933OC
933OD
933OB
933RB
933RA
933PB
933PA
933UB
933UA

Fig. 2

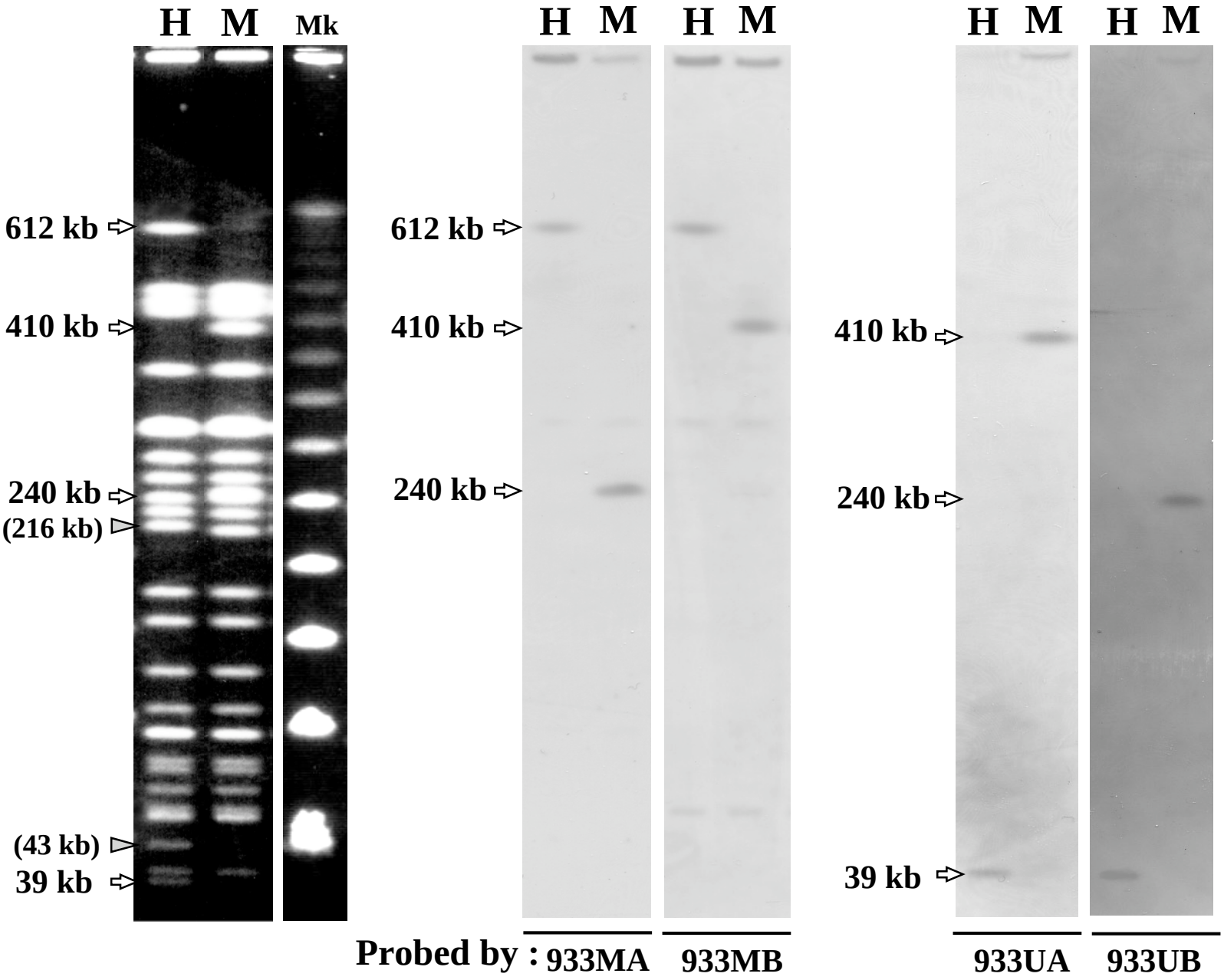

H
M
Mk
612 kb
410 kb
240 kb
(216 kb)
(43 kb)
39 kb
H
M
H
M
612 kb
410 kb
240 kb
H
M
H
M
410 kb
240 kb
39 kb
Probed by :
933MA
933MB
933UA
933UB

Fig. 3 (part 1)

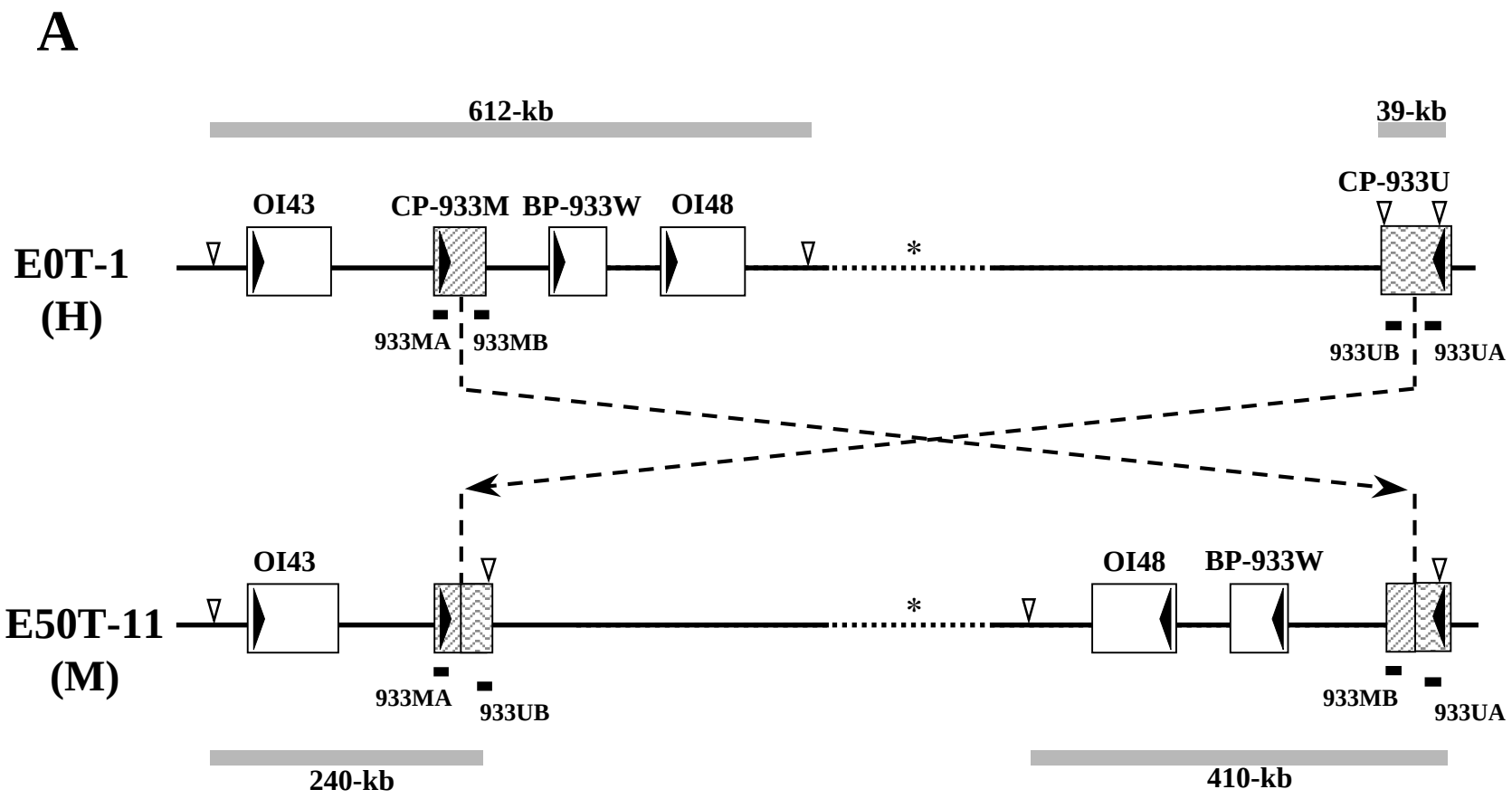

A
612-kb
39-kb
OI43
CP-933M
BP-933W
OI48
CP-933U
E0T-1
(H)
933MA
933MB
933UB
933UA
OI43
OI48
BP-933W
E50T-11
(M)
933MA
933UB
933MB
933UA
240-kb
410-kb

Fig. 3 (part 2)

B

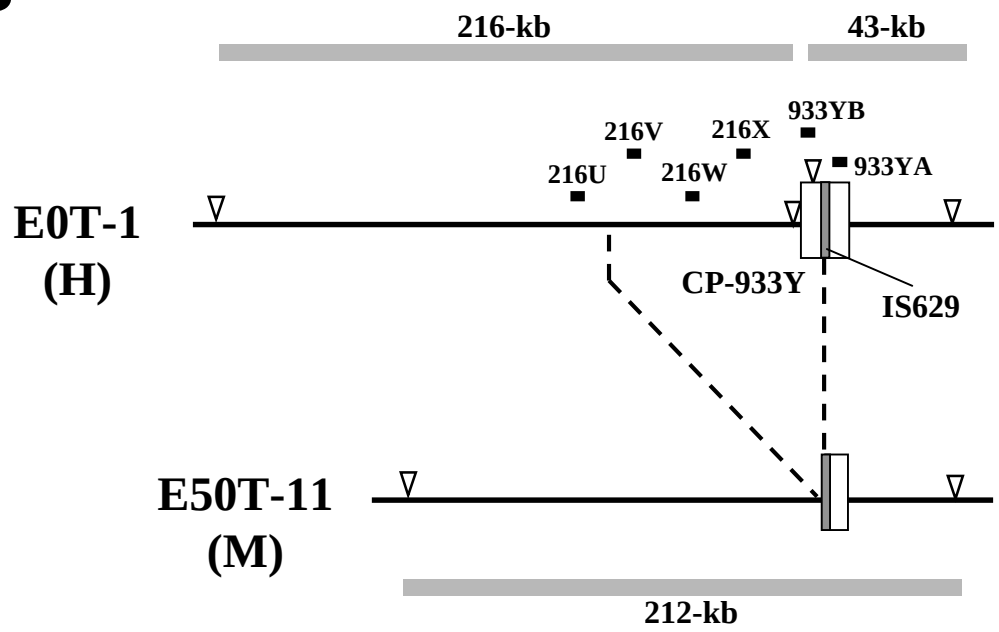

216-kb
43-kb
933YB
216V
216X
933YA
216U
216W
E0T-1
(H)
CP-933Y
IS629
E50T-11
(M)
212-kb

Fig. 4 (part 1)

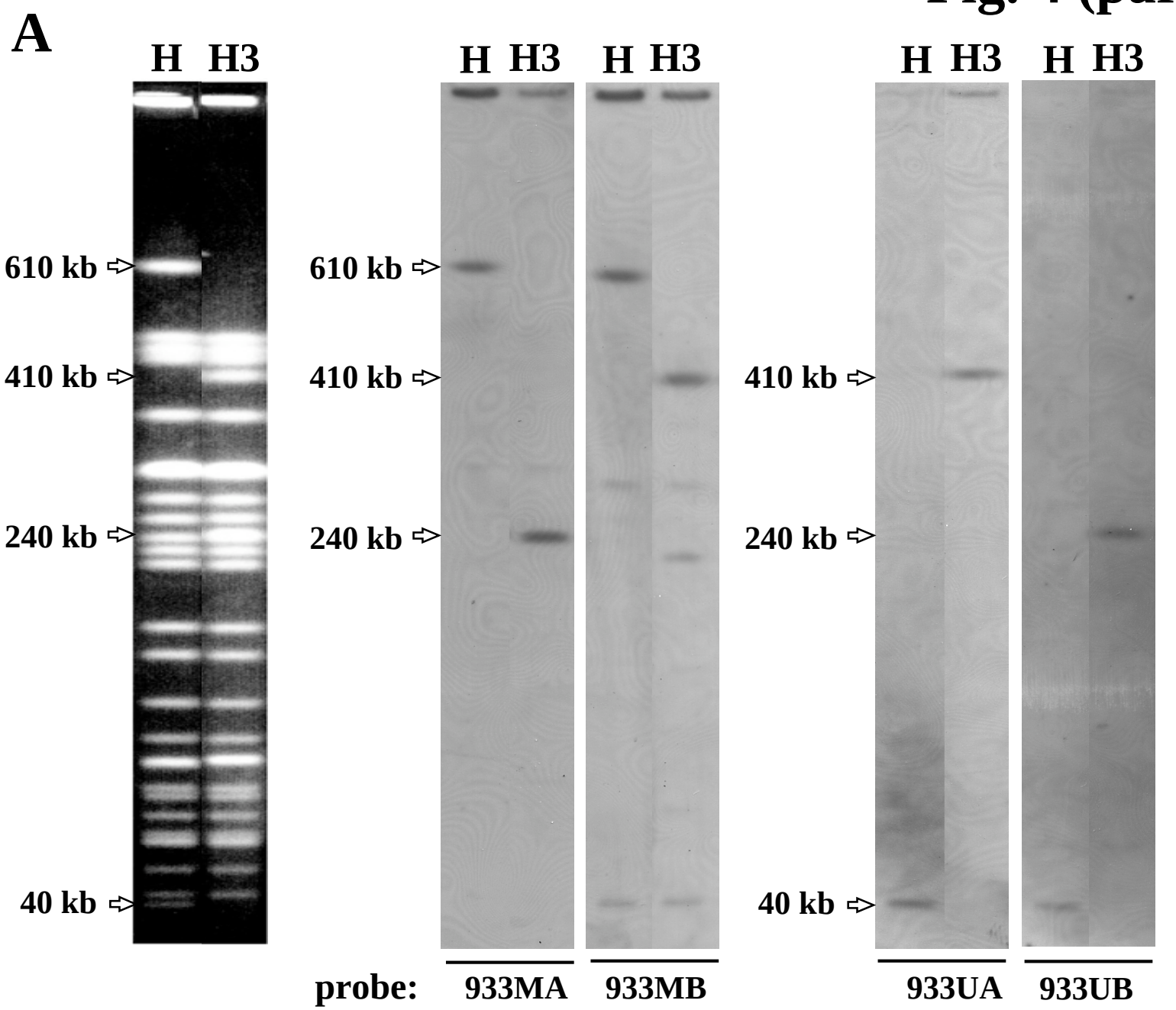


Fig. 4 (part 2)

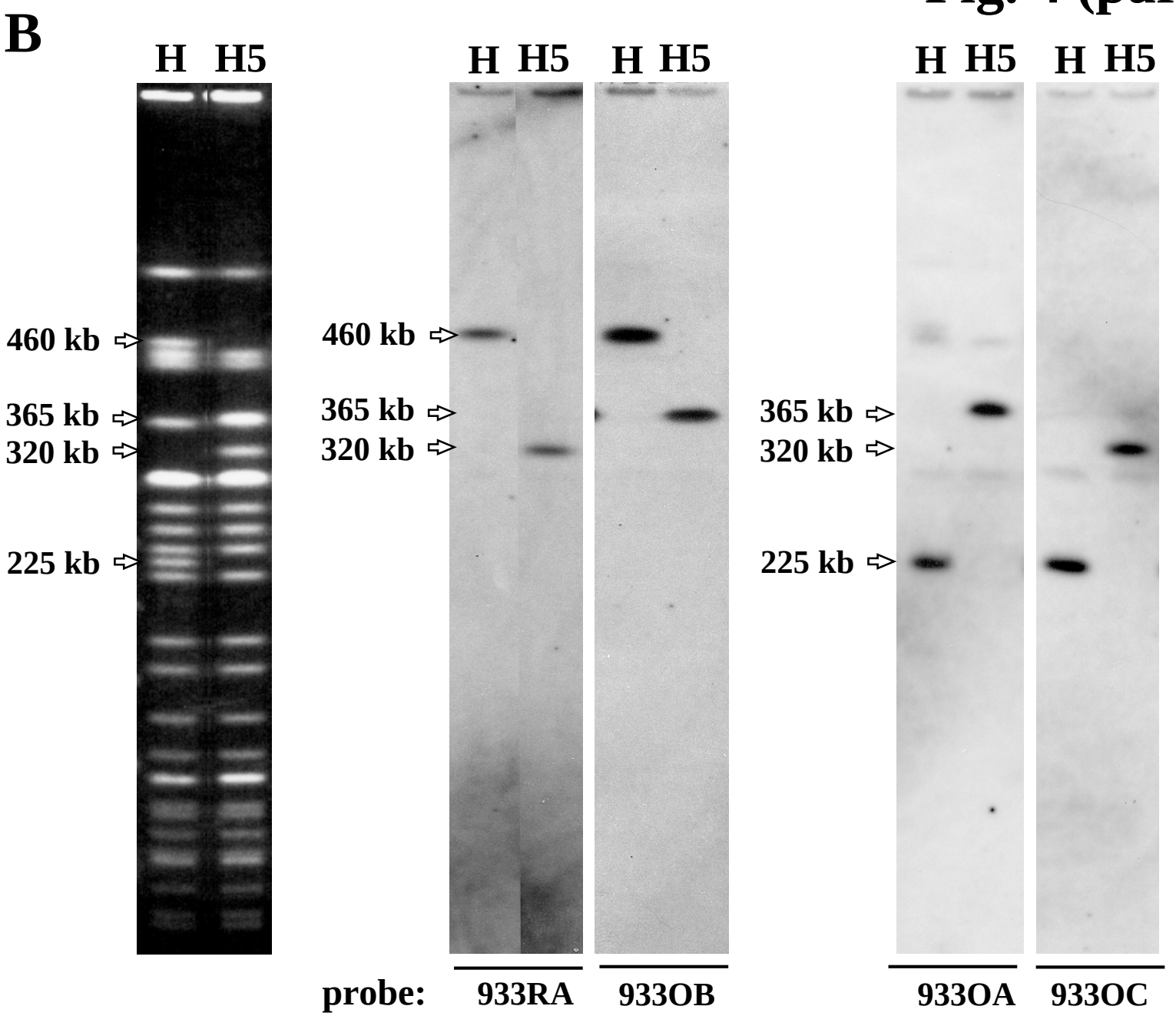

B
H
H5
H
H5
H
H5
H
H5
460 kb
365 kb
320 kb
225 kb
460 kb
365 kb
320 kb
365 kb
320 kb
225 kb
probe:
933RA
933OB
933OA
933OC

Fig. 4 (part 3)

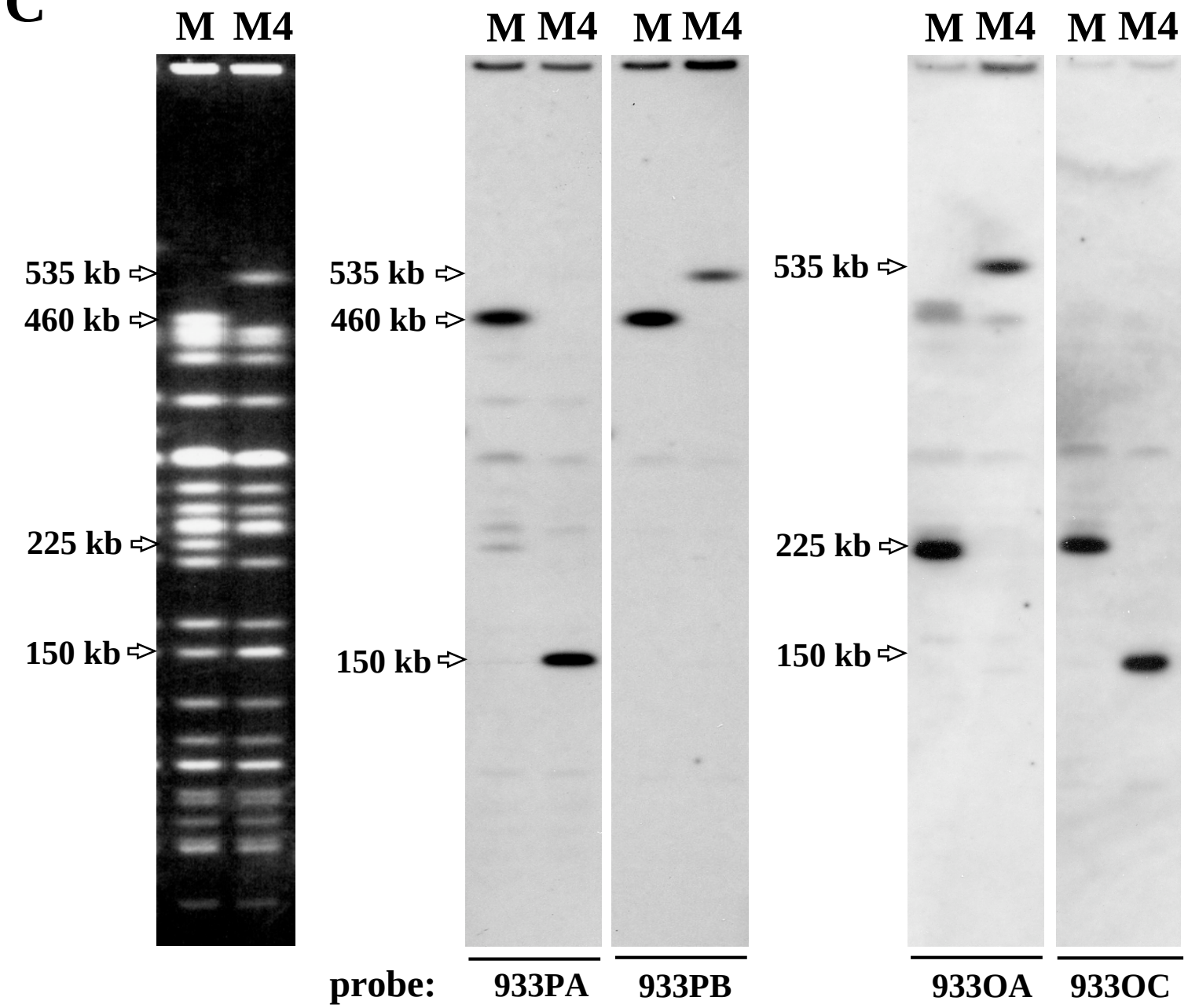

C
M M4
535 kb
460 kb
225 kb
150 kb
M M4 M M4
535 kb
460 kb
150 kb
M M4 M M4
535 kb
225 kb
150 kb
probe:
933PA
933PB
933OA
933OC

Fig. 4 (part 4)

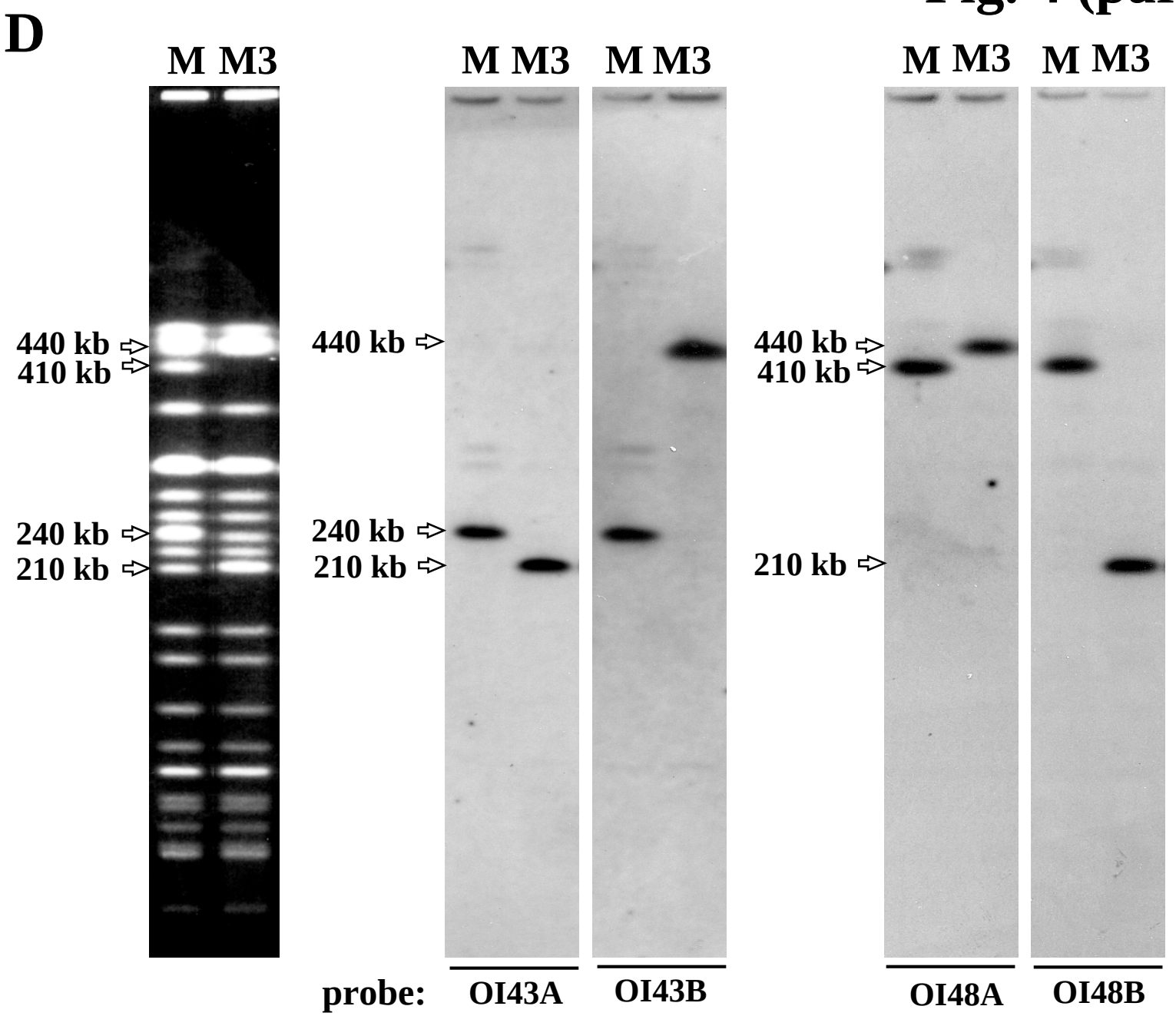

D
M M3
M M3
M M3
M M3
M M3
440 kb
410 kb
240 kb
210 kb
440 kb
240 kb
210 kb
440 kb
410 kb
210 kb
probe:
OI43A
OI43B
OI48A
OI48B

Fig. 4 (part 5)

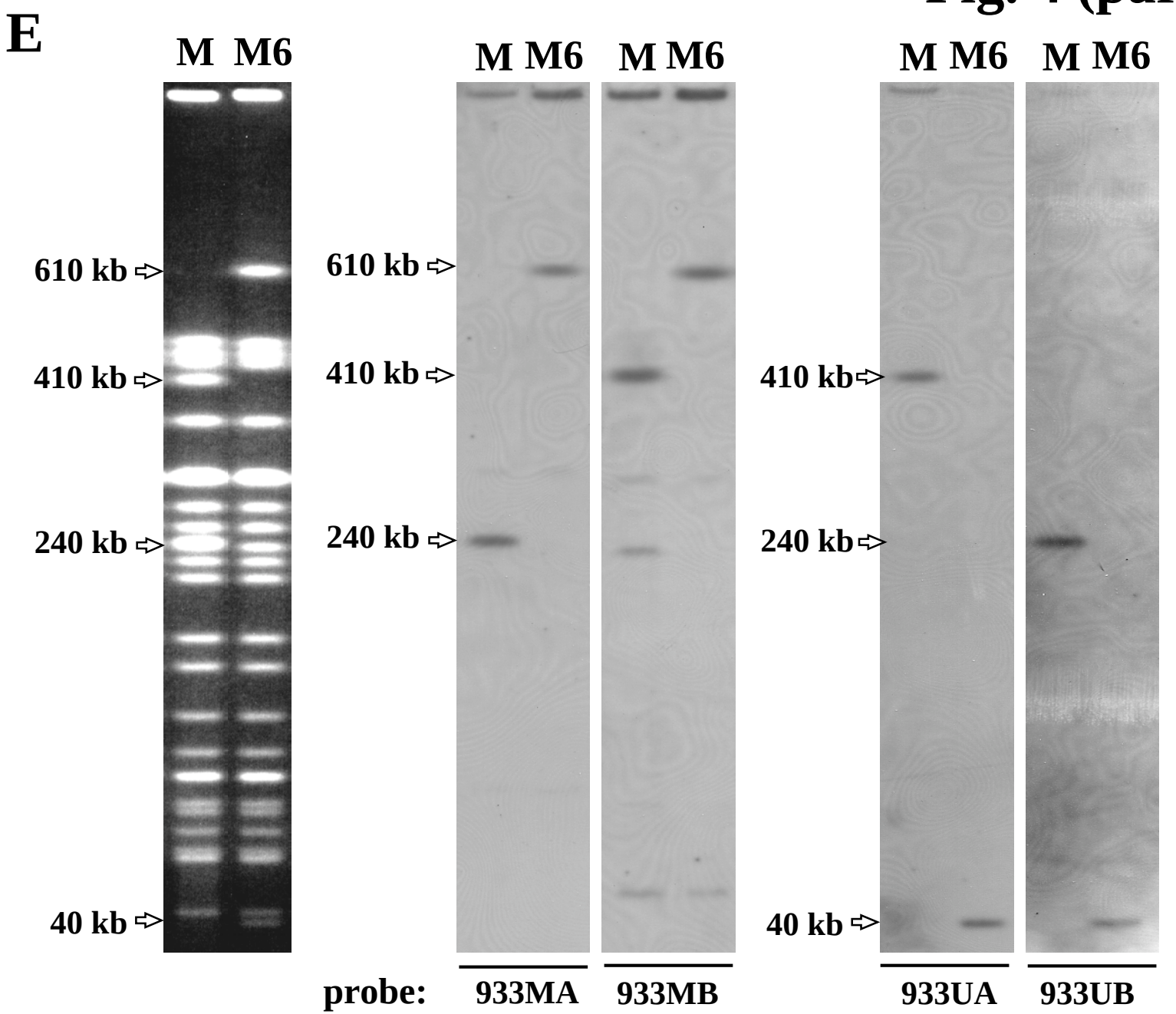

E
M M6
610 kb
410 kb
240 kb
40 kb
M M6 M M6
610 kb
410 kb
240 kb
M M6 M M6
410 kb
240 kb
40 kb
probe:
933MA
933MB
933UA
933UB

Fig. 4 (part 6)

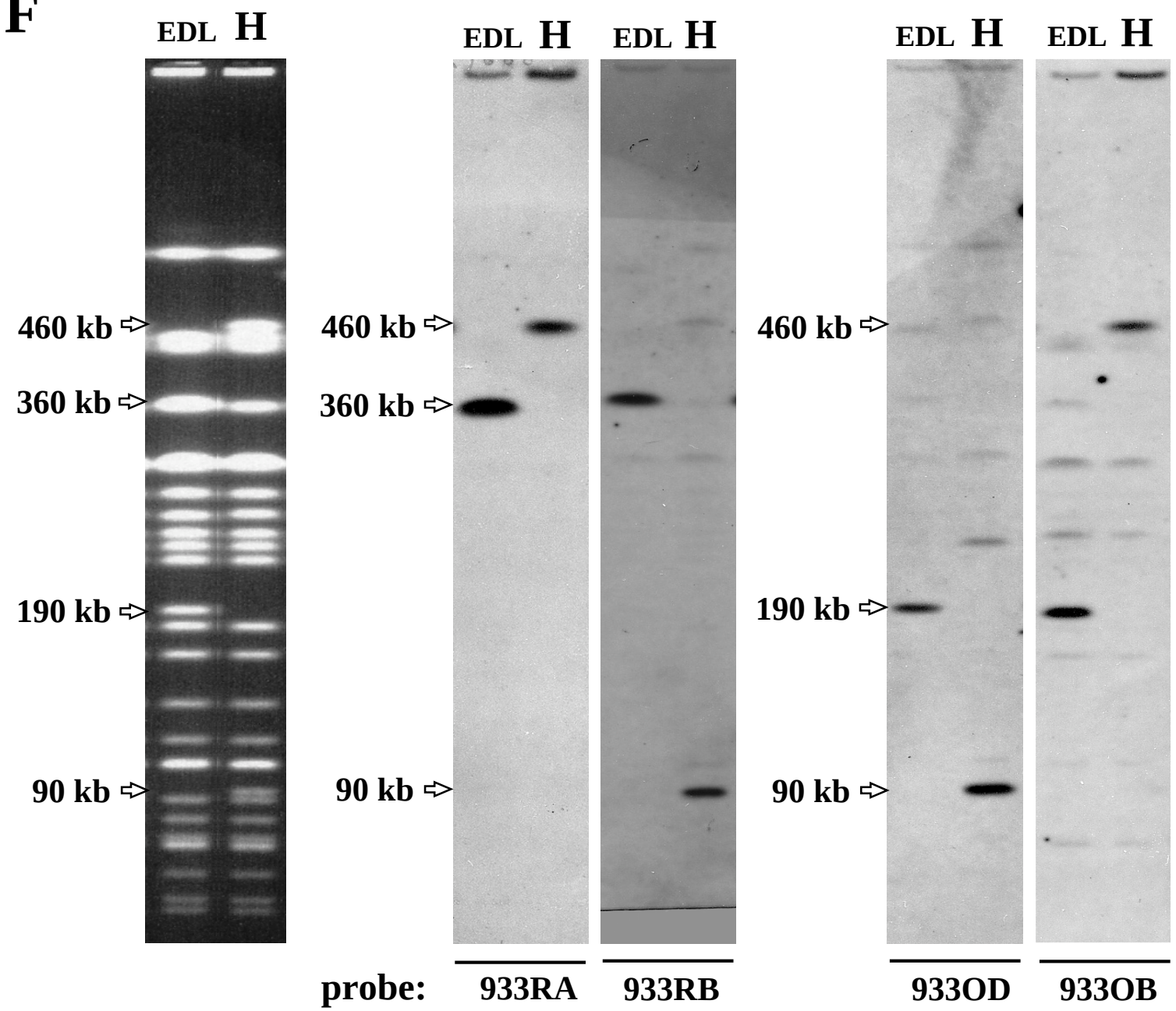


Fig. 5

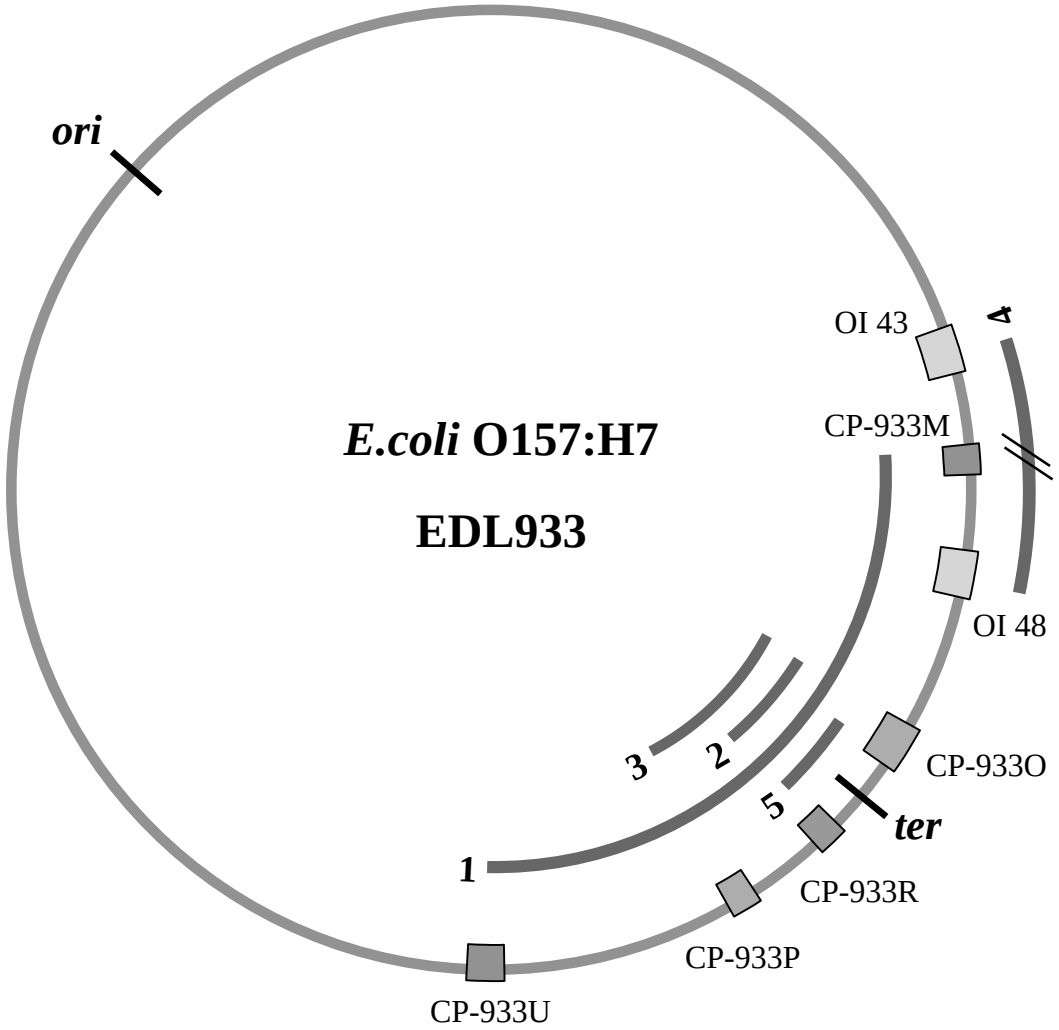
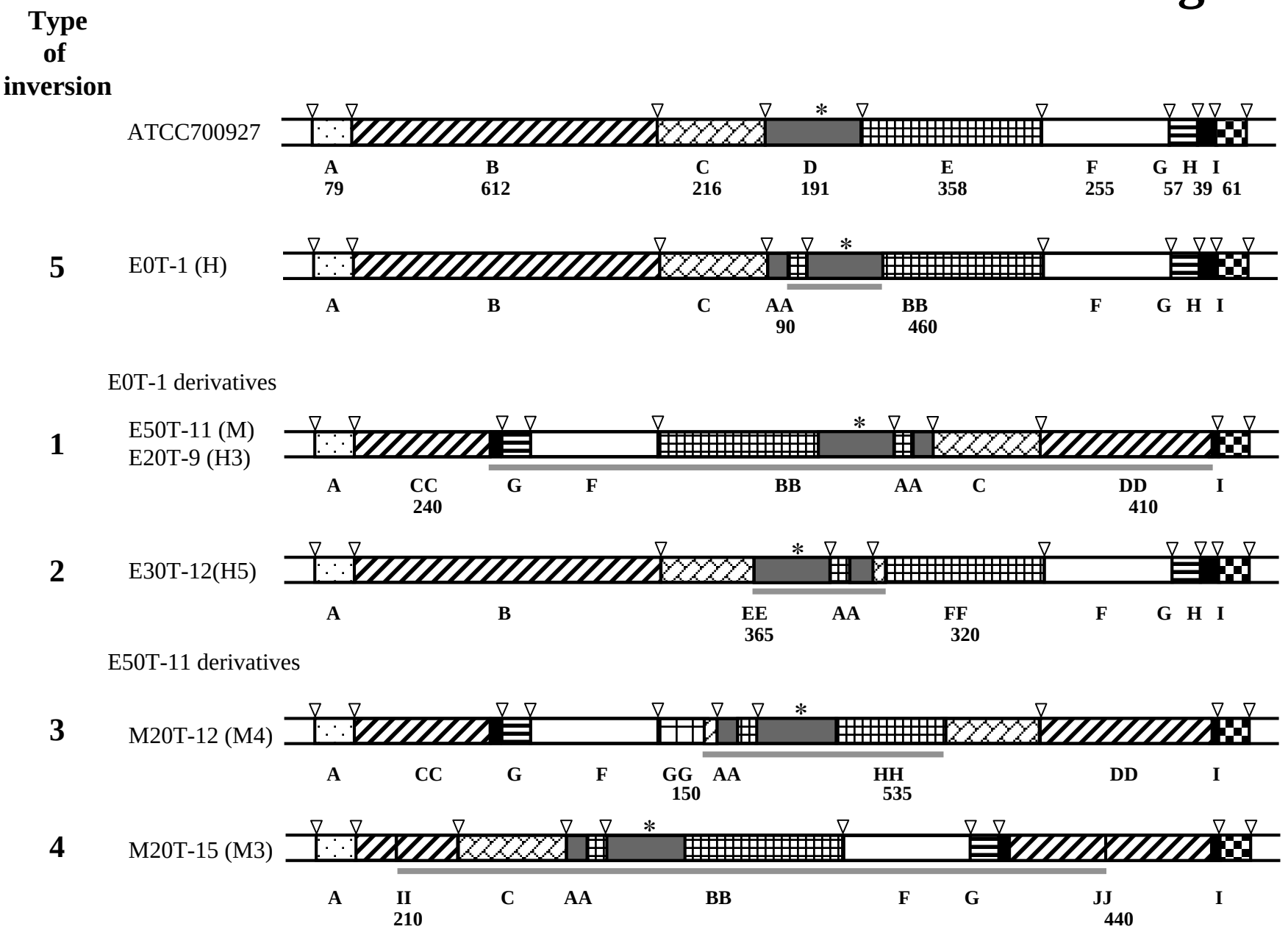

ori
E.coli O157:H7
EDL933
OI 43
4
CP-933M
OI 48
3
2
CP-933O
5
ter
1
CP-933R
CP-933P
CP-933U

Fig. 6


Type of inversion
ATCC700927
A 79
B 612
C 216
D 191
E 358
F 255
G 57
H 39
I 61
5
E0T-1 (H)
A
B
C
AA 90
BB 460
F
G
H
I
E0T-1 derivatives
1
E50T-11 (M)
E20T-9 (H3)
A
CC 240
G
F
BB
AA
C
DD 410
I
2
E30T-12(H5)
A
B
EE 365
AA
FF 320
F
G
H
I
E50T-11 derivatives
3
M20T-12 (M4)
A
CC
G
F
GG 150
AA
HH 535
DD
I
4
M20T-15 (M3)
A
II 210
C
AA
BB
F
G
JJ 440
I