Volume 14 Number 17 1986 Nucleic Acids Research

ISH51: a large, degenerate family of insertion sequence-like elements in the genome of the archaebacterium, *Halobacterium volcanii*

Jason D.Hofman, Leonard C.Schalkwyk and W.Ford Doolittle

Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

Received 30 April 1986; Revised 9 July 1986; Accepted 15 July 1986

ABSTRACT

We describe a new family of repetitive elements in the genome of the archaebacterium Halobacterium volcanii. There are some 20-30 copies of this element, which we designate ISH51. Sequenced copies show typical insertion sequence characteristics (terminal inverted repeats, direct flanking repeats of "target site" DNA). However, members of the ISH51 family are highly heterogeneous, showing on average only 85% primary sequence homology; and some genomic copies appear to be severely truncated. Some ISH51 elements are clustered together in regions of relatively AT-rich DNA. There are at least five such AT-rich "islands" in the H. volcanii genome. Repetitive sequences homologous to ISH51 are found in the genomes of most Halobacterium and Halococcus species.

INTRODUCTION

In 1977, Woese and his collaborators proposed a new universal phylogeny, in which all contemporary living things can be seen as descendent from one of three ancient evolutionary lineages, or "primary kingdoms" (1,2). Members of the first two kingdoms, the eubacteria and eukaryotes (or more precisely the nuclear-cytoplasmic component of eukaryotic cells), are now in general well understood, in terms of basic features of gene and genome structure, gene organization and gene expression. Members of the third primary kingdom, the archaebacteria, have only recently become the subject of molecular biological scrutiny. What little we do know suggests that these organisms (methanogens, halobacteria and certain sulfur-dependent and often thermophilic bacteria) will show a mixture of eubacterial, eukaryotic and unique molecular phenotypes. For instance: archaebacterial tRNA and rRNA genes sometimes have introns; archaebacterial RNA polymerases are complex in subunit composition and immunologically related to eukaryotic polymerases; archaebacterial promoters are neither eukaryotic nor eubacterial in structure; archaebacterial messages may sometimes lack 5' leaders; and archaebacterial ribosomes, although bearing 16S, 23S and 5S rRNAs, are unique in structure and spectrum

of antibiotic sensitivity (for reviews see 3-5).

Among archaebacteria, some species of Halobacterium are further unusual in showing a high level of genetic instability. Spontaneous mutations to readily scored pigment (bacteriorhodopsin, bacterioruberin) or gas vacuole-deficient variants occur at frequencies between 10^{-4} and 10^{-2} (6,7). These organisms also characteristically show, in CsCl gradients, a main ("chromosomal") DNA band of 68 mol % G+C and a "satellite" band of 59 mol % G+C (8,9). These two observations might be interpreted as due to frequent loss of a pigment and gas vacuole-encoding plasmid. In many halobacterial species, much of the satellite DNA is indeed plasmid DNA (6,10). However, most scoreable mutations analysed are associated with insertions into (or rearrangements of) plasmid and chromosomal DNA, and not plasmid loss (6,7,10).

High-frequency phenotypic variation and rearrangement could also be due to the activity of endogenous transposable elements. In an initial attempt to detect such elements as repeated sequences in H. halobium, Sapienza and Doolittle (11) screened random Eco RI clones of genomic DNA, by Southern blotting. They showed that the H. halobium genome harbors as many as fifty different families of repeated sequences, each family having from 2 to 20 members, so that there might be in total as many as 1,000 repetitive elements in this species, distributed between plasmid and chromosomal DNA. Genomic rearrangements associated with these repeated DNAs are detectable by Southern blotting of DNA from randomly-picked colonies, and occur at very high frequencies -- between 10^{-4} and 10^{-3} per repeat sequence family per generation (12).

Many if not all of the repeated sequences in H. halobium DNA are transposable elements. Work from the laboratories of Pfeifer and Boyer and Rajbhandary and Khorana has shown that nearly all mutations affecting the expression of the bacterio-opsin (bop) gene are the result of insertions into or near that gene of insertion sequence-like elements, most of which are present in multiple copies in the genome of H. halobium (13-17). To date, four of these elements (ISH1, ISH2, ISH50 [or ISH23] and ISH1.8, all from H. halobium) have been sequenced in their entirety and several others have been characterized preliminarily (17-20). All except ISH1.8 show "typical" insertion-sequence structure; that is, these elements are from 500 to 3,000 bp long, contain terminal inverted repeats and (usually) one or two open reading frames, and are flanked by direct repeats of target site sequences. ISH1.8, on the other hand, lacks terminal inverted repeats and is not invariably flanked by target site duplications (20).

Pfeifer and Goebel and their coworkers have concluded that most inser-

tion sequence-like elements in H. halobium reside in the relatively AT-rich fraction (fraction II, 59 mol % G+C) of H. halobium DNA, which can be separated from the major (68 mol % G+C) DNA component (fraction I) by density-gradient centrifugation or malachite-green bisacrylamide gel column fractionation. Although much of fraction II DNA is itself plasmid DNA, a significant proportion is probably chromosomal (9,15,21). The notion that insertion sequences are confined to large, relatively AT-rich "islands" within plasmid and chromosomal DNA, and that most movements of transposable elements are within or between such islands, is consistent with available physical data (9,15), and would explain how halobacteria can survive genomic rearrangements as frequent as those observed by Sapienza and coworkers (12). Such rearrangements would, if unrestricted, produce multiple mutations in the majority of cells in a clone after only a few generations.

In their initial report, Sapienza and Doolittle (11) also described random clones of DNA from the distantly-related halobacterial species Halobacterium volcanii which probed multiple bands in Southern hybridization against restriction endonuclease-digested H. volcanii genomic DNA. Several of these random Eco RI clones probed the same set of 20 to 30 Eco RI fragments of H. volcanii genomic DNA, and a more limited number of (different) fragments of H. halobium DNA. The cloned fragment designated v97 is one of these. The H. volcanii cloned fragment designated v147 probed this same set of 20 to 30 bands, and an additional dozen or more fragments in H. volcanii DNA, as if it carried copies of two different repetitive sequences, only one of which was present on v97 (11). Here we report the results of further characterization of the repetitive sequences families borne on fragments v97 and v147.

METHODS

General Methods

Southern hybridizations were performed with nitrocellulose-bound DNA using ^{32}P-labelled M13 (22), nick-translated plasmid and λ or H. volcanii tRNA (23) probes. H. volcanii total DNA (≈100 kb) was fractionated by malachite green bisacrylamide chromatography (6,24) with 2 ml fractions collected from a 40 ml $NaClO_4$ gradient (0-1 M). Dot blots contained 1 μg of FI (68 mol % G+C) or 0.3 μg of FII (59 mol % G+C) DNA taken from single 2 ml fractions eluting at ≈0.4 M and ≈0.6 M $NaClO_4$, respectively. DNA fractionation by pulsed-field gel electrophoresis (25,26) on 1% agarose gels was performed with a 7.5 sec pulse time at 11° and transferred to a nylon membrane (27).

Cloning

Eco RI inserts and Sau 3AI, Taq, I and Hpa II subfragments from v97 and v147 (11) were cloned into M13 (22). Bam HI cleaved H. volcanii DNA was cloned into λ 1059 (28) and λHvo7 was selected by plaque hybridization with a v147 probe. The Bam HI insert from λHvo7 was digested with Xho I and the fragments were cloned into pUC13. The 4.0 kb Xho I fragment cleaved once with Bgl II and the adjacent 1.1 kb Xho I-Bam HI fragments were cloned into M13mp19. M13 deletion clones were generated by exonuclease III and S1 nuclease (29) or exonuclease VII (30) digestion of the replicative form DNA or by T4 polymerase digestion of linearized single strand M13 DNA (31).

Sequence Analysis

The sequence of M13 clones was determined by the dideoxy method (22). Sequence alignments were generated with the Beckman Microgenie program.

RESULTS

Characterization of clones bearing repetitive sequences

Clones v97 and v147 were obtained previously (11), by ligating 0.5-3 kb Eco RI fragments of total H. volcanii DNA into Eco RI-digested pBR322. To analyse the distribution of repeats on v147, the cloned fragment was digested with Sau 3AI, Taq I or Hpa II, and subcloned into M13mp9. M13 subclones were then labelled and used as Southern hybridization probes against Eco RI-digested H. volcanii DNA. Fig. 1A shows hybridization results obtained with three subclones. Subclone S4 produced a pattern identical to that obtained with clone 97 -- at least 27 hybridization bands, of varying intensity (c.f. Fig. 3 of ref. 11). Subclone S5 produced 19 hybridization signals. Thirteen of these signals are clearly different from those obtained with subclone S4; another seven may be identical to S4-hybridizing bands. Clone 9H5 produced a pattern which is the summation of those obtained with S4 and S5, and identical to that found for the parental fragment, v147 (11).

To obtain additional copies of the element(s) carried by v97 and v147, a λ bank was prepared using Bam HI-digested H. volcanii DNA and λ 1059. Plaques were screened by hybridization with labelled clone v147. Approximately 10% of the plaques gave positive signals, a frequency consistent with copy number estimates from Fig. 1. One λ clone, here designated λHvo7, was selected for further study. This 15.7 kb Bam HI fragment can be divided into five subfragments (A-E) by digestion with Xho I. These subfragments were cloned into pUC13, separately labelled by nick-translation, and used to probe Eco RI-digested total H. volcanii DNA (Fig. 1B). Subfragment A of λHvo7 produced a

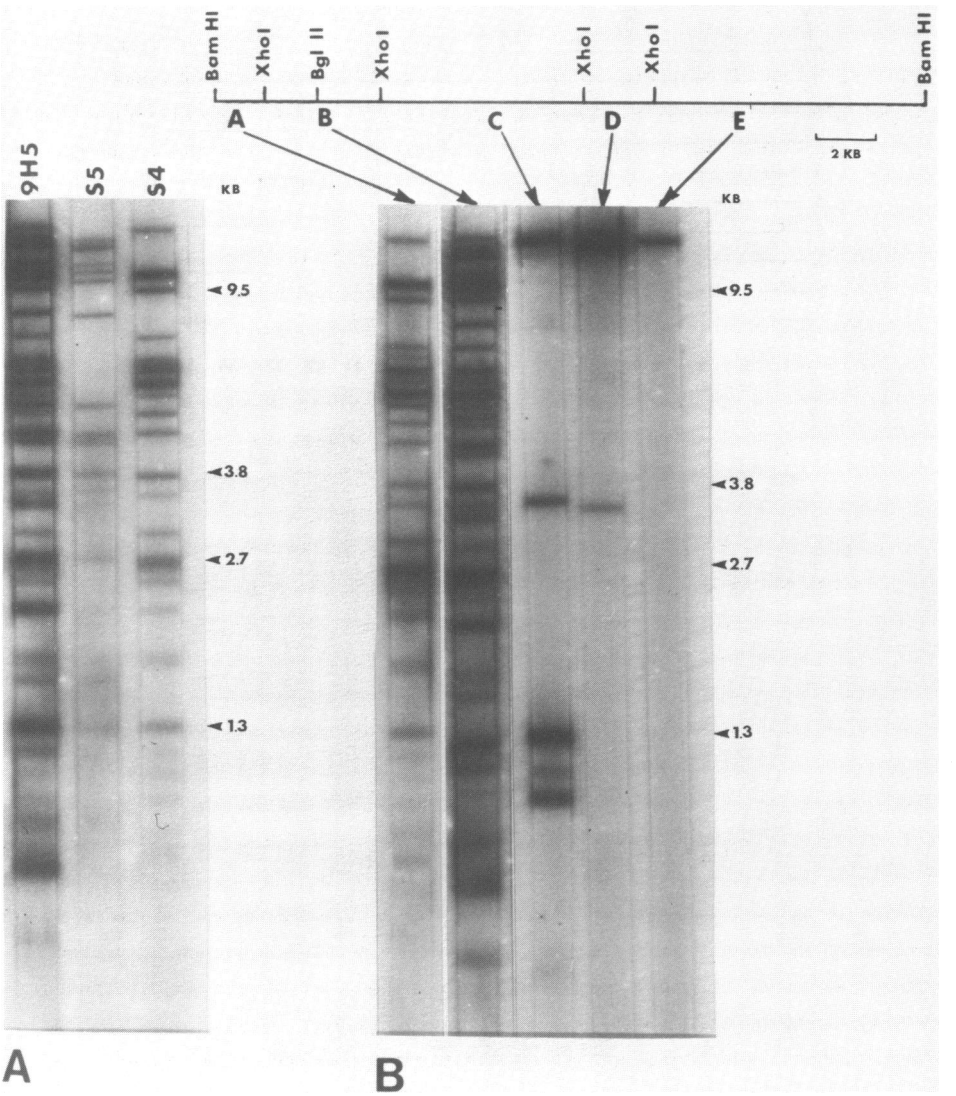


Fig. 1. Southern hybridization against total Eco RI-digested H. volcanii DNA. Panel A: Probes are subclones 9H5, S5 and S4 of fragment v147 (see Fig. 6 for locations of these subclones). Panel B: Xho restriction map of λHvo7, and the results of hybridizing Eco RI-cut genomic DNA with cloned probes of Bam H1-Xho I or Xho I fragments A, B, C, D or E.

hybridization pattern identical to that obtained with v147 subclone S4 (Fig. 1A), except for differences attributable to unique-sequence DNA flanking the repetitive region(s). Subfragment B produced a hybridization pattern iden-

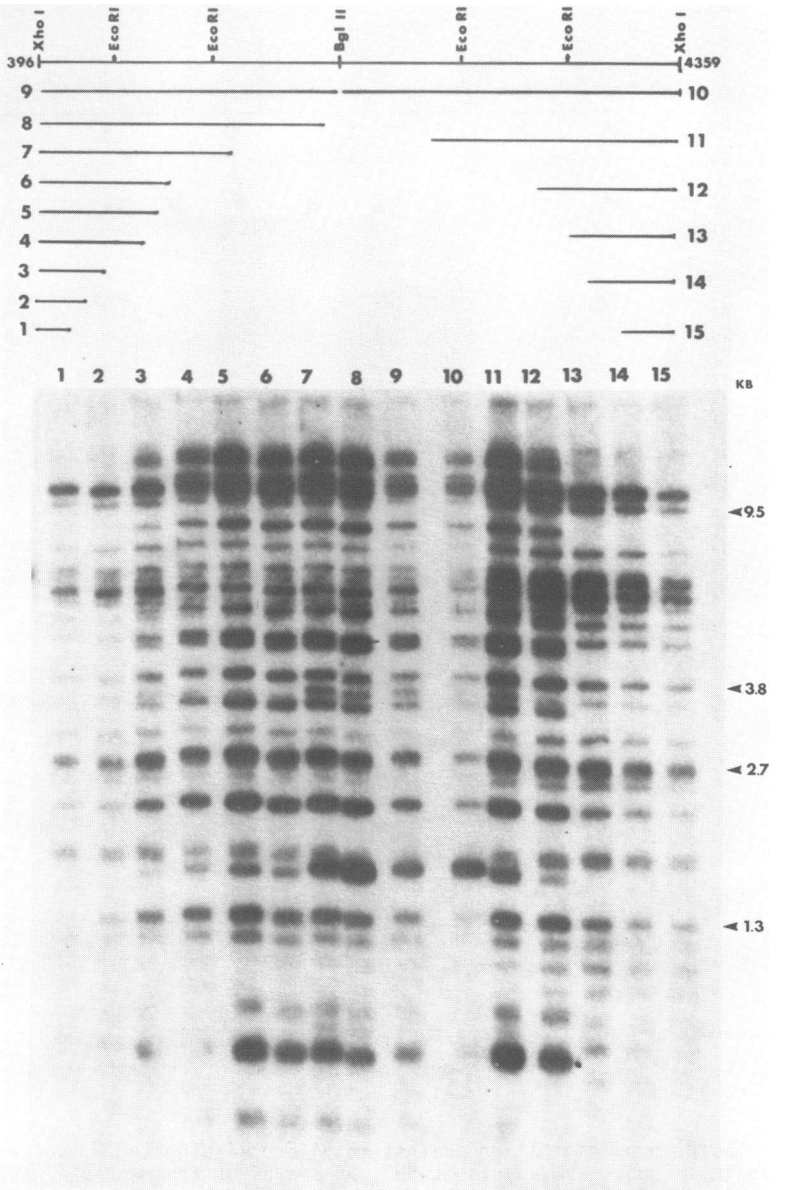


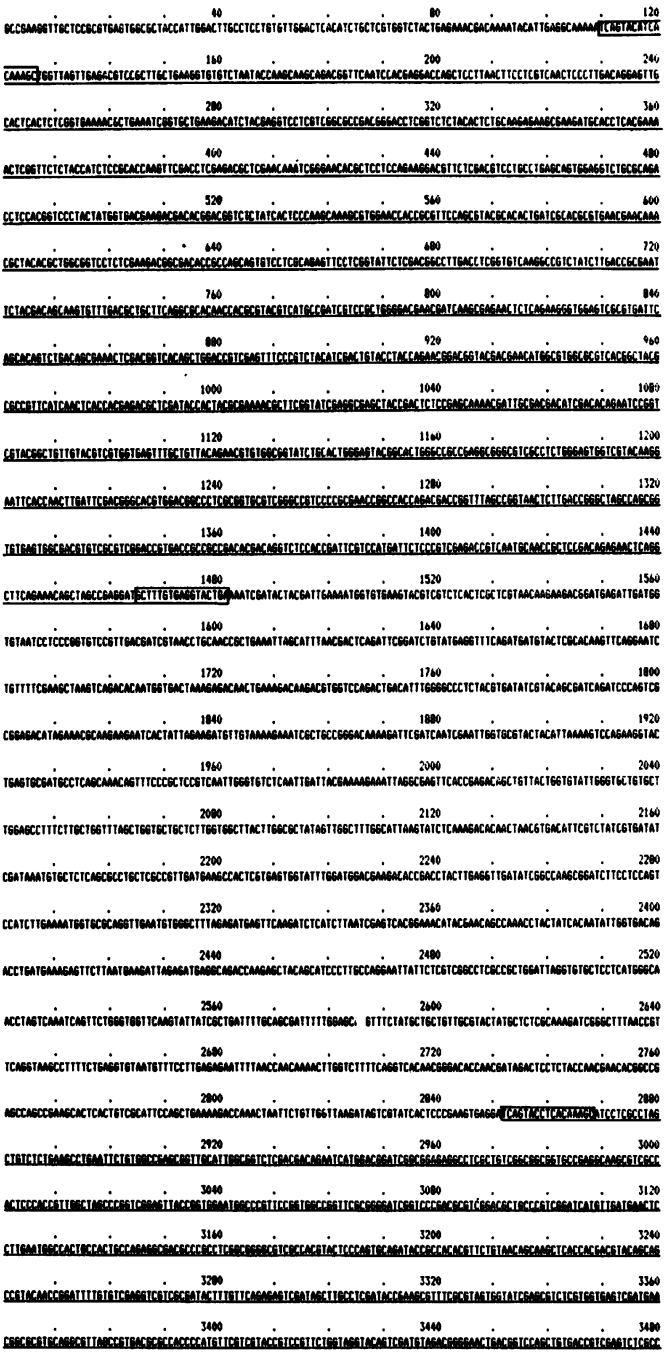
Fig. 2. Southern hybridization against total Eco RI-digested H. volcanii DNA, with deletion probes prepared from fragment B of λHvo7. The length of DNA retained in each deletion clone is indicated by a horizontal bar. The numbers identify the probes and their hybridization pattern. Position numbers from Fig. 3 identify the Xho I termini.

tical to that obtained with v147 subclone 9H5. Subfragment C does not appear (from the results shown in Fig. 1B) to carry copies of the repetitive sequence(s) borne by fragments A and B, but does carry a copy of still another, otherwise uncharacterized repeat.

The locations of sequences on λHvo7 which are repeated in the genome of H. volcanii were further defined by deletion hybridization mapping. Deletions to the left and right of the unique Bgl II site in subfragment B (Fig. 1B) were prepared by exonuclease digestion (see Methods). Hybridization patterns obtained with these deletion fragments as probes are shown in Fig. 2. From these, it is apparent (i) that repetitive elements giving the hybridization pattern of v147 subclone S4 (or of fragment v97) are located, at least, at both ends of subfragment B, and (ii) that a repeated region (or regions) giving the hybridization pattern of v147 subclone S5 (Fig. 1A) is (are) located within the middle one-third to one-half of subfragment B. Similar blotting of deletion fragments extending leftward from the Xho I site in subfragment A of λHvo7 indicated that the repetitive element on the left end of subfragment B extends 50-200 bp into subfragment A (not shown). Subfragments D and E apparently contain only unique sequence DNA.

Sequence analysis

Subfragments A and B were sequenced in their entirety on both strands, using the deletion clones shown in Fig. 2 and a comparable set of deletions from the Xho I sites at the ends of subfragment B and the Xho I and Bam H1 sites of subfragment A. Results are presented in Fig. 3. This sequence shows six interesting features. (i) There are within the 4350 bp shown in Fig. 3, two homologous insertion sequence-like elements, of 1371 bp (left copy) and 1379 bp (right copy) respectively. These lie in inverted orientation with respect to one another, and are underlined in Fig. 3. (ii) Each element contains 16 bp (one mismatch) terminal inverted repeats (boxed in Fig. 3). Each is flanked by short (3 bp) direct duplications, which are presumably target site duplications. We thus consider these elements to be representative of a new family of insertion sequence which we here designate ISH51. This is the first fully characterized insertion sequence described for H. volcanii, and shows no substantial homology to sequenced H. halobium elements. (iii) The distance between insertion sequence copies is 1369 bp -- remarkably close to the size of these elements. There is, however, no obvious homology between the insertion sequences and this central region, and considerable difference in overall G+C content (Fig. 4), so there is no clear explanation for this numerological coincidence. (iv) Neither element contains long open reading



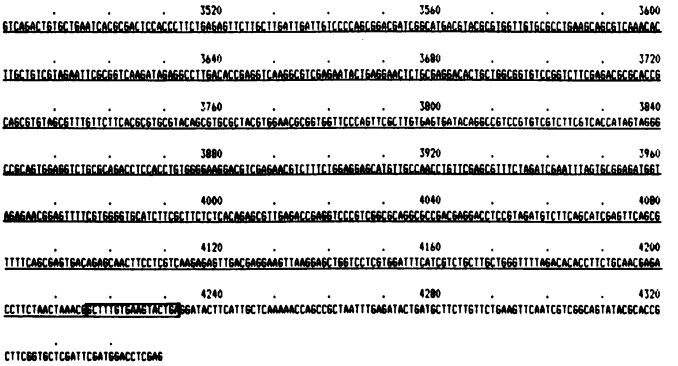


Fig. 3. Sequence of λHvo7 subfragments A and B. The sequence of 660 bp 5' to position 1 is not shown. ISH51 elements are underlined and inverted repeats at the termini of these elements are boxed.

frames, and there are enough base substitutions, insertions and deletions differentiating the two copies to preclude the possibility that both code for a conserved protein. (v) Alignment of the two complete element copies (not shown) reveals only 86.3% sequence homology. (vi) A plot of G+C content (Fig. 4, values determined over 40 bp stretches, and recorded every 20 bp) revealed that ISH51 left and right copies are of substantially higher G+C con-

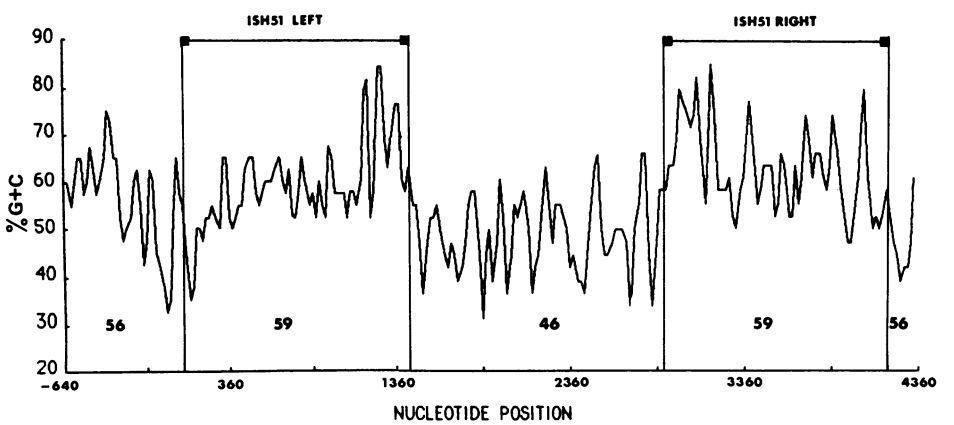


Fig. 4. A plot of G+C content for A and B subfragments of λHvo7. G+C content was calculated over 40 bp lengths, at an interval of 20 bp. Positions ISH51 left and right, and average % G+C content of sequences flanking or between them are shown.

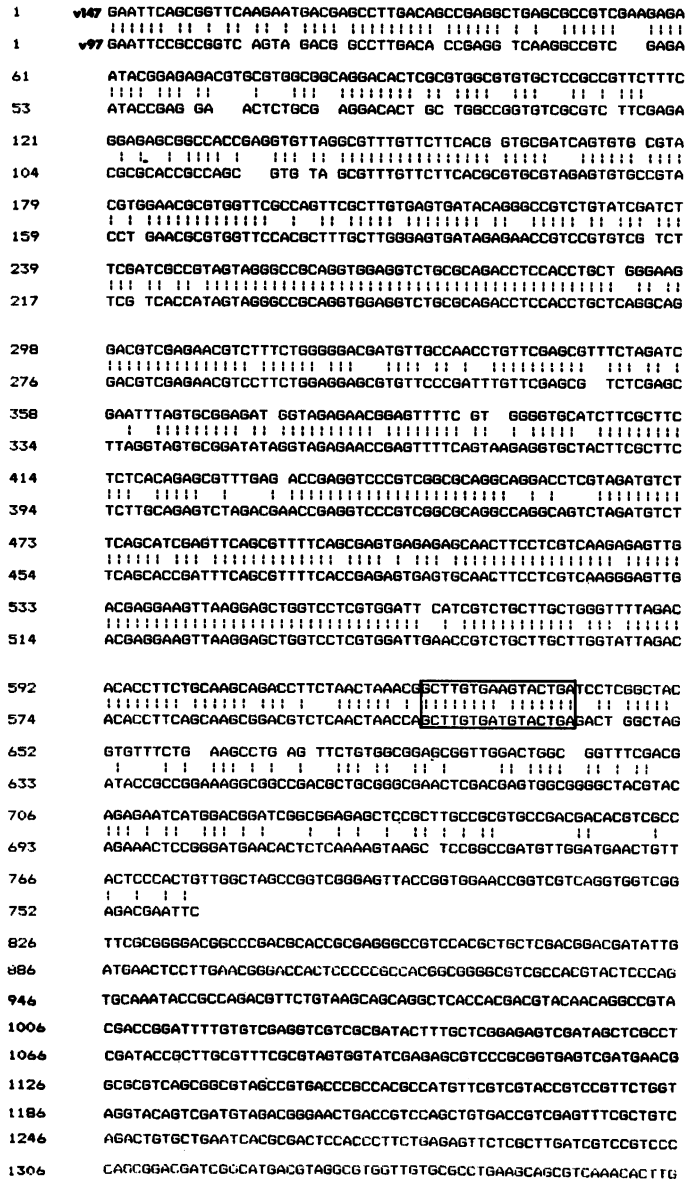


Fig. 5. Sequences of v147 and v97 aligned. The 16 bp sequences identified as terminal repeats in ISH51 left and right are boxed.

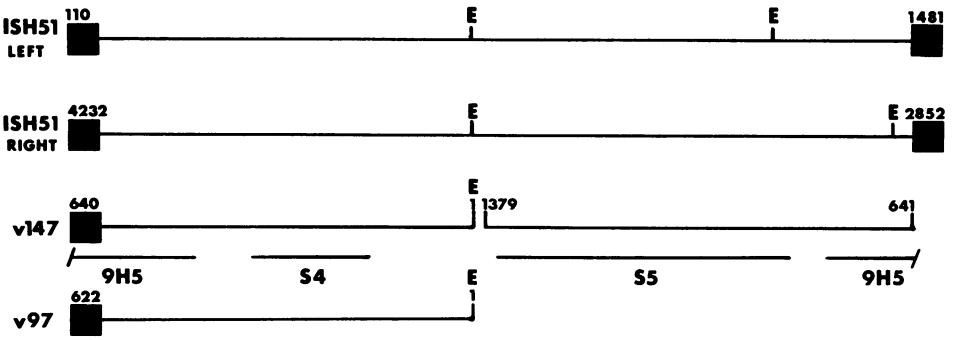


Fig. 6. Alignment of all sequenced ISH51 elements. Numbers refer to position numbers in Figs. 3 and 5. The v147 sequence is broken to indicate that this clone derives from Eco RI digestion of homologous restriction sites in two adjacent, tandemly repeated ISH51 elements. Positions within v147 of subclones S4, 9H5 and S5 are indicated by horizontal bars. Boxed regions represent the terminal inverted repeat sequences. Eco RI sites are indicated by E.

tent than the unique sequence DNAs flanking between them, although the entire region is nevertheless relatively AT rich.

Cloned fragments v97 and v147 were also sequenced on both strands (Fig. 5). Neither contained repetitive sequences which were not also contained within ISH51 left and right copies. Sequences represented are shown schematically in Fig. 6. Fragment v97 begins at an Eco RI site in what is presumably unique-sequence flanking DNA, and terminates at the Eco RI sites corresponding to position 720 in Fig. 3. (There is a conserved Eco RI site at this location in both left and right copies of ISH51.) Fragment v147 carries all but the last 18 bp of ISH51 -- that is, all but one copy of the terminal 16 bp repeat. Fragment v147 is not, however, colinear with ISH51 left or right. Its structure can be most easily interpreted as the result of cloning a fragment between homologous Eco RI sites (near position 720) in two adjacent tandem copies of ISH51, with loss (by some unknown mechanism) of one copy of the terminal 16 bp repeat. Pairwise sequence comparisons involving the ISH51 copies borne on λHvo7, v97 and v147 yield an average of 85.2% homology, with many insertions and deletions required to maintain alignments (Fig. 5). (Note that the conserved 16 bp stretches identified as terminal repeats in ISH51 are relatively well conserved.)

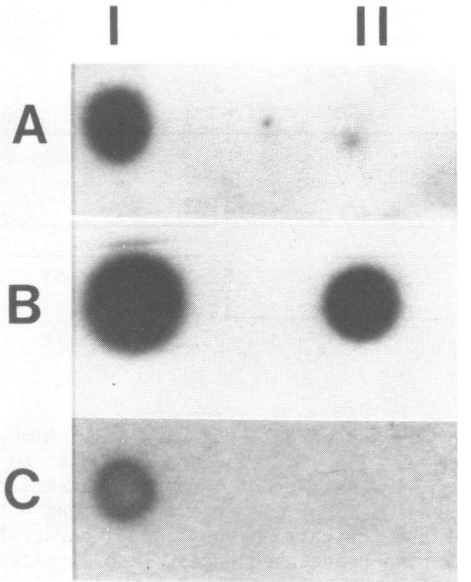


Fig. 7. Dot-blot hybridization to 68 mol % G+C component (fraction I) and 59 mol % G+C component (fraction II) of H. volcanii DNA, obtained by malachite green bisacrylamide gel column chromatography. Panel A: probed with total ^{32}P-pCp end-labelled H. volcanii transfer RNA. Panel B: probed with cloned fragment B of λHvo7. Panel C: probed with a cloned fragment containing H. volcanii 5S rRNA and $tRNA^{Cys}$ genes.

Interpretation of Southern Hybridization Results

The results shown in Fig. 1 and 2 can now be interpreted in light of the sequence information presented in Figs. 5 and 6. Subclone S4 of v14/ derives from the "left" half of the ISH51 sequence and subclone S5 derives from the "right" half (Fig. 6). The ISH51 elements described here contain an Eco RI site at a position corresponding to nucleotide 720 in Fig. 3, and this Eco RI site divides ISH51 roughly in half. S4 and S5 are thus expected to produce different patterns in Eco RI-digested genomic DNA. In a few cases, however S4 and S5 probe what appear to be identical genomic fragments (Fig. 1A). This might indicate that some ISH51 elements lack internal Eco RI sites.

In Fig. 2, all deletion fragments which retain both left and right halves of the ISH51 sequence produce the hybridization pattern obtained with v147 or v147 subclone 9H5, or a mixture of subclones S4 and S5. All deletion fragments which retain only the "left" half of one of the two inverted ISH51 sequences borne on λHvo7 subfragment B (Fig. 1B) produce the S4 hybridization

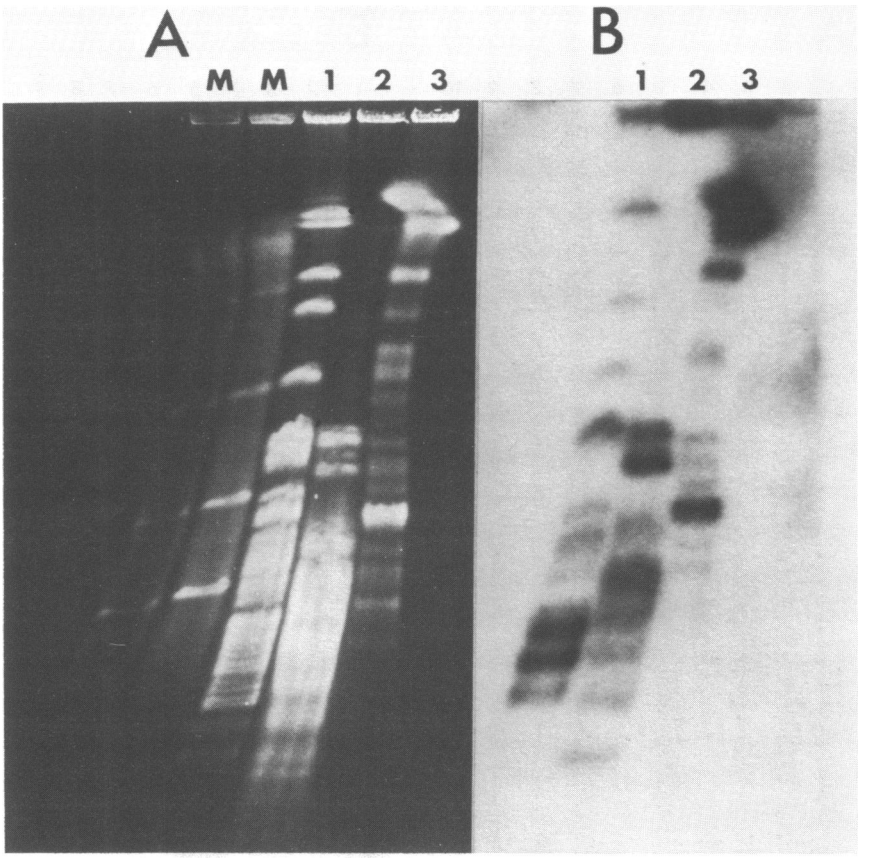


Fig. 8. A. Restriction digests of total H. volcanii DNA fractionated by pulsed-field electrophoresis. Lanes M: multimers of λC1857 Sam 7 DNA; lane 1: Bam HI digest, the smallest fragment is ca. 9 kb and the two largest fragments are 300 and 400 kb; track 2: Sfi I digest; track 3: Xba I digest. B. Southern hybridization with λHvo7 fragment B as probe.

pattern. From these results, it also appears unlikely that the 1369 bp stretch of DNA which separates the left and right full copies of ISH51 or λHvo7 itself contains repetitive sequences. This contention is borne out by hybridization experiments (not shown) using deletions retaining only this central region.

What is unexpected, on the basis of this interpretation, is the observation that subclones S4 and S5 of v147 probe different numbers of Eco RI fragments in genomic DNA (Fig. 1A). (Few, if any, of these fragments is of sufficiently low molecular weight to be derived by cleavage of Eco RI sites

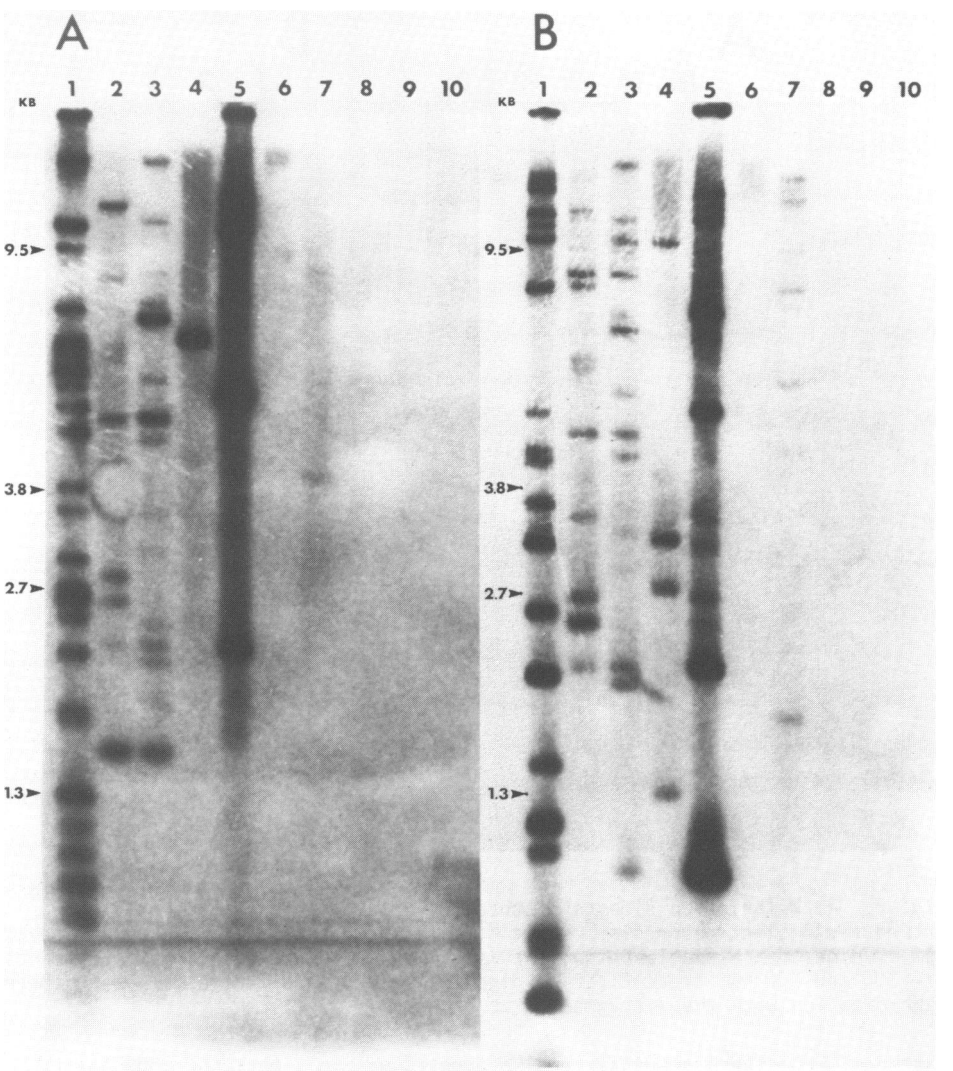


Fig. 9. Phylogenetic distribution of ISH51. Eco RI-digested total DNAs of 1, H. volcanii; 2, H. halobium strain NRC-1; 3, H. halobium strain R1; 4, H. marismortui; 5, H. cutirubrum; 6, H. saccharovorum; 7, Halococcus morrhuae; 8, Methanococcus voltae; 9, Methanococcus vannielli; 10, Methanococcus deltae. A. Probe is subclone S4 of v147. B. Probe is subclone S5 of v147.

internal to the ISH51 sequence, in copies which might contain multiple Eco RI sites.) It is difficult to avoid the conclusion that the H. volcanii genome contains, in addition to many heterogenous full-length ISH51 elements, some

truncated elements containing only sequences derived from the "left half" of ISH51.

Genomic Distribution

Pfeifer and collaborators (13,15) suggest that, in H. halobium, most insertion sequence elements are clustered within relatively AT-rich "islands" of genomic DNA, or are confined to AT-rich plasmids. To look further at possible clustering of ISH51 copies in H. volcanii the experiments illustrated in Figs. 7 and 8 were performed.

Fig. 7 shows the results of dot-blot hybridization to the 68 mol % G+C (fraction I) and the 59 mol % G+C (fraction II, "AT-rich") components of H. volcanii genomic DNA, prepared by malachite-green bisacrylamide column chromatography (6,23) of unsheared (>100 kb) material. The purity of these fractions was confirmed by hybridization with either total ^{32}P-pCp-end labelled transfer RNA (panel A) or nick-translated cloned DNA (24) bearing the 5S and $tRNA^{Cys}$ genes of H. volcanii (panel C). While neither of these two probes gave significant hybridization with the 59 mol % G+C ("AT-rich") fraction II DNA, cloned subfragment B of λHvo7 gave strong hybridization to both fractions. Thus, although some ISH51 copies are found in "AT-rich" fraction II DNA, many are embedded within fraction I (68 mol % G+C) DNA, or located in quite small AT-rich "islands" embedded in such DNA.

The technique of pulsed-field gel electrophoresis (25) makes it possible to separate very large fragments of H. volcanii DNA produced by restriction endonucleases which cut this ca. 4 x 10^3 kb genome fewer than 100 times. if there were but a few very large "islands" containing ISH51, then one would expect relatively few fragments in such a digest to give hybridization signals when probed with ISH51 sequences. Fig. 8A shows Bam HI, Sfi I and Xba I restriction endonuclease digests of H. volcanii DNA resolved by pulsed-field electrophoresis (26). Fig. 8B shows the result of Southern hybridization to a blot of this gel, with subfragment B of λHvo7 as probe. At least ten fragments of between 10 and 500 kb were hybridized, in the Bam HI, Sfi I, and Xba I digests, so that the 20-30 copies of ISH51 (or its truncated derivatives) must be located in a minimum of five widely separated regions of the H. volcanii genome.

Phylogenetic Distribution

Subclones S4 and S5 of v147 were used to probe Eco RI digests of genomic DNAs from several halobacterial strains, and three methanogenic archaebacteria (Fig. 9). All halobacteria, with the possible exception of H. saccharovorum, showed (usually several) hybridizing bands. No methanogens showed cross

hybridization. Patterns observed with S4 as probe were not identical to patterns observed with S5; presumably ISH51 elements in these other halobacterial species also often contain internal Eco RI sites.

DISCUSSION

There are 20-30 copies of the element ISH51 in the H. volcanii genome. ISH51 elements have insertion sequence-like structure, although we have not directly shown that they are mobile. Similar elements in H. halobium have been shown to be highly mobile, by both physical and genetic techniques (11,14).

The ISH51 family is unusual in several ways. Elements appear to lack open reading frames. There is considerable sequence divergence among family members, so that average homologies between the four ISH51 copies we have looked at is only 85%. Furthermore, at least some elements of ISH51 in the H. volcanii genome appear to be truncated, containing only a portion of the sequence we describe as full-length ISH51. We are not aware of any other archaebacterial or eubacterial insertion sequence family which shows such structural degeneracy, although restriction map comparisons between cloned H. halobium elements ISH23 and ISH50 may indicate that these too are part of a heterogeneous insertion sequence family (14,19). It is possible that many members of the ISH51 family are nontransposable. Alternatively, such elements may transpose by means of a transposase encoded by another element family, or by a few functional ISH51 copies.

Pfeifer and collaborators have proposed that in another halobacterial species, H. halobium, most insertion copies are clustered within a relatively few A+T rich "islands" on both plasmid and chromosomal DNA (13,15). We have presented evidence for clustering of some ISH51 elements in H. volcanii (two inverted copies in close proximity in λHvo7, and a presumptive tandem repeat giving rise to genomic fragment v147). There are, however, at least five and perhaps more than ten widely separated genomic regions in which the 20-30 ISH51 copies reside. (Some, but by no means the majority, of the hybridizing bands shown in Fig. 8 can be attributed to the 90 kb plasmid found in our strain.) Furthermore, ISH51 copies are not confined to the relatively AT-rich fraction II, "satellite" DNA fraction of unsheared, high-molecular-weight DNA, in spite of the fact that the two ISH51 copies on λHvo7 are embedded in AT-rich unique sequence DNA. We predict, therefore, that ISH51 copies will be found in many, rather scattered and presumably reasonably short (<20 kb) AT-

rich regions of the H. volcanii genome. Construction of a physical map of the genome, using pulsed-field gel electrophoresis, should allow us to determine whether or not this is so.

ACKNOWLEDGEMENTS

This work is supported by a grant from the Medical Research Council of Canada. L.C.S. is supported by a National Sciences and Engineering Research Council Studentship.

REFERENCES

1. Woese, C.R. and Fox, G.E. (1977). Proc. Natl. Acad. Sci. U.S.A. 74, 5088-5090.
2. Woese, C.R., Magrum, L.J. and Fox. G.E. (1978). J. Mol. Evol. 11, 245-252.
3. Woese, C.R. and Wolfe, R.S. (1985). The Archaebacteria, vol. 8 of The Bacteria, Academic Press, New York.
4. Garrett, R. (1986). Nature 318, 233.
5. Doolittle, W.F. (1985). Trends in Genetics 1, 268-269.
6. Weidinger, G., Klutz, G. and Goebel, W. (1979). Plasmid 2, 377-386.
7. Pfeifer, F., Weidinger, G. and Goebel, W. (1981). J. Bacteriol. 145, 375-381.
8. Moore, R.L. and McCarthy, B.J. (1969) J. Bacteriol. 99, 248-259.
9. Pfeifer, F., Ebert, K., Weidinger, G. and Goebel, W. (1982). In: Kandler, O. (ed.), Archaebacteria, Fisher, Stuttgart, pp. 110-119.
10. Pfeifer, F., Weidinger, G. and Goebel, W. (1981). J. Bacteriol. 145, 369-379.
11. Sapienza, C. and Doolittle, W.F. (1982). Nature 295, 384-389.
12. Sapienza, C., Rose, M.R. and Doolittle, W.F. (1982). Nature 299, 182-185.
13. Pfeifer, F., Betlach, M., Martienssen, R., Friedman, J., and Boyer, H.W. (1983). Mol. Gen. Genet. 191, 182-188.
14. Pfeifer, F., Friedman, J., Boyer, H.W. and Betlach, M. (1984). Nucl. Acids Res. 12, 2489-2497.
15. Pfeifer, F. and Betlach, M. (1985). Mol. Gen. Genet. 198, 449-455.
16. Betlach, M., Pfeifer, F., Friedman, J. and Boyer, H.W. (1983). Proc. Natl. Acad. Sci. U.S.A. 80, 1416-1420.
17. DasSarma, S., RajBhandary, U.L. and Khorana, H.W. (1983). Proc. Natl. Acad. Sci. U.S.A. 80, 1416-1420.
18. Simsek, M., DasSarma, S., RajBhandary, U.L. and Khorana, H.G. (1982). Proc. Natl. Acad. Sci. U.S.A. 79, 7286-7273.
19. Xu, W.-L. and Doolittle, W.F. (1983). Nucl. Acids Res. 11, 4195-4199.
20. Schnabel, H., Palm, P., Dick, K. and Grampp, B (1984). EMBO J. 3, 1717-1722.
21. Ebert, K., Goebel, W. and Pfeifer, F. (1984). Mol. Gen. Genet. 194, 91-97.
22. Messing, J. (1983) Methods Enzymol. 101, 20-78.
23. Daniels, C.J., Hofman, J.D., MacWilliam, J.G., Doolittle, W.F., Luehrsen, K.R and Fox, G.E. (1985) Mol. Gen. Genet. 198, 270-274.
24. Bunemann, H. and Muller, W. (1978) Nucl. Acids Res. 5, 1059-1074.
25. Schwartz, D.C. and Cantor, C.R. (1984) Cell 37, 67-75.
26. Carle, G.R. and Olson, M.V. (1984) Nucl. Acids Res. 12, 5647-5664.

27. Reed, K. and Mann, D. (1985) Nucl. Acids Res. 13, 7207-7221.
28. Karn, J., Brenner, S., Barnett, L. and Cesarini, G. (1980) Proc. Natl. Acad. Sci U.S.A. 77, 51-72-5176.
29. Henikoff, S. (1984) Gene 28, 351-359.
30. Yanish-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103-119.
31. Dale, R., McClure, B. and Houchins, J. (1985) Plasmid 13, 31-40.