GENETIC CHARACTERIZATION OF *TRICHINELLA SPIRALIS* ISOLATES BY ANALYSIS OF REPETITIVE DNA

by

Juanita Pearl Thiessen

a thesis presented to the

Faculty of Graduate Studies

University of Manitoba

In partial fulfillment of the

requirements for the degree

Master of Science

Department of Microbiology

May 1987 

ISBN 0-315-37440-3

GENETIC CHARACTERIZATION OF *TRICHINELLA SPIRALIS* ISOLATES BY ANALYSIS OF REPETITIVE DNA

BY

JUANITA PEARL THIESSEN

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

ACKNOWLEDGEMENTS

I would like to thank Drs. Glen Klassen and Terry Dick for the time they took to instruct and encourage me and for the financial assistance they provided. I also value their support of my future endeavors. I am gratefully indebted to my supervisor, Dr. Glen Klassen, for his assistance throughout my pre-master's and master's studies. I would also like to acknowledge Dr. Peter Loewen for his help with this thesis.

Thanks to Sally McNabb for her friendship, her patient teaching, and her ingenius solutions to constant research problems. The friendship and the technical assistance of Barb Triggs-Raine, Theo DeVos, and Dave Boyd are also recognized as invaluable to the success of this project.

A great big thanks to my parents, parents-in-law, and brothers and sisters. Every one of them has had their own way of being encouraging. A special thanks to Bev Loewen, a most inspiring friend. Mom, Dad and Bev insisted on understanding what I was doing and their fascination with the details of my research continually renewed my enthusiasm for the project.

Finally, a thank you to my husband Doug for his continuing support of my education. He has always demonstrated full faith in my abilities and reminded me to try again when something was unsuccessful.

Thanks for your help - here's to the future!

ABSTRACT

The restriction endonuclease profiles of 21 isolates of Trichinella spiralis were compared with respect to restriction fragment length differences in their repetitive DNA. Three strain-specific patterns were found. The P_1 strain includes pig and black bear isolates. The AF_1 strain contains 15 isolates forming 9 groups. Six of the groups are very similar while the remaining 4 sylvatic isolates are significantly different from this AF_1 core group. Comparison of host species and geographical location indicates that neither of these parameters has a great influence on the genetic similarities between 2 isolates. TP is the single isolate in the third strain. pPRA, a cloned member of the P_1 EcoRI 1.7 kb repetitive family, is minimally dispersed in direct tandem arrays and has a copy number of about 2800, representing 2% of the genome. Although a 1.9 kb variant of the sequence also occurs, the family is highly homogeneous. When the P_1 family members are probed with pPRA hybridization is identical in pattern and intensity with self-hybridization. No hybridization of pPRA to any of the AF_1 family isolates can be detected. However, faint hybridization of the probe to DNA from TP is observed. This indicates that the 1.7 kb sequence has been conserved in the course of strain evolution.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A	deoxyadenosine
bp	base pairs
BSA	bovine serum albumin
C	celsius
cm	centimeters
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotide phosphate
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
EDTA	ethylenediamine-tetra-acetic acid
EtBr	ethidium bromide
g	gram(s)
h	hour(s)
kb	kilobase pairs
L	liter(s)
M	molar
MCS	multiple cloning site

mg	milligram(s)
min	minute(s)
mL	milliliter(s)
mm	millimeter(s)
mM	millimolar
uCi	microcuries
ug	microgram(s)
uL	microliter(s)
N	normal
nm	nanometer(s)
O.D.	optical density
rDNA	ribosomal DNA
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
rRNA	ribosomal RNA
SDS	sodium dodecyl sulphate
sec	second(s)
T	deoxythymidine
Tris	tris(hydroxymethyl)aminomethane
V	volts

INTRODUCTION

INTRODUCTION

Isolates of Trichinella spiralis have been identified and compared with one another on the basis of morphology and infectivity and other types of phenotypic variation. However, this approach has left many questions regarding the global population structure of the parasite unanswered. It is vital to study these questions in order to gain a better understanding of the parasite's host range, distribution and speciation. These answers are also important in order to determine antigenic variability of the isolates so that the feasibility of vaccine development can be assessed.

An analysis of T. spiralis at the molecular level will aid in addressing these questions. More specifically, this problem will be approached by the analysis of the repetitive DNA fraction of the genome. Since repetitive DNA evolves rapidly, differences between two closely related isolates will be reflected in this genomic fraction. It is particularly useful to study T. spiralis isolates in this way since 42% of its genome consists of repetitive sequences (Searcey and MacInnis, 1970).

In this study genetic characterization will be done using two approaches. The repetitive bands in restriction profiles of various isolates will be compared, and a cloned highly-repetitive DNA sequence will be used as a strain-specific probe.

The results obtained can be used to group isolates into

strains. One can then attempt to correlate genetic identity with geographical location or host species. Finally, characterization of a strain-specific probe will enable the rapid identification of new isolates of Trichinella spiralis.

HISTORICAL

HISTORICAL

1. Repetitive DNA in eucaryotes

The discovery of repetitive DNA sequences in eukaryotic genomes marked the beginning of a continuing study into the complexity of these sequences. Approximately 30-40% of the average eukaryotic genome is made up of repetitive DNA (Britten and Kohne, 1968).

Eucaryotic DNA can be divided into three classes. Single copy DNA contains many protein coding sequences. Segments of DNA that are found again and again in the genome are referred to as repetitive DNA. Although these repeats are very similar to each other they need not be exact duplicates and will likely be variants of each other. The moderately repetitive DNA class is composed of sequences repeated up to 10^4 times. Included in this class are the genes coding for ribosomal RNA, transfer RNA, histones, actin, β-globin, and immunoglobulins (Lewin, 1982). Any sequence that is repeated more than 10^4 times is classified as highly-repetitive DNA (Singer, 1982).

Highly-repetitive DNA is further divided into two types based on the distribution of the sequence within the genome. Satellite DNA refers to those highly-repetitive sequences that are tandemly repeated while interspersed DNA is distributed between other DNA sequences. It should be noted that originally the term satellite DNA referred to DNA which could be separated from the bulk DNA by virtue of its unique

density in CsCl or $CsSO_4$ gradients during isopycnic centrifugation. It has since been demonstrated that this method does not necessarily separate out all of the tandemly-repeated DNA present in any given genome.

Satellite DNA has been found to make up anywhere from a few percent (human) to over 50% (Kangaroo rat) of mammalian genomes (Singer, 1982). A few common characteristics of satellite DNA include association with heterochromatin, lack of measurable transcription, replication late in S-phase, and underreplication in polytene chromosomes (Singer, 1982).

There has been wide speculation regarding the origin and evolution of the tandem repeat. Some form of amplification of these sequences must account for their high numbers. The extent of the similarities between satellite sequences of related species led to the library hypothesis (Fry and Salser, 1977). This theory states that a library of sequences is available to related species and that various members of this set are chosen to be amplified in the various species. There are several proposed mechanisms by which this amplification occurs. One of the most widely accepted proposals is that of random unequal crossing over (Hardman, 1986; Singer, 1982). When homologous or nonhomologous chromsomes pair there is seldom a perfect matching of the tandem arrays. The imperfect pairing and subsequent unequal crossing-over results in reciprocal amplification and deletion of sequences.

At present there are no conclusive findings regarding

the function of satellite DNA. Putative functions include involvement in chromosome pairing, control of gene expression, processing of messenger RNA precursors, and participation in DNA replication (Lewin, 1982). Bostock has stated that it is most probable that satellite DNA functions in germ line processes (Bostock, 1980). The association of satellite DNA with heterochromatin may provide more clues after further research.

There are two basic types of patterns of interspersed DNA (Singer, 1982). The sequences termed SINES (short interspersed repeated sequences) contain families with unit lengths under 500 base pairs and are present in hundreds of thousands of copies. The well characterized Alu family in the human genome is representative of a SINE. LINES, or long interspersed repeated sequences, are usually several kilobase pairs in length and are found up to 10^4 times in a genome.

Very little is known about how these interspersed sequences are amplified or about how they function. Gene conversion has been proposed as the mechanism of sequence amplification (Singer, 1982). This is a type of nonreciprocal recombination, whereby a DNA sequence is duplicated at a homologous site within the genome without being removed from the original site. It is feasible that some SINES are mobile or transposable sequence elements or that they may serve as origins for DNA replication (Hardman, 1986). Understanding the processes that maintain LINES in

smaller numbers of copies than SINES will also aid the understanding of the function of these sequences.

2. Repetitive DNA as a phylogenetic tool

It is very useful to study the repetitive DNA fraction of an organism's genome in order to better understand its phylogenetic relationship to other organisms. The theory of concerted evolution, proposed by Smith (Smith, 1973), is based on the idea that repetitive sequences will display more intra-specific homogeneity than inter-specific homogeneity. In other words, a sequence which is selectively amplified in a certain population will likely continue to evolve in concert in that population. That same sequence will probably change in a different way in a second population. Therefore the evolved forms of this sequence will soon be different between two distant populations. Over time these particular sequences may aid in establishing reproductive barriers thus playing a role in speciation. Conversely, once reproductive barriers have been established by any route, the repetitive sequences present in the two species will evolve in a diverging manner, thus further establishing the evolutionary distance between them. Several models have been formulated to describe how speciation may be the result of turnover processes involving repetitive DNA sequences (Rose and Doolittle, 1983).

Species which are closely related and which have recently diverged from a common ancestor will have more highly-repetitive DNA in common than species which are not closely related. Therefore the study of this fraction of an organism's genome will allow one to analyze its evolutionary

relationship to a second species. On an even larger scale, it is possible that the study of the organization of the repetitive DNA fraction of several genomes can provide information on which to build a phylogenetic tree (Flavell, 1982).

Classical taxonomy has relied on criteria such as morphological, biochemical and immunological characteristics. The use of the highly-repetitive DNA fraction as a taxonomic tool has several advantages over the more traditional methods (Curran et al, 1985). Most significantly it becomes possible to analyze the genome directly, thereby eliminating problems associated with phenotypic variation. Furthermore if restriction endonucleases are used it is possible to obtain a variety of characters by using many different enzymes. This approach has a broad use since it can be applied to any life-cycle stage of a species and is applicable to any organism. It is also of practical value in that analysis is rapid and extracted genomic DNA is stable for extended periods of time.

There are two basic methods that can be used to compare the highly-repetitive fractions of different genomes. In the first approach genomic DNA is digested with a restriction endonuclease and the fragments separated by gel electrophoresis. The highly-repetitive sequences will all be cut at the same position and therefore can be visualized as distinct bands on the gel. Differences in the spacing of

restriction sites in the repetitive fraction of a genome will appear as differences in the size distribution of these bands. In the second approach labelled probes of cloned highly-repetitive DNA fragments are made. These are then hybridized to Southern blots of genomic digests and the extent of sequence homology is visualized by autoradiography

Although the methods described above are relatively new this type of molecular analysis has been employed in a variety of cases. The phylogeny of cereal plant species (Flavell et al, 1977, 1979; Rimpau et al, 1978, 1980; Smith and Flavell, 1974), cetacean species (Arnason, 1982; Arnason et al, 1982; Arnason and Widegren, 1984; Arnason and Widegren, 1986; Widegren et al, 1985), deer species (Lima-de-Faria et al, 1984), and slime mold species (Richter and Ennis, 1985; Shaw et al, 1986) has been studied by the analysis of highly-repetitive DNA.

Flavell has studied the highly-repetitive DNA component of cereal plant genomes in order to construct the phylogenetic tree for this group of plants. Two types of experiments have been done. The first type involves renaturation of single strands from two species, while in the second type of experiment random fragments of repeated DNA are cloned, labelled and used as probes against restriction fragment digests that have been transferred onto nitrocellulose blots. These methods of analyses make it possible to determine the extent of similarity in the repetitive fraction of the two genomes.

Three different species of wheat were studied and it was found that essentially all of the highly-repetitive DNA present in Aegilops squarrosa is also present in Ae. speltoides and Triticum monoccum (Flavell et al, 1979). However, a very small proportion of the highly-repetitive DNA fraction in T. monoccum (0.5% of the total DNA) is not present in the other two species. A larger proportion of the Ae. speltoides genome (2-3% of the total DNA) consists of highly-repetitive DNA that is not present in the Ae. squarrosa and T. monoccum species. Since these differences are relatively minor it is apparent that the majority of the highly-repetitive fraction is common to all three species and was probably present in the common ancestral species. Studies with several repetitive DNA probes from the wheat genome indicate that although these sequences belong to common families, there are differences in copy number and sequence organization (Flavell et al, 1979).

Comparisons between the repetitive fractions of more distantly related cereal plants were also made. Renaturation experiments were carried out between wheat, rye (Secale cereale), barley (Hordeum vulgare), and oats (Avena sativum). Species-specific repeated sequences account for 16%,22%, 28% and 58% of the respective genomes (Rimpau et al, 1978, 1980). It appears that most of the repetitive DNA fraction in the oats genome is unrelated to that of the other three species. These experiments have led to the conclusion that wheat and rye are closer phylogenetically

than wheat or rye and barley. Analysis of sequences that are present in wheat, rye, and barley indicates that here too sequence structure and organizational differences exist. Studies on thermal stabilities of reannealed repeated sequences point to the existence of many subfamilies within species (Smith and Flavell, 1974). All these data have been used to confirm the cereal phylogenetic tree that had been determined using classical taxonomic methods.

Arnason has used highly-repetitive DNA components to study mysticete/odontocete phylogeny and pinniped phylogeny. The primary tool is a cloned highly-repetitive fragment which is used as a probe against the genomic DNA of a variety of related species.

All whales, dolphins and porpoises are classified as cetaceans. The odontocete (toothed whale) and mysticete (whalebone whale) lineages are thought to have separated about 40-50 million years ago. There is still some disagreement as to whether the group should be considered mono- or diphyletic.

RNAs complementary to two unrelated DNA satellites of balenopterid (mysticete) species were hybridized to the DNA of several cetacean species (Arnason et al, 1982). Sequences related to the light satellite were not present in the odontocete species. However, sequences related to the heavy satellite were found in both odontocete and mysticete genomes. These satellites were also used in intrageneric cRNA/DNA hybrids in order to obtain thermal melting data.

The results obtained confirmed the hybridization findings. In a more recent study the satellites themselves were isolated and hybridized to mysticete and odontocete species with similar results (Arnason and Widegren, 1984). The light satellite did not hybridize outside the genus Balaenoptera while the heavy satellite was found in mysticete and odontocete species.

Several highly-repetitive DNA fragments have been used as probes to compare odontocete and mysticete genomes. A fragment from killer whale was hybridized to several other odontocete species and two mysticete species. Hybridization to the three odontocete DNAs was evident but was limited or nonexistent in the mysticete ones (Arnason 1982). Conversely,a 1740-bp balenopterid fragment was found to be highly conserved in all cetaceans (Arnason and Widegren, 1984). It is thought that a 1579-bp component isolated from killer whale and found in only one family of cetaceans has evolved from this 1740-bp cetacean fragment (Widegren et al, 1985). All of the data obtained is consistent with a monophyletic origin of odontocetes and mysticetes.

The pinniped group includes true seals, walrus, and sea lions. Otters, ferrets, minks, badgers and skunks are classified as mustelids. Arnason's research team used four cloned highly-repetitive DNA fragments from the Weddell seal to study pinniped and mustelid phylogeny (Arnason and Widegren, 1986). Each of these four components was identical in all pinnipeds and three of the four also

hybridized to corresponding fragments in most mustelids. The existence of a pinniped-specific component suggests that the pinnipeds are monophyletic, separating from the mustelids as one lineage that later evolved into the various species. The hybridization results also confirmed that among the terrestrial carnivores the mustelids are the closest relatives to the pinnipeds. Although the skunk is classified as a mustelid its DNA did not hybridize to the pinniped probes as the other mustelids did, thus raising a question about its true relationship to this group.

Repetitive DNA sequences have also been used to study the Cervidae family, which contains the deer species. This family is unique in that, among the mammals, it exhibits the largest variation in chromosome number (Lima-de-Faria et al, 1984). Analysis of repetitive DNA sequences has also found similarities and differences between these species. Very similar band patterns were obtained in five species of Cervidae after digestion with four different restriction enzymes. When a repetitive DNA sequence from reindeer DNA was cloned and used as a probe it was found that homologies exist between reindeer and all the other species. However, a greater degree of homology was found between the reindeer, elk and roe deer than between the reindeer and fallow deer or muntjac. These results agree with previous genetic analysis and with the geographic distribution of the deer (Lima-de-Faria et al, 1984).

Similar studies have been carried out using repetitive

DNA sequences in the slime mold, Dictyostelium discoideum. Two different repetitive sequences were cloned and hybridized to genomic digests of three other members of the same taxonomic family (Richter and Ennis, 1985; Shaw et al, 1986). Both sequences were found in all Dictyostelium species but not in Polysphondylium, a closely related species. These results are consistent with the taxonomic classification based on the morphological and biochemical properties of these cellular slime molds.

3. Repetitive DNA in helminths

3.1 Freeliving helminths

The most well-studied freeliving helminth is the small nematode, Caenorhabditis elegans. This worm's genome contains 8×10^7 base pairs, which is only about 20 times the amount found in the E. coli genome. Only a few thousand genes are encoded. Cot curves based on renaturation studies indicated that 83% of the DNA sequences are unique and that 17% are repetitive (Sulston and Brenner, 1974).

The arrangement of repeated sequences in the C. elegans genome has been further studied by electron microscopy and reassociation kinetics (Emmons et al, 1980). It was found that most of the repetitive sequences are only a few hundred base-pairs long and that they are highly interspersed throughout the genome. While there is no distinct class of moderately repetitive DNA, there are many sequences which are repeated less than 100 times. Based on the electron microscopic analysis of renatured DNA, it was tentatively concluded that the satellite sequences represent no more than a few percent of the total repetitive fraction. However, it has been noted that satellite sequences can be preferentially lost during phenol extraction (Skinner and Triplett, 1967) or during chromatin diminution (Moritz and Roth, 1976).

Characterization of the genes coding for the ribosomal RNAs in C. elegans and related species showed that restriction cleavage patterns of the rDNA could be used to

differentiate between several strains and species of Caenorhabditis. It was first demonstrated that the rDNA has been highly conserved between two strains of C. elegans, Bristol (N2) and Bergerac (Bo). These worms were isolated in different locations and 10 years apart. After hybridizing 13 probes to a variety of digests, only 5 out of a total of 37 bands showed size differences between the two strains (Files and Hirsh, 1981). More recent studies have shown further restriction fragment length differences in the 5S rRNA genes of the two strains (Nelson and Honda, 1985).

Comparisons of the rDNA of C. elegans and C. briggsae reveal differences (Files and Hirsh, 1981). Heteroduplex analysis demonstrated that most of the divergence is in the non-transcribed spacer. Comparison of the restriction maps of the rDNA show that there are several restriction sites in each species which are not found in the other. The majority of restriction fragments in the coding regions are the same in the two species. However, the restriction maps show no fragments in common in the non-coding regions.

Restriction fragment length differences in two C. elegans strains, Bristol (N2) and Bergerac (Bo), have been used to map these genomes (Rose et al, 1982). Randomly cloned fragments were used to probe genomic restriction digests from hybrid populations. Six restriction fragment differences were found by the use of 27 probes. The linkage mapping of these differences is providing a better way to characterize these genomes.

Curran et al (Curran et al, 1985), have done some initial work on the highly-repetitive DNA fraction of Caenorabditis by comparing restriction fragment length differences in several strains and species after digestion with one restriction enzyme. Their results demonstrate that three different strains of C. elegans have similar banding patterns while C. briggsae, a distinct species, has several different major bands. Similar results were obtained with several species of the nematodes Romanomermis, Steinernema and Meliodogyne. In all cases restriction fragment length differences were detected between species but not between strains or populations. These results point to a correlation between the degree of restriction fragment length differences in repetitive DNA and the potential for two populations to interbreed.

Initial studies on Panagrellus silusiae have shown that its genome is similar in size to C. elegans and that about 26.1% of it consists of repetitive DNA. The repeated sequences are clustered in the genome with segments extending up to 10,000 nucleotide pairs (Beauchamp et al, 1979).

3.2 Parasitic helminths

3.2.a. Ascaris/Parascaris

Studies of chromatin diminution in Ascaris have provided some interesting information about the distribution of the highly repetitive DNA in this genome (Streeck et al, 1982; Tobler et al, 1985). About 27% of the total germ line genome is eliminated in the soma cells. Reassociation kinetic experiments have shown that the majority of these eliminated sequences are highly repetitive sequences (Tobler et al, 1985). Although only a very small percentage of repetitive sequences are retained in somatic cells it appears that these sequences are distinct from the eliminated ones.

Additional research has shown that the highly-repetitive DNA that is eliminated is satellite DNA (Roth and Moritz, 1981). This germ line limited satellite DNA is composed entirely of two related families. These tandemly arranged AT-rich variants are 121 bp long and differ in about 20% of their base sequence. The fact that satellite DNA is not transcribed and that it is germ line limited may provide further clues about its function (Tobler et al, 1985).

Work on Parascaris equorum has shown that about 85% of the DNA is eliminated during chromatin diminution (Teschke, 1985). The germ line limited DNA is composed of two satellites (Moritz and Roth, 1981). Peculiar results were obtained when the satellite DNA was digested with a variety of restriction endonucleases. Homogeneous smears with no

distinct bands were observed on all the gels (Teschke, 1985). Since this DNA is tandemly repeated it was concluded that these smears were only possible if the satellite DNA was based on a core sequence of about 20 bases which was free of any recognition sites of the enzymes used. The satellite DNA could have been obtained after repeated amplification of the basic sequence. Random deletions could provide occasional recognition sites for some of the enzymes.

3.2.b. Brugia

Because there are a variety of tropical diseases caused by filarial nematodes it is crucial that accurate identification of the parasite is possible. However all the methods currently in use are insensitive or time consuming. The characterization of a member of a DNA repeated-sequence family from Brugia malayi is proving helpful in distinguishing several filarial nematodes (McReynolds et al, 1986). There are approximately 30,000 copies of this tandemly arranged 320-base-pair sequence. This represents about 12% of the total genome.

The cloned repeated DNA sequence is a sensitive probe that can detect DNA isolated from a single parasite in an aliquot of blood. The sequence is also a valuable probe since it does not hybridize to the parasites, Dirofilaria immitis, Dipetalonema viteae, Litomosoides carinii, or Onchocerca volvulus. This would imply that this repeated sequence is the result of a large-scale amplification that

took place after Brugia diverged from the other filarial parasites.

Cross-hybridization between the clone from B. malayi and a sequence in B. pahangi indicates that they are partially homologous. Sequence analysis shows that there are some differences in restriction sites. These minor differences will allow for the construction of important species specific hybridization probes.

3.2.c. Schistosomes

Schistosomes are parasitic trematodes that cause a variety of human diseases. Although intraspecies diversity may play an important role in characterizing the disease, few methods exist for distinguishing parasites within a species. The use of cloned DNA markers is proving to be helpful in diffferentiating schistosomes by species, strain and sex. A segment of the S. mansoni ribosomal gene has been used as a probe and has been found to be present in all 8 strains probed (McCutchan et al, 1984). However the copy number or the surrounding DNA sequence is variable. These differences can be used to distinguish and identify the various strains. Differences in the length of the major repeating unit of the ribosomal gene can also be used to distinguish between S. haematobium and S. japonicum. A variety of minor bands has been detected in each species. These may provide a future tool for differentiating the strains within each species.

3.2.d. Trichinella

Although many Trichinella isolates have been studied, the question of speciation is still unresolved. Since approximately 42% of the T. spiralis genome consists of repetitive sequences (Searcy and MacInnis, 1970) recent studies have used this portion of the genome to attempt to solve this dilemma.

Restriction fragment length differences were observed between two Trichinella species (T. spiralis, T. spiralis var. pseudospiralis) after digestion with a single restriction endonuclease (Curran et al, 1985). These differences reflect differences in the highly-repetitive component of the genome. Although it has been proposed that T.spiralis and T. pseudospiralis are distinct species, the data is not conclusive (Dick, 1983a). The differences recorded in this study provide further evidence of a degree of reproductive isolation, but incomplete speciation.

A more complete analysis of repetitive sequences has been carried out with 6 Trichinella isolates and 12 restriction endonucleases (Klassen et al, 1986a). Restriction profiles were obtained with 12 endonucleases on a pig isolate, an arctic fox isolate, and a T. spiralis var. pseudospiralis isolate. With the exception of two major bands, a unique pattern of bands was observed for each isolate. Identical profiles were obtained for two pig isolates using four enzymes. Isolates from an arctic fox, a wolverine and a polar bear were also compared using four

endonucleases. Again there were only two departures from identity in the profiles. In both exceptions the wolverine isolate was different from the other two. This implies that there may have been an isolating event occurring in northern Canada very recently.

A 1.7 kb repetitive DNA sequence in T. spiralis has been used to address the speciation question (Klassen et al, 1986b). This tandemly-repeated sequence cloned from P_1, an isolate from domestic pig, has a copy number of about 2800, representing 2% of the genome. On the basis of hybridization patterns this family is present in the repetitive fraction of two other pig isolates and two black bear isolates from Pennsylvania. Further studies will provide information about the transmission of the nematode between these two hosts. No hybridization was detected with isolates from martin, pen fox, arctic fox, wolverine or polar bear. Therefore this sequence is either absent from these isolates or present in very low copy numbers. However, faint hybridization to the T. spiralis var. pseudospiralis was observed. The difference in degree of hybridization intensity indicates that the sequence has been amplified in the pig strain after divergence of the two strains. This result is consistent with the library hypothesis for the generation of satellite DNA families (Fry and Salser, 1977).

Not only do restriction profiles provide a way of distinguishing Trichinella strains but some repetitive

families such as the 1.7 kb one studied may provide strain specific probes. This should help to answer the question of speciation in Trichinella.

4. The speciation problem in Trichinella

The genus Trichinella belongs to the family Trichinellidae, superfamily Trichuroidea, order Anoplida, class Aphasmidia, and phylum Nemathelminthes.

Although some researchers feel that Trichinella spiralis is the only well-defined species it has been proposed that there are up to 4 different ones, i.e. T. spiralis, T. nativa (Britov and Boev, 1972), T. nelsoni (Britov and Boev, 1972), and T. pseudospiralis (Garkavi, 1972). These species distinctions were based on morphological differences that included adult, larvae, and capsule length, the texture of the external surface and the presence or absence of a cyst (Boev et al, 1979). Differences were also noted in the cuticle of each species and in the resistance to low and high temperatures. Evidence of the failure to interbreed or produce fertile offspring was also presented. Because so little is known about the degree of intraspecific variation in Trichinella, the data accumulated thus far does not appear to be conclusive enough to justify the establishment of four different species (Dick 1983a). Therefore it has been suggested that the various worms be referred to as geographic isolates until more conclusive data has been gathered.

Worms can be distinguished from each other on the basis of a variety of genetic, morphological, biochemical, and immunological characteristics (reviewed by Dick et al, 1984;

Dick, 1983a; Dick, 1983b). The genetic features that have been studied include the karyotype and breeding potential. All female Trichinella analyzed have six chromsomes and all males have five (Penkova and Romanenko, 1973). Slight differences in the chromsome shape, length, and centromere index have been detected between some isolates (Mutafova and Komandarev, 1976). Since reproductive isolation is so important in the designation of species it would seem that the results of breeding experiments should provide the most conclusive distinctions. However, the method of infection, the use of single or multiple breeding pairs, and the criteria for sexing Trichinella larvae differ in the various laboratories. Therefore the data obtained is not consistent and is not complete.

As mentioned above, morphological criteria include the length and width of the larvae, the adults, and the cysts, the appearance of the external surface of the worm, and the configuration of some of the worm's external anatomy. Although some differences in adult length have been reported, it has also been observed that there is a decline in size with passage through experimental hosts (Chadee and Dick, unpublished). Some differences in the copulatory appendages of various isolates have also been cited (Hulinska and Saikenov, 1980; Dick, unpublished data). However none of these differences are pronounced enough to support species separation on morphological grounds (Dick, 1983a).

Starch gel electrophoresis work with 11 enzyme systems has shown that seven isolates can be segregated into 3 basic patterns (Mydynski and Dick, 1985). Enzyme polymorphisms were found between the isolates from pigs, wild carnivores and T. spiralis var pseudospiralis. In addition minor differences were observed between genotypes within wild carnivores.

Although minor differences in the amino acid composition of a few isolates has been observed (Dick, 1983a), no consistent biochemical differences have been found for any of the comparisons. Similarily, a variety of immunological results have not provided conclusive findings. Methods used include blood serum fraction analysis, immunodffusion, immunoeletrophoresis, in vitro microprecipitation, cross-absorption of antisera, and indirect immunofluorescence.

Sensitivity to the drug, thiabendazole, varies among isolates of Trichinella. In one recent study the polar bear and wolverine isolates and T. spiralis var. pseudospiralis were the most sensitive (Chadee et al, 1983). The pig isolate was significantly less sensitive to the drug. However, it has been observed that these differences may partially be due to the length of time that an isolate has been passaged in experimental hosts (Dick, 1983a).

The degree of infectivity as measured by the reproductive-capactiy index (RCI), is influenced by the hosts. Problems arise when an isolate has been passaged

through a variety of hosts before analysis. It has also been observed that the host can influence the level of infection in subsequent infections (Dick and Chadee, 1981). If an isolate is maintained in one host the RCI will be relatively stable (Belosevic and Dick, 1979). However one isolate may behave quite differently in different hosts.

Other characteristics that have been used to compare isolates include cyst morphology, resistance to freezing, virulence of isolates, intestinal distribution, in vitro larval release by 7-day-old females over a 24-hr period, duration of the intestinal stage, and viability of muscle larvae.

The four proposed species fall into approximate geographic zones. The arctic isolate or T. nativa, is relatively resistant to freezing, has an opaque appearance and has a low infectivity in experimental hosts. The temperate-to-north-temperate form is associated with pigs and rats and has a high infectivity and low resistance to freezing. These isolates are referred to as T. spiralis. The southern species, T. nelsoni, is poorly defined and shares several characteristics with some holarctic forms.

The isolate which is most likely to be a separate species is T. pseudospiralis. It has been obtained from the U.S.S.R. although it may also be present in birds in North America. It is characterized by its small size, infectivity to birds and lack of cyst in the muscle stage.

Since geographic isolation is one of the major

isolating mechanisms in the process of speciation, it is possible that the geographic grouping of isolates may be valid. However, the geographic zones in which Trichinella are found overlap and some of the host animals are certainly capable of moving through more than one region. Therefore although some of the biological differences reported may be a result of this separation, speciation is likely not complete. It would seem appropriate then to conclude that there is only one true species, T. spiralis. While T. spiralis var pseudospiralis may eventually progress to recognition as a separate species the minor differences in the other isolates will likely be absorbed into the natural variation of T. spiralis (Dick, 1983a, Dick, 1983b).

The information that has just been presented does not include any molecular data. Yet it would appear that an analysis at this level could help establish species delineation. It is evident that many of the isolates are in a state of change where they may become more or less unique. These modifications will require an alteration in the DNA and, as proposed in Section 1, these changes may be reflected in the highly-repetitive DNA fraction. Therefore the study of this portion of the genome may provide a molecular basis for the grouping of Trichinella isolates into strains and species. By following the changes that take place over the years it may even be possible to trace the evolution of various isolates.

MATERIALS AND METHODS

MATERIALS AND METHODS

1. Buffers, media and solutions

10X Borate buffer - 0.089 M Tris-HCL, pH 8.3, 0.089 M boric acid, 2.5 mM EDTA.

Chloroamphenicol - 34 mg/mL of 100% ethanol, stored at -20°C.

Chloroform:isoamyl-alcohol - 24 volumes chloroform and 1 volume isoamyl-alcohol.

Denaturing solution - 0.4 N NaOH, 0.6 M NaCl.

DNase I solution - 0.1 ug/mL DNase I in nick translation buffer containing 50% glycerol, stored at -20°C.

1X Denhardt's solution - 0.2% Ficoll, 0.2% BSA, 0.2% polyvinylpyrrolidone in distilled water, stored at -20°C.

GET solution - 25 mM Tris-HCL, 20 mM EDTA, 50 mM glucose.

Hybridization buffer - 6X SSC, 0.5% SDS, 5X Denhardt's solution, 0.1 M EDTA, 100 ug/mL denatured salmon sperm DNA.

Hybridization wash solutions - #1 - 2X SSC.
#2 - 2X SSC, 1% SDS.
#3 - 0.1X SSC.

LB medium - 10.0 g tryptone, 5.0 g yeast extract, 5.0 g NaCl, per L of distilled water. 10.0 g of agar was added for plates and 7.0 g of agar for top agar.

10X ligase buffer - 500 mM Tris-HCL, pH 8.0, 70 mM MgCl, 10 mM DTT.

NaOH/SDS solution - 200 mM NaOH, 1% SDS.

Neutralizing solution - 0.5 M Tris-HCL, pH 8.0, 1.5 M NaCl.

10X Nick translation buffer - 100 uM dCTP, 100 uM dGTP, 100 uM dTTP, 250 ug/mL BSA, 0.05 M $MgCl_2$, 0.005 M DTT, 0.25 M Tris-HCL, pH 7.8, stored at $-20^{o}C$.

Pepsin solution - 1% pepsin in 1% HCL.

Phenol - Freshly distilled phenol was saturated with TE buffer and stored in foil-covered bottles at $-20^{o}C$.

Phenol washing buffer - 10 mM Tris-HCL, pH 8.0, 1mM EDTA.

Prehybridization buffer - 6X SSC, 0.5% SDS, 5X Denhardt's solution, 100 ug/mL denatured salmon sperm DNA.

Proteinase K solution - 2 mg proteinase K in 1 mL of buffer. The buffer was 100 mM Tris-HCL, pH 8.0, 50 mM EDTA, 200 mM NaCl, 1% SDS.

RNase - 2.0 mg/mL RNase A in proteinase K buffer, stored at $-20^{o}C$.

Saline - .85% NaCl.

SM - 100 mM NaCl, 8mM $MgSO_4$ $7H_2O$, 50 mM Tris-HCL, pH 7.5, 0.01% gelatin.

1X SSC - 0.15 M NaCl, 0.015 M sodium citrate.

Sucrose solutions - These were made in a buffer containing 1 M NaCl, 20 mM Tris-HCL, pH 8.0, 5 mM EDTA.

TAES solution - 50 mM Tris-HCL, 0.1% SDS, 1 mM EDTA,

100 mM sodium acetate.

TE buffer - 10 mM Tris-HCL, pH 8.5, 1 mM EDTA.

Water - All the water used was distilled and deionized.

2. Trichinella DNA isolation

2.1 Trichinella isolation from mice

The Trichinella spiralis isolates were maintained in outbred Swiss Webster mice by Dr. T.A. Dick, University of Manitoba, Winnipeg (Mydynski and Dick, 1985).

Approximately 40-60 days after infection 12-24 mice were killed, skinned and eviscerated. Each animal was homogenized in a Braun blender for a few seconds with 100 mL of the pepsin solution. This mixture was poured into a 250 mL Erlenmeyer flask and shaken at 37^{o}C for 2 h. After this digestion the flasks were allowed to stand for 15 min before the fat layer on the top of each flask was removed with a water aspirator.

The mixture was then poured through two interlocking sieves. The #80 at the top trapped large pieces of debris and the #250 at the bottom caught the worms on its surface. The sieves were washed with one water rinse and one saline rinse before the worms were washed into a test tube with about 20 mL of saline. Ten min later, when the larvae had settled to the bottom of the test tube, about 18 mL of the saline was removed by water aspirator. These worms were washed five times with saline and three times with water. The larvae were stored in water in 2 mL Cryo vials at a concentration of about 500,000 worms/mL. The vials were then frozen in liquid nitrogen until use.

2.2 DNA isolation

The worms were digested at 65^{o}C for 30 min with an

equal volume of proteinase K solution. The mixture was divided into 500 uL aliquots in 1.5 mL microcentrifuge tubes. The phenol was washed 3 times with the phenol wash buffer before being used. The proteinase K mixture was then extracted with the phenol for 30 min at room temperature 3 times. One 30 min chloroform:isoamyl-alcohol extraction was done at room temperature. For each 30 min incubation the tubes were inverted every 3 min. The DNA was then precipitated out of solution by adding 1 mL of ethanol to each tube and pouring this into a dish of ethanol on a bed of ice. At this stage the DNA aliquots could be pooled according to their quality. Good quality DNA was very viscous and remained in one piece. The poor quality DNA was thin and appeared broken up into many fragments.

After a 30 min precipitation at -20^{o}C the tubes were spun down in a microcentrifuge for 5 min and the ethanol decanted. The pellet was air dried for 30 min and then resuspended in 100 uL of TE buffer. 1 uL of RNAse A solution was added and incubated at room temperature for 30 min. An extraction was done with an equal volume of chloroform:isoamyl-alcohol. A final ethanol precipitation was carried out with 2 volumes of ethanol. After being spun down for 30 min the pellets were dried and resuspended in 100 uL of TE buffer. Finally, the mixture was again spun down for about 15 min and the supernatant transferred to a new microcentrifuge tube for further use.

3. Restriction enzyme digestion and gel electrophoresis

Restriction enzymes were purchased from Bethesda Research Laboratories, Boehringer Mannheim Canada or Pharmacia.

Digestion mixtures were prepared in 1.5 mL microcentrifuge tubes using 0.5-1.0 ug of Trichinella DNA and about 5 units of enzyme in buffers recommended by the supplier of the enzyme. In order to ensure complete digestion an additional 5 units of enzyme was added after several hours of incubation. The reaction was continued for another half hour before being stopped. Although the total time of incubation varied from 2-4 h it was maintained at a constant time for any series of comparative gels. Series of partial digestions were carried out with constant amounts of enzyme and varying lengths of time. Digestions were stopped by the addition of 0.3 volumes of a solution containing 40% sucrose, 120 mM EDTA and 0.1% bromophenol blue.

The gels used for electrophoresis were 0.5-2.0% agarose and were poured as 15 x 16 cm horizontal slabs. They were run at 30 V for 15-20 h or at 85 V for 4 h. During electrophoresis they were submerged in 1X borate buffer. Gels were stained either by running them in buffer containing 1 ug/mL of ethidium bromide or by submerging them in this staining mixture for 20 min after electrophoresis.

Photographs of gels were taken by placing the gel directly onto an ultraviolet transilluminator. Instant photographs were obtained using Polaroid type 667 film and

negatives were made with Kodak type 4147 Plus-X pan film. Both types of prints were obtained using a Wratten 22A filter.

In order to provide the appropriate standard size markers, restriction digest fragments of bacteriophage lambda DNA or the Bethesda Research Laboratories (BRL) 1 kb ladder were run on each gel. Fragment sizes were estimated by graphical comparison to standards.

4. Library construction

4.1 Partial digest of DNA

The preparation of partially digested DNA was as described below (Maniatis, et al, 1982).

About 100 ug of DNA were digested with 0.01 units of the restriction enzyme Sau3A1 (see Methods and Materials section 3). After a 1 h incubation at 37°C this reaction was stopped by cooling the mixture to 0°C for 10 min and by adding EDTA to a concentration of 20 mM EDTA. An aliquot of about 1 ug of DNA was run on a 0.4% agarose gel to confirm that the desired sizes of fragments had been obtained in the partial digest. The DNA was extracted twice with a phenol: chloroform:isoamyl-alcohol (25:24:1) mixture and then precipitated out with 95% ethanol. After a wash with 70% ethanol it was redissolved in 70 uL of TE buffer.

A 10-40% sucrose density gradient was prepared in a 5 mL polyallomer tube. After the DNA sample was heated at 68°C for 10 min and then cooled to 20°C it was loaded onto the gradient. The sample was spun at 32,000 rpm for 16 h at 20°C in a Beckman swinging bucket rotor. Starting at the bottom of the centrifuge tube, 70-100 uL aliquots were collected from the tube using a Beckman fraction recovery system. Ten uL of every third fraction was run on a 0.4% agarose gel to determine which aliquots contained 13-20 kb fragments. The standard markers run on the gel were brought to a concentration of about 20% sucrose. The DNA from the desired aliquots was then pooled and diluted to about 10%

sucrose by the addition of TE buffer. Another ethanol precipitation was done and the DNA was resuspended in TE buffer to a concentration of about 0.2 ug/uL.

4.2 Ligation of inserts to arms

The ligation reaction was carried out using the conditions recommended by the EMBL3 arms manufacturer, Stratagene.

One ug of EMBL3 arms, 0.5 ug of insert DNA, 0.6 uL 10X ligase buffer, 0.6 uL 10mM ATP, pH 7.5, 200 units T4 DNA ligase, and water up to 5 uL final volume were combined in a 1.5 mL microcentrifuge tube. The ligation proceeded at 4^{o}C for 16 h. An aliquot of 0.5 uL was electrophoresed to ensure that the inserts were ligated to the arms.

4.3 Packaging and plating of the library

The instructions provided by Amersham, the manufacturer of the DNA *in vitro* packaging kit, were followed.

The remaining 10 uL of ligation mixture was added to a 10 uL aliquot of an extract prepared from *E. coli* strain BHB2688. Fifteen uL of an extract prepared from *E. coli* strain BHB2690 was added and the entire mixture incubated at room temperature for 90-120 min. Five hundred uL of SM and 10 uL of chloroform were added. The phage stock was then stored at 4^{o}C for further use.

A 6 h culture of P2392 was grown up in LB with 10 mM $MgSO_4$ and 0.2% maltose. Ten uL of diluted phage were combined with 100 uL of P2392 to a final concentration of about 5,000 plaques/plate. They were then incubated at 37^{o}C

for 15 min. Three mL of top agar containing 10 mM $MgSO_4$ was added to the cells and the mixture poured onto LB plates. The plates were incubated at 37°C for 16 h.

5. Large scale plasmid isolation

The large scale plasmid isolation was done by the method described below (Birnboim and Doly, 1979).

The E. coli strain containing the plasmid was grown up overnight at 37°C in 10 mL of LB medium. This was used to inoculate 500 mL of LB medium in a 2 L shaker flask. After being incubated on a platform rotary shaker at 37°C until an $O.D._{600}$ of 0.6-0.9 had been reached, 2 mL of the chloramphenicol solution was added. The culture was then incubated overnight.

In order to collect the cells, the culture was centrifuged in a Sorvall GSA rotor at 5,000 rpm for 5 min at 4°C. The pellets were then resuspended in a total of 20 mL of lysis buffer containing 2 mg/mL lysozyme. After a 30 min wait on ice 40 mL of the NaOH/SDS solution was added and the mixture placed on ice for another 5 min. The suspension was held on ice for a further 60 min after 30 mL of 3M sodium acetate, pH 4.8, was added. The mixture was then centrifuged at 5,000 rpm for 15 min in a Sorvall SS-34 rotor at 4°C. The supernatant was carefully transferred to another tube and respun under similar conditions. Two volumes of 95% ethanol was then added to the supernatant and the mixture precipitated at -60°C for 30 min. The nucleic acids were collected by a centrifugation at 5,000 rpm for 10

min in a Sorvall SS-34 rotor at 0°C. The pellet was resuspended in 20 mL of TAES solution and then combined with an equal volume of phenol:chloroform:isoamyl-alcohol (25:24:1). A centrifugation at 5,000 rpm for 10 min in a Sorvall SS-34 rotor at 4°C was carried out and the aqueous phase collected. The lower phase was re-extracted with another 20 mL of TAES and the aqueous phases pooled. This phase was then precipitated with ethanol overnight.

The nucleic acids were collected by centrifugation at 5,000 rpm for 10 min in a Sorvall SS-34 rotor at 0°C. This pellet was redissolved in 5 mL of water before 2 mL of 1 M sodium acetate, pH 8.0, was added. Another ethanol precipitation was carried out. Water and sodium acetate were added again and a similar precipitation done. The final pellet was redissolved in 2 mL water and 100 uL of a RNase A solution that had been boiled for 10 min. This mixture was incubated at 37°C for 30 min. Forty uL of 4 M sodium acetate, pH 6.0, and one volume of 95% ethanol was added and the solution allowed to sit at room temperature for 10 min. The plasmid DNA was then collected by centrifugation at 10,000 rpm for 5 min in a Sorvall SS-34 rotor at 20°C. This extraction and centrifugation were carried out several times until a clear supernatant was obtained. The final pellet was dried in a dessicator at room temperature before being resuspended in 0.2-1.0 mL of TE buffer and stored at 4°C.

6. DNA-DNA hybridization

6.1 Northern Blots

DNA that had been digested and electrophoresed was transferred onto Genescreen filters according to the manufacturer's instructions.

The gel was trimmed to keep only the desired areas and incubated in 200 mL of denaturing solution for 30 min at room temperature. It was transferred to 200 mL of neutralizing solution and incubated under similar conditions. After being cut to exactly the same size as the gel, the Genescreen was wet in water and placed onto a 10X SSC solution for 15 min.

For transfer the gel was laid onto a glass plate so that no air bubbles were trapped underneath. The membrane was placed onto the gel with the correct side in contact with the gel and with no air bubbles trapped between the two surfaces. Five pieces of dry Whatman 3MM paper, the same size as the gel, were placed onto the membrane. A 2-3 in stack of paper towels, the same size as the gel, was laid over the paper. Finally a weight (approximately 500 g) was placed on the towels and the transfer allowed to continue for 12-24 h.

After transfer the membrane was dried at room temperature for about 1 h, sandwiched between Whatman 3MM paper and wrapped in aluminum foil. The blots were stored in a dessicator under vacuum at room temperature for further use.

6.2 Plaque blots

Plaque lifts of the phage plated out from the libraries were done according to the procedure described by Benton and Davis (Benton and Davis, 1977).

Using forceps, an 82 mm nitrocellulose disc was placed onto a replica-plating block and marked in an asymmetric pattern with a waterproof marking pen. A pre-chilled phage plate was placed onto the disc and marked by tracing the pattern with a waterproof marking pen. A pre-chilled phage plate was placed onto the disc and marked by tracing the pattern from the disc. After 30 sec, the filter was transferred face-up to 200 mL of denaturing solution. After floating the filter on the surface, the upper side was covered with fluid. The blot was soaked 1 min, transferred to 200 mL of 200 mM Tris, pH 7.5, and left for 1 more min. This wash was repeated before a final wash with 200 mL of 2X SSC was done. The filter was then dried for 1 h on Whatmann 3MM paper and baked at 80^{o}C for 2 h. These blots were stored between pieces of Whatmann 3MM paper and wrapped in aluminum foil. They were kept in a dessicator under vacuum at room temperature until use.

The corresponding phage plates were sealed with Parafilm and stored at 4^{o}C until the results of the hybridizations had been obtained.

6.3 Dot-blots

Preparation of the dot blots was carried out according

to the instructions accompanying the Genescreen membrane and the BRL dot-blot apparatus (Kafatos et al, 1979).

The membrane was floated on 2 washes of water and placed in the apparatus. The DNA was denatured by boiling for 10 min and held on ice. One to ten uL of DNA were applied to the filter under vacuum. Where varying amounts of DNA were added to produce a standard curve, salmon sperm DNA was added to each well to keep the total amount of DNA in each sample constant. After all the samples had been loaded, each well was washed with 100 uL of TE. The vacuum was slowly released and the membrane dried at room temperature for several hours. It was then sandwiched between 2 pieces of Whatmann 3MM paper, wrapped in aluminum foil, and stored in a dessicator under vacuum at room temperature.

6.4 Radioactive labelling of DNA, hybridization and autoradiography

DNA was radioactively labelled by nick translation (Rigby, et al. 1977) and separated from the nonincorporated nucleotides by chromatography as described below (Maniatis, et al, 1982).

Five uL of the nick translation buffer containing the the dNTPs, 1 ug of DNA, 1 uL of DNase I solution, 1 uL of E. coli DNA polymerase I (5 units/uL), 5 uL of [alpha-^{32}P]dATP (166 pmoles at 500 uCi/pmole), and HPLC water to bring the total volume up to 50 uL, were combined in a 1.5 mL microcentrifuge tube. The reaction was

incubated at 15°C for 60-90 min.

A siliconized glass wool plug was stuffed into the tip of a plastic disposable 5 mL pipette. The pipette was then filled with 5 mL of Sephadex G-50 that had been buffered with TE. After the reaction mixture was loaded, the column was kept from drying by the continuous addition of TE buffer. The radioactivity was followed with a Geiger counter and the labelled DNA collected in the first peak eluted.

Prehybridization, hybridization, washing and autoradiography of the filters was carried out according to the Genescreen instructions.

The filters were incubated at 65°C for 2-4 h in heat-sealed plastic bags containing 50 mL of prehybridization solution. The fluid was then replaced by an equal volume of hybridization solution. Labelled DNA was denatured by boiling for 10 min and then added to the hybridization bag. Hybridization was carried out for 16-24 h at 65°C with continuous agitation. Following hybridization, the filters were washed 2 times for 5 min at room temperature in wash #1. They were held in wash #2 two times for 30 min each at 65°C and in wash #3 two times for 30 min each at room temperature.

Filters were immediately wrapped in Saran Wrap and placed on Kodak X-Omat RP film in Kodak X-Omatic cassettes with intensifying screens. They were autoradiographed at -60°C for 24-72 h.

When autoradiography was not needed, the dot-blot membranes could be cut up and the individual spots of hybridization analyzed quantitatively by scintillation counting.

RESULTS

RESULTS

1. Comparison of three distinct strains of Trichinella

1.1 Restriction endonuclease profiles

DNA was extracted and isolated from the P_1, AF_1, and TP strains of Trichinella spiralis (Table 1). Restriction endonuclease digestion mixtures were prepared using 0.5-1.0 ug of DNA. The twelve restriction profiles obtained after electrophoresis are given in Figure 1.

The fluorescent background smear in each lane consists of nonrepetitive DNA sequences in a variety of lengths. In contrast, the repetitive DNA seqeuences are represented by the distinct fluorescent bands since restriction endonucleases recognize and cut each member of a repetitive family identically. All of these digested sequences will be the same length and will therefore electrophorese to the same location on the gel thus forming a distinct band. The intensity of each band is determined by both the copy number of the repetitive sequence in the genome and the length of the repeat unit. Mitochondrial DNA may also account for some of the bands present in the profile.

A unique profile was obtained for each of the three genomes with all 12 enzymes. In two cases major bands line up between isolates: 1) the major HindIII bands at 0.5 kb in P_1 and TP, and 2) the AccI bands at 1.3 kb in P_1 and AF_1. There are also a variety of minor bands that correspond between isolates and this information is summarized in Table 2.

To obtain the values in Table 2 the total number of discernable bands in regions of the gel in which all the lanes are legible was compared between strains. Estimates of genetic relatedness between isolates or groups of isolates were made by calculating F, the fraction of restriction fragments shared by the pair of isolates being compared (Nei and Li, 1979). Low values of 0.27 similarity between P_1 and AF_1, 0.15 between P_1 and TP, and 0.14 between AF_1 and TP, were obtained.

Table 1. *T. spiralis* isolates.

Isolate	Host	Latitude	Longitude	Year Isolated
AF_1	arctic fox	69°15'N	105°00'W	1980
AF_2	arctic fox	69°15'N	105°00'W	1980
AF_3	arctic fox	69°15'N	105°00'W	1980
AF_4	arctic fox	69°15'N	105°00'W	1980
PF_1	pen fox	42°00'N	78°00'W	1982
UPB_3	black bear	42°00'N	78°00'W	1982
MSIL	marten	56°00'N	99°00'W	1980
SL	wolverine	55°00'N	100°00'W	1979
TC	polar bear	58°00'N	95°00'W	1976
M_{289}	marten	45°30'N	78°00'W	1983
M_2	marten	45°30'N	78°00'W	1983
M_3	marten	45°30'N	78°00'W	1982
M_5	marten	45°30'N	78°00'W	1982
F_1	fisher	46°30'N	77°30'W	1982
F_{424}	fisher	45°30'N	78°00'W	1982
P_1	pig	43°00'N	81°00'W	1952
P_2	pig	44°00'N	63°00'W	1980
PB	pig-obtained from Institute of Parasitology, Beltsville, Maryland - history unknown			
UPB_6	black bear	42°00'N	78°00'W	1982
UPB_8	black bear	42°00'N	78°00'W	1982
TP	racoon	43°00'N	77°00'E	1972

Figure 1.

Restriction endonuclease profiles of DNA from *T. spiralis* isolates P_1, AF_1, and TP, using 12 enzymes as shown. The standard (STD) is the BRL 1 kb ladder. Electrophoresis was performed in 1% agarose for 4.5 h at 80 V.

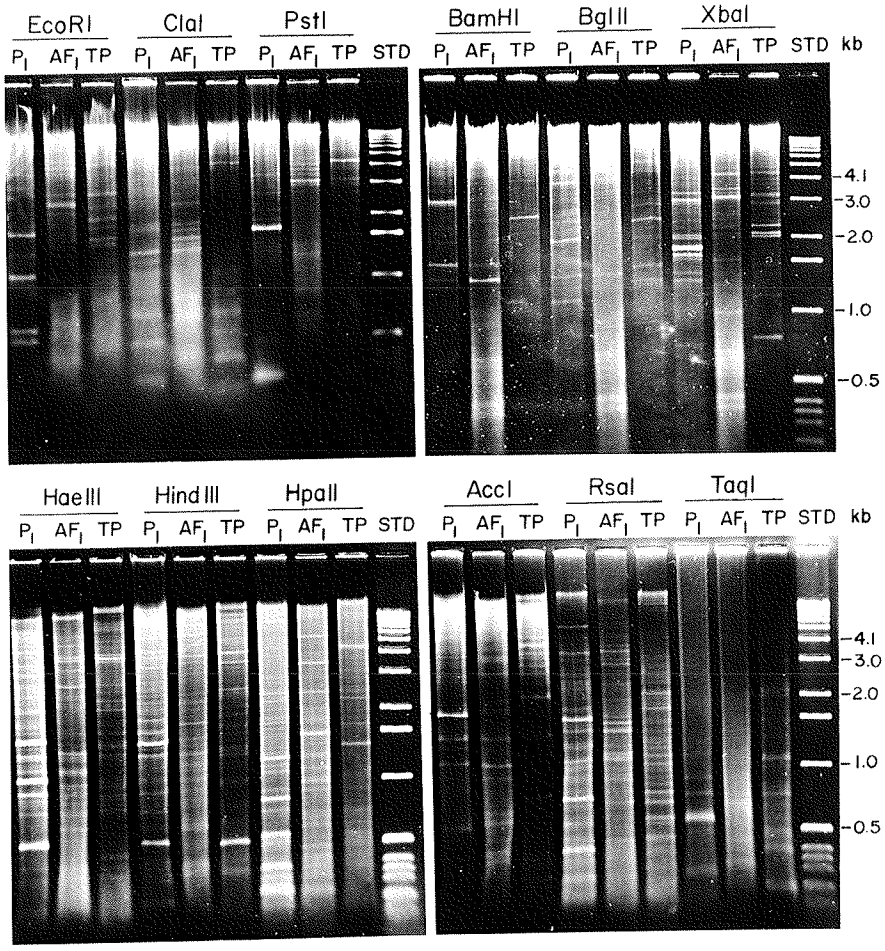

EcoRI
ClaI
PstI
BamHI
BglII
XbaI
P1 AF1 TP P1 AF1 TP P1 AF1 TP STD P1 AF1 TP P1 AF1 TP P1 AF1 TP STD kb
4.1
3.0
2.0
1.0
0.5
HaeIII
HindIII
HpaII
AccI
RsaI
TaqI
P1 AF1 TP P1 AF1 TP P1 AF1 TP STD P1 AF1 TP P1 AF1 TP P1 AF1 TP STD kb
4.1
3.0
2.0
1.0
0.5

Table 2. Relationships between repetitive DNA families of the three strains of T. spiralis isolates.

Isolates	N_x	N_y	N_{xy}	F
P_1 - AF_1	69	50	16	0.27
P_1 - TP	69	51	9	0.15
AF_1 - TP	50	51	7	0.14

N_x = number of bands in restriction profiles of first isolate

N_y = number of bands in restriction profiles of second isolate

N_{xy} = number of bands shared by both isolates

$F = 2N_{xy}/N_x+N_y$ (Nei and Li, 1979)

1.2 The plasmid probe for the 1.7 kb repetitive family

1.2.a. Source of recombinant plasmid

The plasmid, pPRA, was received from Dr. Glen Klassen, University of Manitoba. The insert is a member of the 1.7 kb repetitive family in the P_1 genome. It was cloned into the EcoR1 site in pUC9 and maintained in JM103 competent cells.

1.2.b. pPRA restriction mapping

The fragment sizes obtained by digestion of pPRA with various endonucleases are presented in Tables 3-5. The restriction map of pPRA is shown in Figure 4. The insert in pPRA was found to contain no sites for BamHI, BglII, HincII, and HindII. When the plasmid is digested with EcoRI, two bands of 2.7 kb and 1.7 kb are produced. Since the linearized pUC9 vector is 2.7 kb the 1.7 kb band must represent the insert.

i. AccI site

The only AccI site in the vector is in the multiple cloning site (MCS), downstream from the insert. The presence of two bands (3.89 and 0.51 kb) after AccI digestion indicates that there must be one site in the insert. Therefore the 3.89 kb fragment must begin in the MCS (near the 3' end of the insert) and extend around the vector, past the 5' end and into the insert. Since the vector is 2.7 kb this fragment must end 1.19 kb from the 5' end. The 0.51 kb fragment makes up the distance between the 1.19 kb site and the MCS (see Figure 2a).

Table 3. Molecular size (kb) of fragments generated by restriction enzyme cleavage of pPRA.

Fragment number	Restriction enzyme								
	AccI	PstI	ClaI	XbaI	HaeIII	Sau3AI	HindIII	HpaII	TaqI
1	+3.89	+3.79	+4.40	+4.40	1.13	1.15	+2.91	0.97	1.55
2	0.51	0.61			0.72	v0.96	1.45	0.66	*0.76
3					v0.59	v0.59	0.20	v0.54	0.62
4					v0.47	0.48		v0.49	v0.48
5					v0.44	v0.34		v0.41	0.17
6						v0.26		0.39	0.13
7								v0.24	0.11
8								v0.19	

Only the fragments that were visualized on gels are listed.

+ = Fragment seen on gel but not accurately measured by gel.

* = Fragment represented in twice molar amount.

v = Fragment located in vector only.

Table 4. Molecular size (kb) of fragments generated by double digestion of pPRA with restriction enzymes.

Fragment number	Restriction enzymes								
	ClaI x BamHI	XbaI x BamHI	HaeIII x BamHI	Sau3AI x PstI	Sau3AI x AccI	HindIII x EcoRI	HindIII x PstI	HpaII x ClaI	HpaII x HindIII
1	2.85	4.10	0.95	1.12	1.10	+2.70	+2.91	0.92	0.97
2	1.55	0.35	0.75	v0.93	v0.90	1.45	0.82	*0.52	v0.58
3			v0.63	v0.58	v0.55	*0.20	0.62	v0.44	*0.53
4			v0.49	0.47	0.36		0.27	v0.41	v0.42
5					v0.34				

Only the fragments that were visualized on gels are listed.

+ = Fragment seen on gel but not accurately measured by gel.

* = Fragment represented in twice molar amount.

v = Fragment located in vector only.

Table 5. Molecular size (kb) of fragments generated by double digestion of pPRA with TaqI and a second restriction enzyme.

Fragment number	Restriction enzymes: TaqI x EcoRI	TaqI x XbaI	TaqI x PstI	TaqI x HindIII	TaqI x HpaII
1	+v 1.44	+v 1.44	v 1.50	+v 1.44	0.59
2	+* 0.76	+* 0.76	v 0.76	+* 0.76	0.47
3	+ 0.51	v .49	0.64	0.56	v 0.44
4	+v 0.48	0.43	0.55	v 0.48	v 0.41
5	0.16	0.20	v 0.48	0.15	v 0.39
6	0.12	0.17	0.16	0.14	v 0.33
7	0.10	0.13	0.12		0.32
8		0.11	0.10		v 0.24
9					v 0.19
10					0.16
11					v 0.15
12					v 0.14

Only the fragments that were visualized on gels are listed.

+ = Fragment seen on gel but not accurately measured by gel.

* = Fragment represented in twice molar amount.

v = Fragment located in vector only.

Figure 2.

Fragment sizes (kb) generated after digestion of pPRA with restriction enzymes. The insert is represented by a solid line and the vector by a dotted line. MCS = multiple cloning site, A=AccI, B=BamHI, C=ClaI, E=HaeIII, N=HindIII, P=PstI, R=EcoRI, S=Sau3AI, X=XbaI.

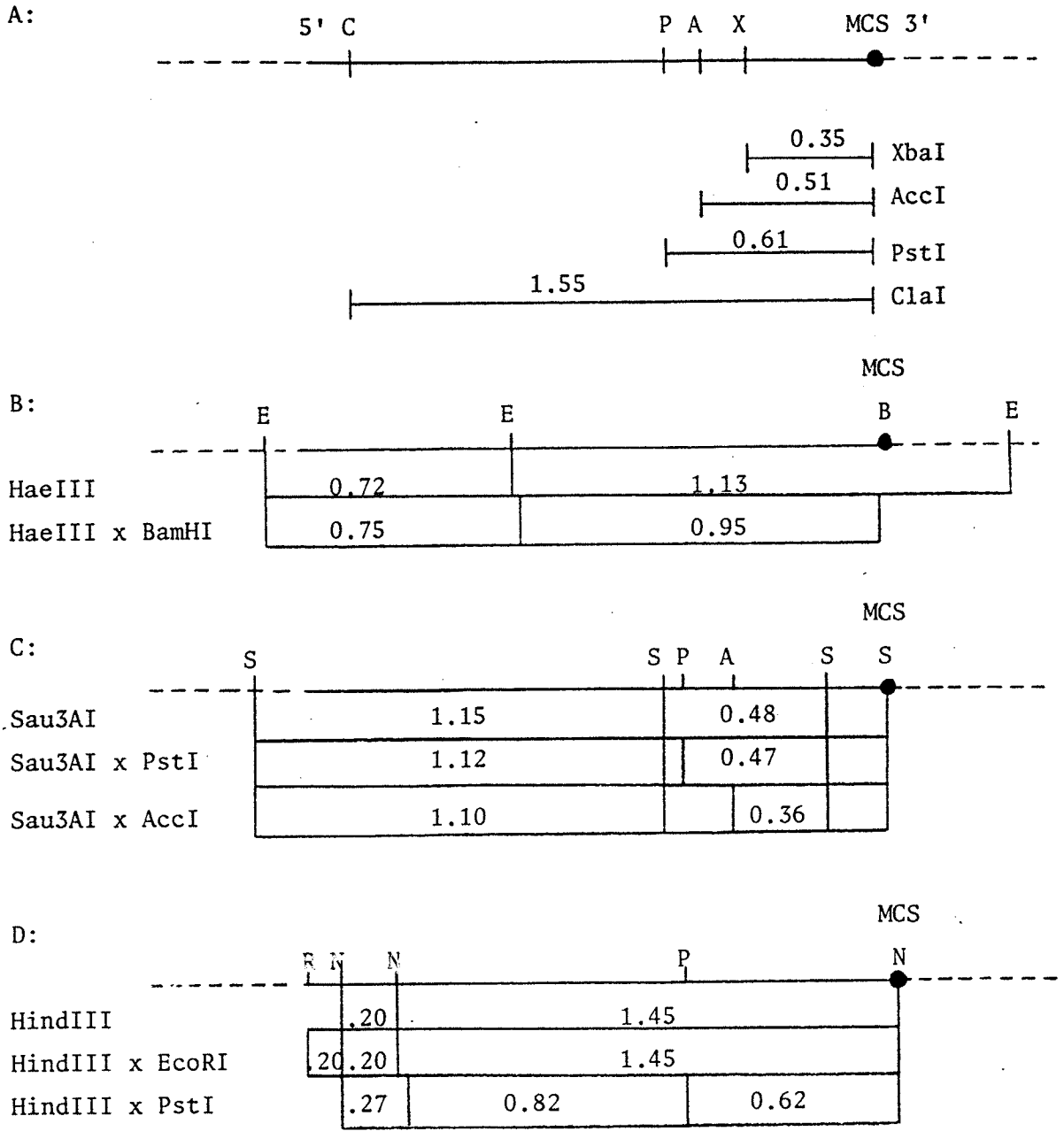
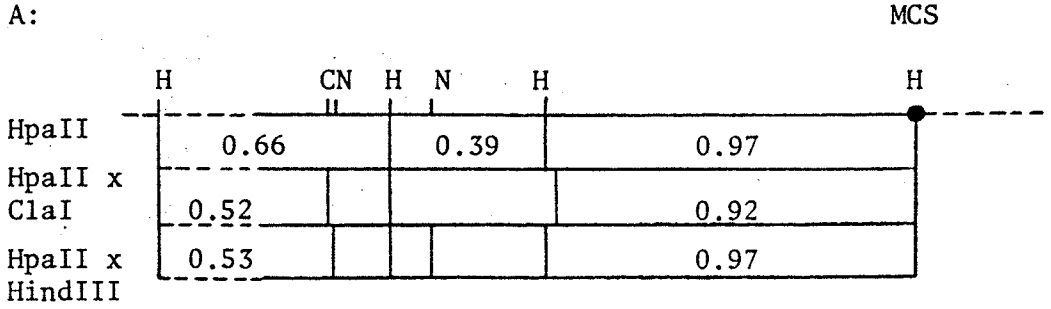

A:
5' C
P A X
MCS 3'
0.35
XbaI
0.51
AccI
0.61
PstI
1.55
ClaI
B:
MCS
E
E
B
E
HaeIII
0.72
1.13
HaeIII x BamHI
0.75
0.95
C:
MCS
S
S P A
S
S
Sau3AI
1.15
0.48
Sau3AI x PstI
1.12
0.47
Sau3AI x AccI
1.10
0.36
D:
MCS
P
N
HindIII
.20
1.45
HindIII x EcoRI
.20
.20
1.45
HindIII x PstI
.27
0.82
0.62

Figure 3.

Fragment sizes (kb) generated after digestion of pPRA with restriction enzymes. The insert is represented by a solid line and the vector by a dotted line. MCS = multiple cloning site, C=ClaI, H=HpaII, N=HindIII, P=PstI, R=EcoRI, T=TaqI, X=XbaI.

A:


MCS
H
CN
H
N
H
H
HpaII
0.66
0.39
0.97
HpaII x ClaI
0.52
0.92
HpaII x HindIII
0.53
0.97

B:

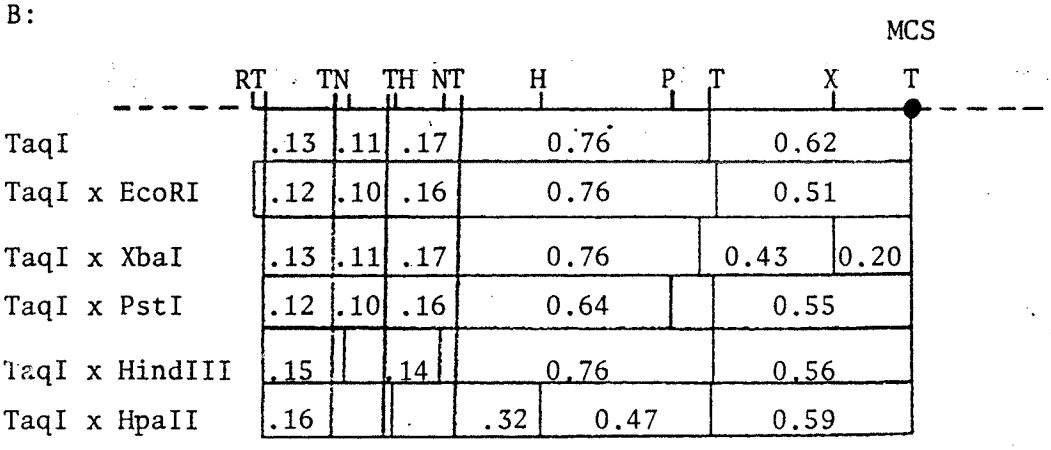
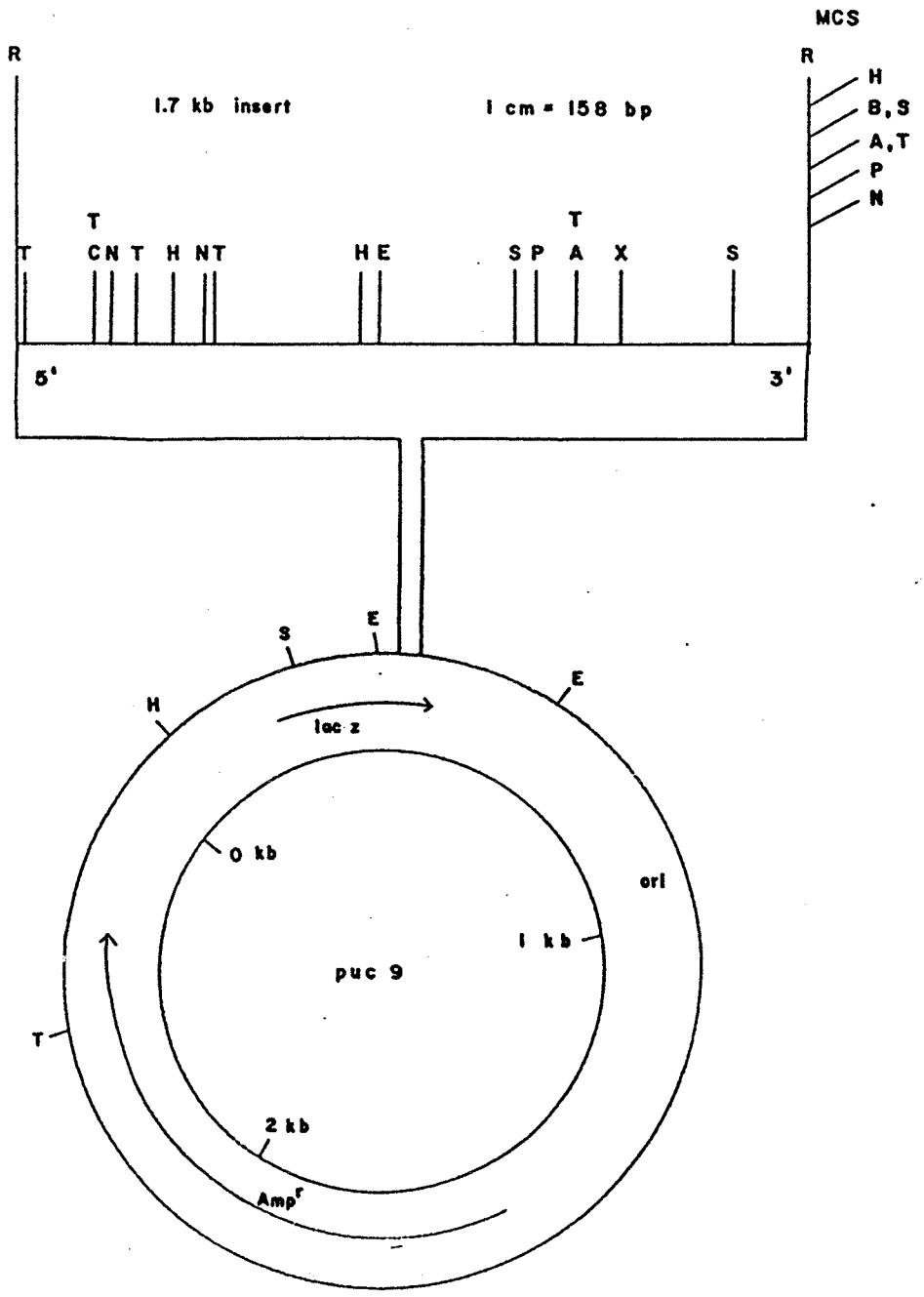

MCS
RT
TN
TH
NT
H
P
T
X
T
TaqI
.13
.11
.17
0.76
0.62
TaqI x EcoRI
.12
.10
.16
0.76
0.51
TaqI x XbaI
.13
.11
.17
0.76
0.43
0.20
TaqI x PstI
.12
.10
.16
0.64
0.55
TaqI x HindIII
.15
.14
0.76
0.56
TaqI x HpaII
.16
.32
0.47
0.59

Figure 4.

Restriction site map of plasmid pPRA. Only some of the restriction sites in pUC9 are indicated. The gene for ampicillinase (Amp^R) and the region of the origin of replication (ori) are shown. MCS=multiple cloning site, A=AccI, B=BamHI, C=ClaI, E=HaeIII, H=HpaII, N=HindIII, P=PstI, R=EcoRI, S=Sau3AI, T=TaqI, X=XbaI.


MCS
R
R
H
B, S
A, T
P
N
1.7 kb insert
1 cm = 158 bp
T
T
C N T H N T
H E
S P
T
A
X
S
5'
3'
S
E
H
E
lac z
0 kb
ori
1 kb
puc 9
T
2 kb
Amp^r

ii. PstI site

The only PstI site in the vector is in the MCS and so the large 3.79 kb fragment must include pUC9. Since pUC9 is only 2.7 kb the fragment must extend the remaining 1.09 kb into the insert. The 0.61 kb band represents the distance between the PstI site 1.09 kb from the 5' end of the insert and the PstI site in the MCS at the 3' end (see Figure 2a).

iii. ClaI site

A ClaI digestion linearizes the plasmid. Since the vector does not contain any ClaI sites, a single ClaI site must exist somewhere within the insert. The MCS contains the only BamHI site in the vector or insert. A ClaI x BamHI digestion produces two fragments confirming the presence of a single site in the insert. Since the vector is 2.7 kb the 2.85 kb fragment must begin at the BamHI site at the 3' end of the insert, include the vector and extend 0.15 kb into the 5' end. The 1.55 kb band makes up the distance between the ClaI site 0.15 kb from the 5' end of the insert and the MCS at the 3' end (see Figure 2a).

iv. XbaI site

The vector does not contain any XbaI sites. Since a XbaI digestion linearizes the plasmid, the insert must contain a single XbaI site. The only BamHI site in the vector or insert is in the MCS. A XbaI x BamHI digestion produces fragments of 4.10 kb and 0.35 kb. The 4.10 kb band must begin at the BamHI site in the MCS at the 3' end of the insert. It includes the vector and extends 1.40 kb into the

5' end of the insert. The 0.35 kb band represents the distance between the XbaI site 1.40 kb from the 5' end and the MCS at the 3' end (see Figure 2a).

v. HaeIII

pUC9 contains 11 HaeIII sites with one 0.01 kb from the 5' end of the insert and one 0.25 kb from the 3' end. Since a HaeIII digestion of pPRA yields two new fragments of 1.13 kb and 0.72 kb, there must be one HaeIII site in the insert. The 0.72 kb fragment is retained after a HaeIII x BamHI digestion. Therefore this fragment must begin at the 5' end of the insert where the HaeIII site in the vector is very close to the start of the insert. The 1.13 kb fragment has been replaced by a 0.95 kb band. Therefore the single HaeIII site in the insert lies 0.95 kb from the 3' end of the insert (see Figure 2b).

vi. Sau3AI sites

pUC9 contains many Sau3AI sites. One is located in the MCS and another is 0.13 kb upstream from the 5' end of the insert. When pPRA is digested with Sau3AI two new bands (1.15 kb and 0.48 kb) appear. These two bands do not change significantly in size when a Sau3AI x PstI digest is carried out. Therefore there must be a Sau3AI site close to the PstI site. This is only possible if the 1.15 kb fragment begins at this Sau3AI site and moves 5' to include the rest of the insert and the 0.13 kb in the vector. This means that there is a Sau3AI site 1.02 kb from the 5' end of the insert.

Since the two bands seen on the gel only account for 1.50 kb of the insert there is likely another fragment 0.20 kb in length that could not be visualized. In the Sau3AI x AccI digest the 0.48 kb fragment is missing and there is a new 0.36 fragment instead. Therefore AccI cuts the 0.48 kb fragment and this fragment must lie adjacent to the 1.02 kb fragment. This places the final Sau3AI site 1.50 kb from the 5' terminus or 0.20 kb from the 3' terminus. This confirms that there is likely a 0.20 kb Sau3AI fragment next to the MCS (see Figure 2c).

vii. HindIII sites

The only HindIII site in the vector is in the MCS. Therefore the 2.91 kb fragment, generated after a HindIII digest, must include pUC9 and begin at the MCS. Since the vector is 2.7 kb in length the fragment must include the entire vector and extend 0.21 kb into the 5' end of the insert. The presence of two more bands (1.45 kb and 0.20 kb) indicates that there is one more site in the insert.

In a HindIII x EcoRI digest the 2.91 kb band is cut into bands of 2.70 kb and 0.20 kb. This confirms that there is a HindIII site 0.20 kb from the 5' end of the insert. It should be noted that although these large bands could not be accurately sized on the gel the largest band in the HindIII x EcoRI digest was significantly smaller than the largest fragment in the HindIII digest.

The HindIII x PstI digestion splits the 1.45 kb band into pieces of 0.82 kb and 0.62 kb. Therefore the large HindIII fragment must begin at the 3' end of the insert and extend to the HindIII site 1.45 kb from the 3' end. The 0.20 kb band represents the fragment between the sites 0.20 kb from the 5' end and 1.45 kb from the 3' end (see Figure 2d).

viii. HpaII sites

There are numerous HpaII sites in pUC9. One of these sites is in the MCS and another one lies 0.33 kb upstream from the 5' end of the insert. When pPRA is digested by HpaII the 0.33 kb band disappears and three new bands are seen (0.97 kb, 0. 66 kb, 0.39 kb). Therefore there must be two HpaII sites in the insert.

A double digestion of HpaII x ClaI cleaves the 0.66 kb band and yields a new band of 0.52 kb. Therefore the 0.66 kb band in the HpaII digest must extend into the vector and there must be a HpaII site in the insert 0.33 kb from the 5' end.

In the HpaII x HindIII digestion the 0.97 kb band is not cut. This indicates that this fragment must lie at the 3' end of the insert. It begins at the MCS and extends 0.97 kb into the insert. The other 0.39 kb HpaII band is positioned between the HpaII sites 0.33 kb from the 5' end and 0.97 kb from the 3' end (see Figure 3a).

ix. TaqI sites

There are several TaqI sites in pUC9. One of these

sites is in the MCS and another one lies 0.76 kb upstream from the 5' end of the MCS. When pPRA is digested by TaqI the 0.76 kb band becomes a double band and four more bands are seen (0.62 kb, 0.17 kb, 0.13 kb, and 0.11 kb). Therefore there must be five TaqI sites in the insert. Since the 0.76 kb vector band at the 5' end of the insert does not change significantly in size there must be a TaqI site at the extreme 5' end of the insert. A double digestion of TaqI x EcoRI yields the very same set of bands that are found in the TaqI digestion. This confirms that there must be a TaqI site just inside the 5' end of the insert. Since TaqI recognizes the same sequence that is recognized by ClaI there must be another TaqI site at the ClaI site, 0.15 kb from the 5' end of the insert.

When pPRA is cut by TaqI x XbaI the 0.62 kb band is cut into two smaller pieces of 0.43 kb and 0.20 kb. The only XbaI site in pPRA is 0.35 kb from the MCS. Therefore it is likely that the 0.62 kb TaqI band lies at the extreme 3' end of the insert. If this is correct then the TaqI site is 0.62 kb from the MCS. This would place the TaqI site at approximately the same position as the AccI site. This is feasible since TaqI can recognize one of the same sequences recognized by AccI.

A TaqI x PstI digestion cuts one of the 0.76 kb fragments in the 0.76 kb TaqI double band. This 0.76 kb TaqI band is replaced by a smaller band 0.64 kb in length. Since the proposed TaqI site 0.62 kb from the MCS is about

0.10 kb from the PstI site the 0.76 kb band lies adjacent to the 0.62 kb band. This would place the TaqI site 0.43 kb from the 5' end of the insert.

With the two largest TaqI bands accounted for the amount of the insert that remains to be mapped is 0.43 kb. The three remaining TaqI bands that can be visualized on a gel total up to approximately 0.41 kb. Therefore the TaqI site located just inside the 5' end of the insert is likely about 0.02 kb from the end. It has already been suggested that there must be a TaqI site at the ClaI site 0.15 kb from the 5' end. This accounts for the 0.13 kb band between the two TaqI sites. It has also been established that there is a TaqI site 0.43 kb from the 5' end of the insert. There is only one TaqI site left to map and it must produce two TaqI bands of 0.17 kb and 0.11 kb. Double digestions of TaqI x HindIII and TaqI x HpaII indicate that the 0.17 kb TaqI band is cut by HindIII and by HpaII. Therefore the final TaqI site must be located 0.26 kb from the 5' end of the insert.

Compilation of all the information presented above indicates that the TaqI sites are located as proposed, 0.51 kb, 1.27 kb, 1.44 kb, 1.53 kb, and 1.68 kb from the MCS at the 3' end of the insert (see Figure 3b).

1.2.c. Evidence that pPRA is a typical member of 1.7 kb family

To demonstrate that a 1.7 kb sequence from P_1 had been cloned, a P_1 EcoR1 profile was probed with labelled pPRA. As seen in Figure 5 lane b, the probe hybridized strongly at

Figure 5.

Hybridization of pPRA to T. spiralis P_1 DNA. Digested DNA was run on 0.8% agarose at 30 V for 16 h. Sizes given (kb) are based on the BRL 1kb ladder.

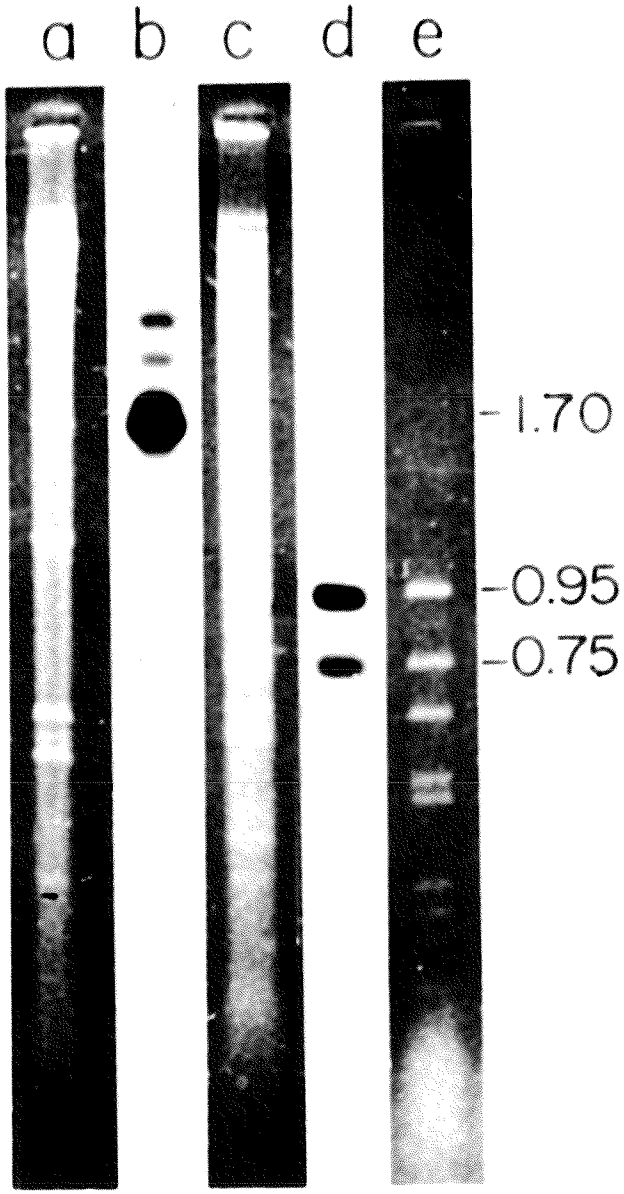
(a) EcoRI digestion of P_1 DNA.

(b) Hybridization of pPRA to the profile in lane a.

(c) EcoRI x HaeIII digestion of P_1 DNA.

(d) Hybridization of pPRA to the profile in lane c.

(e) EcoRI x HaeIII digestion of pPRA.


a
b
c
d
e
1.70
0.95
0.75

the 1.7 kb position. The probe also hybridizes to 2 bands above 1.7 kb. These bands are likely multimers of the 1.7 kb sequence.

Once the restriction map had been constructed it was possible to confirm that pPRA is a member of the 1.7 kb repetitive family. The map predicts that an EcoR1 x HaeIII digestion will produce two fragments, 0.95 kb and 0.75 kb. When the P_1 bulk DNA was digested with these two enzymes and electrophoresed, the 1.7 kb band disappeared and the two bands of predicted size appeared (Figure 5, lane c). To demonstrate that these two bands represent the sequence found in pPRA, the P_1 EcoR1 x HaeIII profile was probed with labelled pPRA. Both bands hybridized strongly to the pPRA sequence (Figure 5, lane d). A similar experiment was performed using EcoR1 and HindIII (results not shown). Again the expected results were obtained.

Further confirmation that pPRA is a representative of the 1.7 kb family comes from digestion profiles of P_1 bulk DNA. The four restriction enzymes that have a single site in pPRA (PstI, XbaI, HaeIII, and AccI) should also cut all the other members of the family only once. Therefore digestion with any one of these enzymes should produce 1.7 kb bands. This 1.7 kb band is observed in the PstI (Figure 1A, lane g), XbaI (Figure 1B, lane g), HaeIII (Figure 1C, lane a), and AccI (Figure 1D, lane a) profiles. These results indicate that the insert in pPRA shares at least five restriction sites with the majority of the 1.7 kb

repetitive family members, and thus can be considered a typical member of the family.

1.2.d. Arrangement of the 1.7 kb family

Individual members of a repetitive family may be interspersed throughout the genome or may lie adjacent to each other in a tandem arrangement. Two experiments were carried out to determine the arrangement of the 1.7 kb repetitive family.

A series of partial P_1 genomic DNA digests were obtained by digesting the DNA with EcoR1 for varying lengths of time. A blot of these digests was then probed with labelled pPRA. A ladder of bands appeared on the autoradiogram (Figure 6). Up to 30 min digestion time only high multiples of the 1.7 kb family are observed. After 60 min dimers can be visualized and after 5 h the 1.7 kb monomer is the primary band. In this experiment there is a gradual increase in the number of EcoR1 sites that are recognized and digested over time. If the members of the family do not lie adjacent to each other then dimers, trimers and other multimers could not be formed from partial digests (Horz et al, 1974; Horz and Zachau, 1977). Only if the family is tandemly arranged can a limited EcoR1 digest produce multimers of the repetitive sequence.

The partial digests also provide evidence of 2 types of heterogeneity within the 1.7 kb family. First, a significant number of dimers (3.4 kb) are not digested to monomers even after a complete 5 h EcoRI digestion (Figure

Figure 6.

pPRA hybridization to EcoRI time course digestion of T. spiralis P_1 DNA. Sizes given (kb) are based on the BRL 1 kb ladder. Digestion times (lanes a-i):

(a) 0 min (b) 15 min

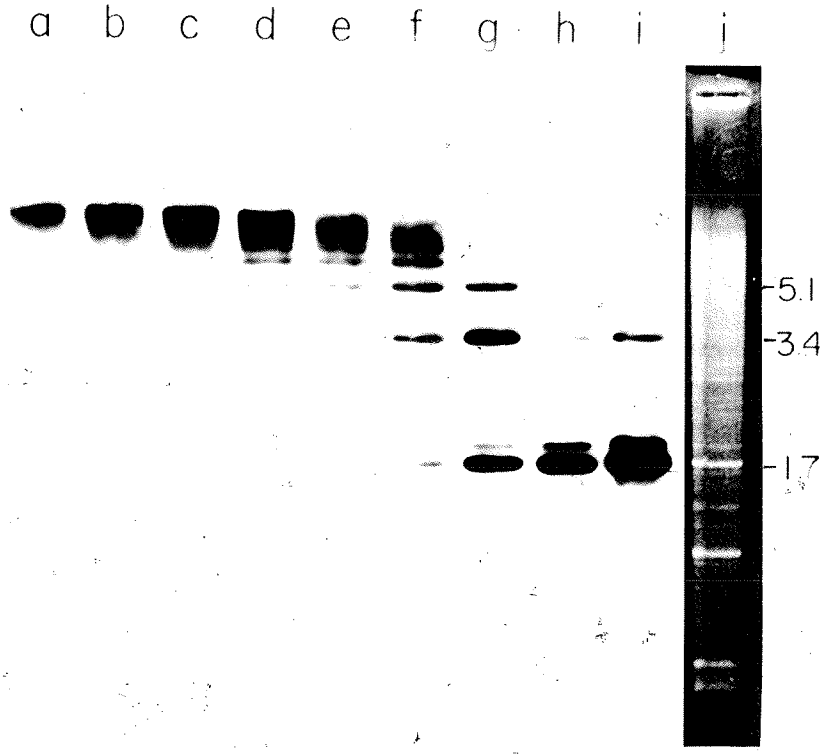
(c) 30 min (d) 45 min

(e) 60 min (f) 1.5 h

(g) 3 h (h) 5 h

(i) 20 h

(j) EtBr stained profile of lane i.


a
b
c
d
e
f
g
h
i
j
5.1
3.4
1.7

6, lane i). A few trimers (5.1 kb) also survive exhaustive digestion. These results confirm similar information about multimers outlined in section 1.2.c. It appears that the EcoRI site is not present in every member of the 1.7 kb family. A second type of heterogeneity within the 1.7 kb family is indicated by the prominent 1.9 kb band (Figure 6, lane i). This band appears opposite a distinct but minor band in the ethidium bromide profile (Figure 6, lane j).

The tandem arrangement is confirmed by the P_1 profiles obtained after digestion with PstI, XbaI, HaeIII, or AccI (Figure 1). According to the restriction site map of pPRA these four enzymes have single cutting sites in the sequence. Therefore, if the 1.7 kb family is tandemly arranged, genomic digestion with any one of these enzymes should produce 1.7 kb bands. As noted earlier, a prominent 1.7 kb band is observed in digestions with PstI (Figure 1A, lane g), XbaI (Figure 1B, lane g), HaeIII (Figure 1C, lane a), and AccI (Figure 1D, lane a). These results confirm that the 1.7 kb repetitive family is organized into tandem arrays characteristic of satellite DNA.

Tandemly arranged repeats may exist in one or a few large arrays in the genome or they may be dispersed in many smaller arrays throughout the genome. In order to discover the degree of dispersion of the 1.7 kb repeats, a λ bacteriophage EMBL3 library of P_1 DNA was constructed. In this way a sampling of the entire P_1 genome is obtained with each phage containing 13-20 kb of P_1 DNA. Host cells are

infected with the phage and plated to form plaques. Labelled pPRA was used to screen 2000 plaques for the presence of the 1.7 kb repetitive family sequence. The probe hybridized strongly to 20 or 1% of the 2000 plaques (Results not shown).

1.2.e. Copy Number of the 1.7 kb family in the P_1 genome

Two experiments were carried out to determine how many copies of the 1.7 kb repetitive sequence are present in the P_1 genome.

In the first experiment the dot-blot technique was used. In order to prepare a standard curve a series of pPRA DNA concentrations were chosen to be equivalent to copy numbers ranging from 1000 to 6000. P_1 bulk DNA was also blotted with the amount of DNA in each dot determined by assuming a P_1 genome size of 2.53×10^8 haploid bp. The series of unlabelled pPRA DNA and eight replicates of unlabelled P1 DNA were probed with labelled pPRA DNA. After exposing the dot-blot to a film for an hour in order to produce an autoradiogram, the individual dots were cut out and measured in a scintillation counter. The series of hybridizations to the unlabelled pPRA DNA was used to construct a standard curve of the number of counts vs. copy number (Figure 7). The results indicate that the haploid P_1 genome contains approximately 2800 copies of the 1.7 kb repeat with a standard deviation of 230. Therefore the 1.7 kb family represents about 4760 kb or 2% of the P_1 genome.

A densitometer reading (Figure 8) was used to quantitate the relative distribution of the four highly repetitive EcoRI families seen in the P_1 profile (Figure 1A, lane a). The peaks on the graph paper that represented the repetitive families were cut out, weighed and compared to the total weight of the entire profile. The 1.7 kb peak represents approximately 4.3% of the total weight. Assuming a P_1 genome size of 2.53×10^8 haploid bp the 1.7 kb repetitive family makes up about 10.9×10^6 bp. This corresponds to a copy number of about 6400 per genome.

1.2.f. Presence of the 1.7 kb repetitive family in P_1, AF_1, and TP

In order to determine whether sequences homologous to the 1.7 kb repetitive family in P_1 exist in the other strains of Trichinella, 0.035 ug of bulk DNA from each isolate was spotted onto nitrocellulose. This dot-blot was prepared in duplicate and one membrane probed with pPRA at 65°C and the second one hybridized at 55°C. pPRA hybridizes strongly at both temperatures to P_1 bulk DNA (Figure 9A, lane 1a). There is no detectable hybridization to the AF_1 isolate at either temperature (Figure 9A, lane 2a). A small degree of hybridization is observed for TP at 65°C. When probed under less stringent conditions, a slightly stronger reaction is observed (Figure 6A, lane 3d and Figure 6B, lane 3d).

Based on the information obtained from the dot-blot, a hybridization to restriction profiles was done. A blot of P_1,

Figure 7.

Dot-blot copy number determination of pPRA in P_1 genomic DNA. pPRA samples with copy numbers ranging from 1000 to 6000 were probed with labelled pPRA. These samples were counted and the results plotted as a standard curve (o). Eight P_1 samples were also probed and the obtained counts plotted to correspond to a copy number of 1400 per haploid genome (●). The error bar indicates a standard deviation of 115. This corresponds to a genome copy number of 2800 $\pm$ 230. CPM=counts per min.

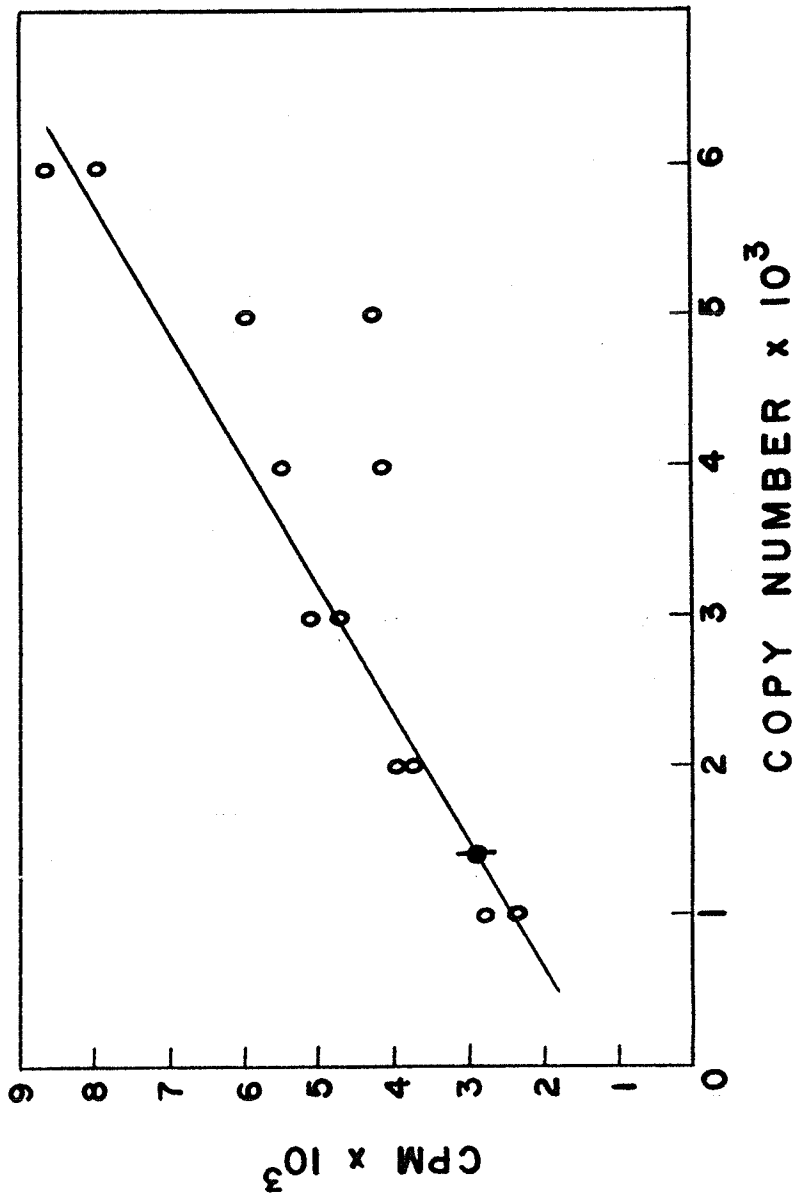
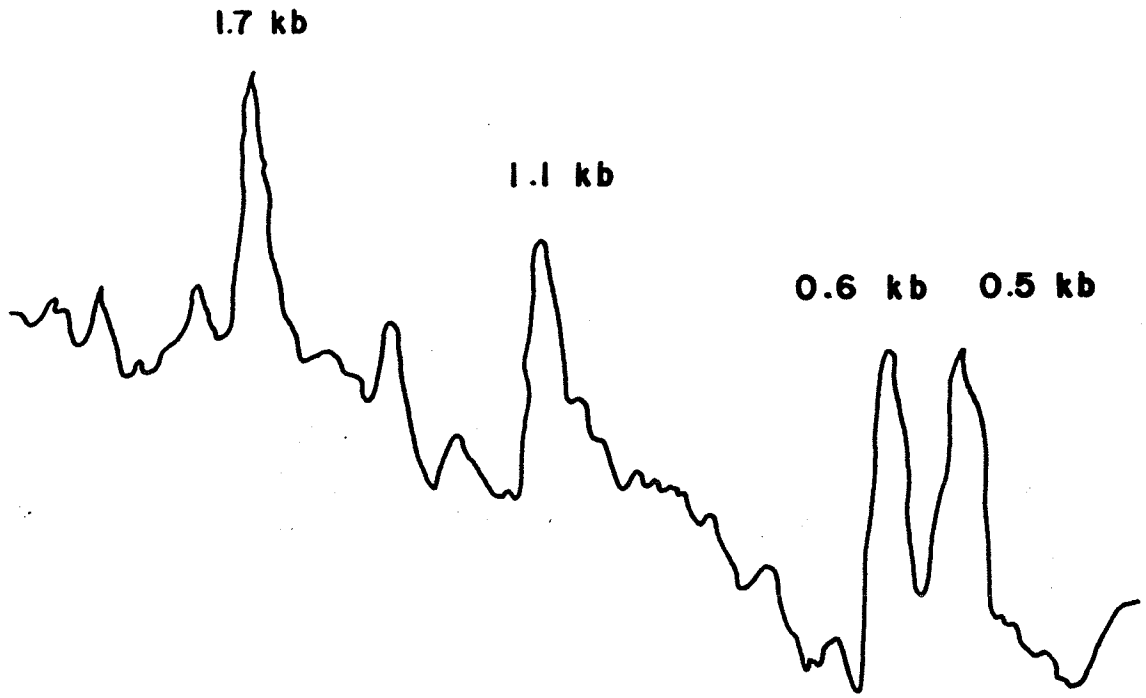

COPY NUMBER x 10^3
0
1
2
3
4
5
6
CPM x 10^3
1
2
3
4
5
6
7
8
9

Figure 8.

Densitometric scan of a photograph of EtBr fluorescence in the bands of the EcoRI digestion of P_1 genomic DNA shown in Figure 5, lane a.



P_1 x Eco RI

AF_1 and TP bulk DNA digested with EcoRI was prepared. This blot was then probed with labelled pPRA. After normal exposure of the autoradiogram strong hybridization was seen between the probe and its source, P_1 (Results not shown). No bands were observed in the lanes representing the AF_1 and TP profiles. However, when the autoradiogram was overexposed a single band appeared at the 1.7 kb position in the TP profile (Figure 10, lane 3). The TP EcoRI restriction profile in lane 4 indicates that there is no corresponding band visible at the 1.7 kb position. This indicates that the sequence which is hybridizing to the probe must be a part of the single copy DNA of the genome or one that has a relatively low copy number.

Figure 9.

Dot-blot hybridization of ^{32}P-labelled pPRA to 0.035 ug samples of bulk denatured DNA from various isolates of T. spiralis. Placement of samples:

(1a) PB_1 (1b) P_1 (1c) MSIL (1d) PF_1

(2a) AF_1 (2b) AF_2 (2c) AF_3 (2d) AF_4

(3a) UPB_6 (3b) UPB_8 (3c) SL (3d) TP

Figure 9A shows hybridization at 65°C and Figure 9B at 55°C. Only the third row of dots is shown in Figure 9B.

Figure 10.

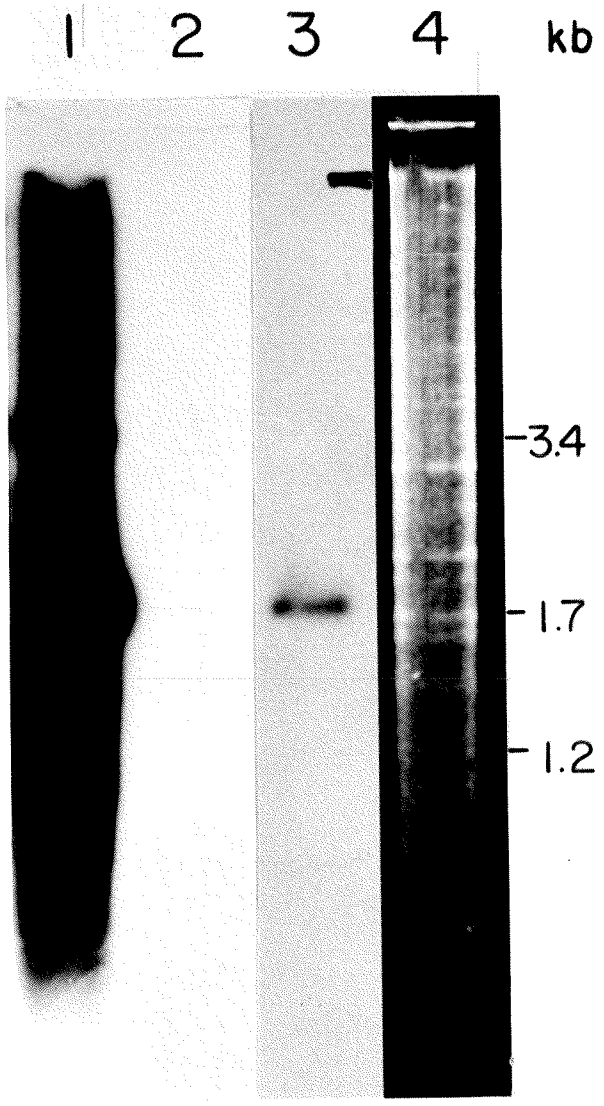
Hybridization of pPRA to EcoRI profiles of bulk DNA from a selection of T. spiralis isolates. The 0.8% agarose gel was run at 80 V for 4 h and blotted on Genescreen. Sizes given (kb) are based on the BRL 1 kb ladder.

(1) P_1

(2) AF_1

(3) TP

(4) EtBr stained profile of lane e before blotting.


1
2
3
4
kb
3.4
1.7
1.2

2. Comparison of 5 isolates in the P_1 family

2.1 Restriction endonuclease profiles

P_1 and P_2 genomic DNA were digested with 5 enzymes and restriction profiles obtained after electrophoresis (Figure 11). No differences can be seen.

One more pig isolate (PB_1) and two black bear isolates (UPB_6 and UPB_8) were digested with 6 enzymes in order to compare profiles with P_1. As demonstrated in Figure 12, there are no detectable differences between these 4 strains.

2.2 Presence of the 1.7 kb repetitive family in the P_1 family isolates

P_1, PB_1, UPB_6, and UPB_8 genomic DNA was spotted onto nitrocellulose. It was then probed with labelled pPRA in order to determine whether the 1.7 kb repetitive family is represented in these isolates. pPRA hybridizes strongly at both 65°C and 55°C to all 4 samples (Figure 9A, lane 1a, 1b, 3a, 3b and Figure 9B, lane 3a, 3b).

To confirm the similarities found within this family a blot of P_1, PB_1, UPB_6, and UPB_8 bulk DNA digested with EcoRI was prepared. This blot was then probed with labelled pPRA. Strong hybridization to the 1.7 kb band is observed in all 4 patterns (Figure 13). Hybridization is also observed to the 3.4 kb dimer in all 4 lanes. However there are differences in the intensity of the hybridization. These differences may reflect differences in copy number between isolates, sequence heterogeneity, or simply inaccuracies in measuring DNA amounts.

Figure 11.

Restriction endonuclease profiles of DNA from T. spiralis isolates P_1 and P_2, using the enzymes shown. Electrophoresis was performed in 1% agarose for 4 h at 70 V. Sizes given are based on the BRL 1 kb ladder.

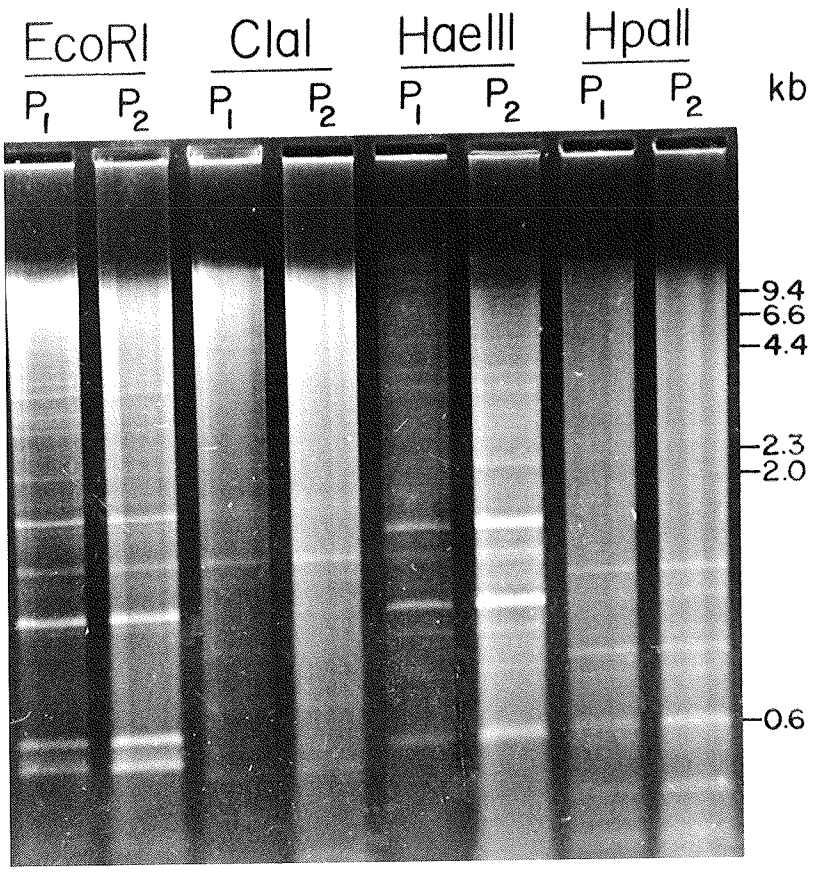
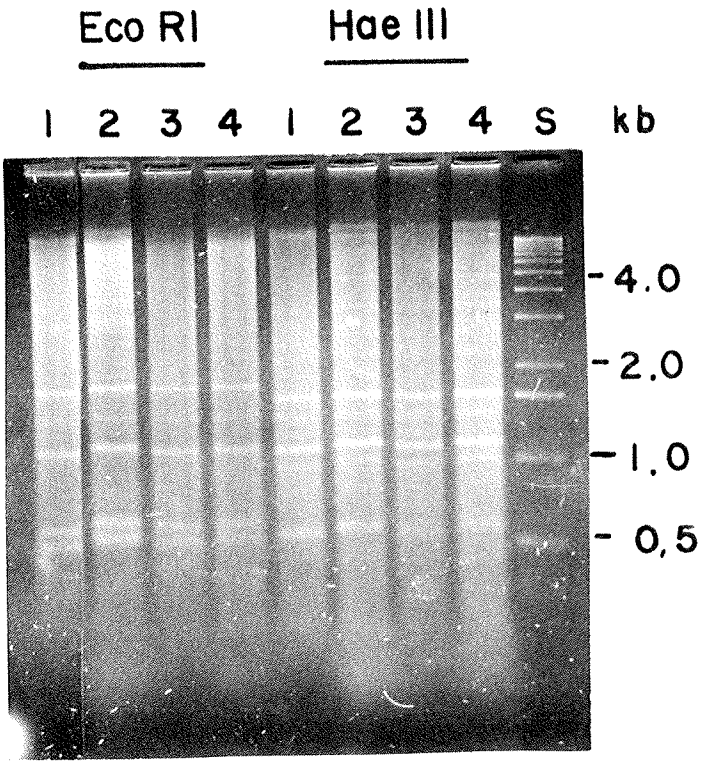

EcoRI
ClaI
HaeIII
HpaII
P1 P2 P1 P2 P1 P2 P1 P2
kb
9.4
6.6
4.4
2.3
2.0
0.6

Figure 12.

Restriction endonuclease profiles of DNA from T. spiralis isolates using the enzymes shown. The standard (S) is the BRL 1 kb ladder. Electrophoresis was performed in 1% agarose for 4 h at 80 V.

(1) P_1 (2) PB (3) UPB_6 (4) UPB_8

A


Eco RI
Hae III
1 2 3 4 1 2 3 4 S kb
4.0
2.0
1.0
0,5

B

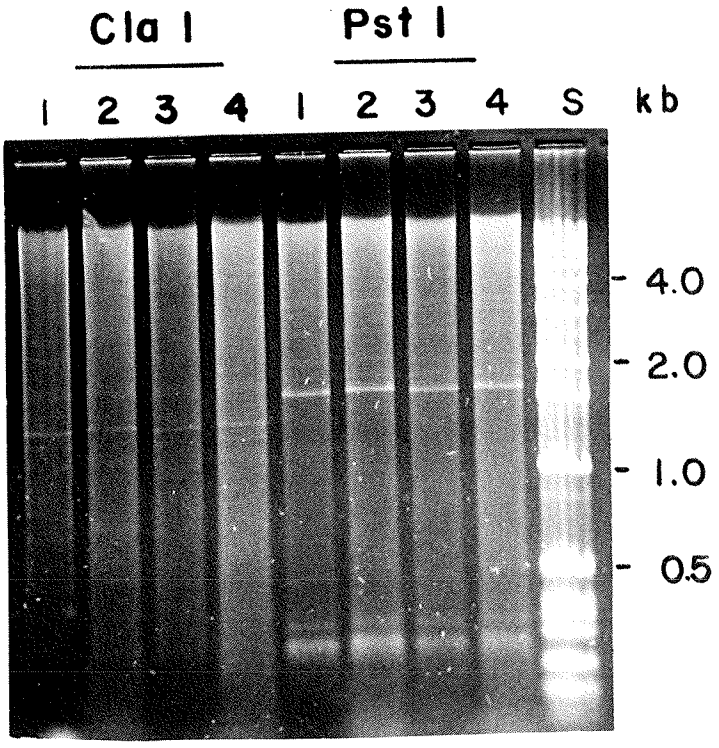
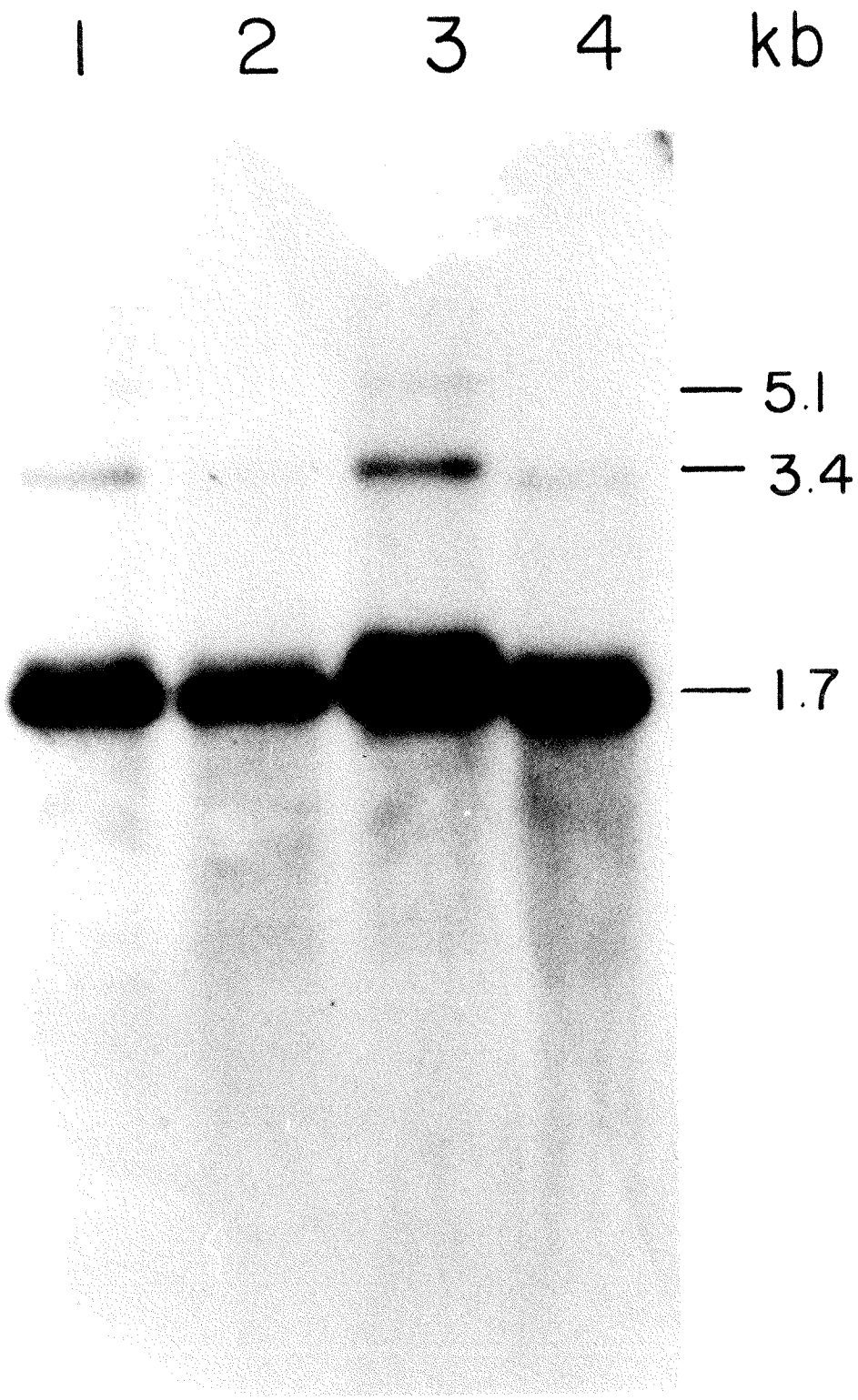

Cla I
Pst I
1 2 3 4 1 2 3 4 S kb
4.0
2.0
1.0
0.5

Figure 13.

Hybridization of pPRA to EcoRI profiles of bulk DNA from 2 T. spiralis isolates taken from domestic pig hosts (P_1, PB_1) and from 2 isolates from black bears in Pennsylvania (UPB_6, UPB_8). The 0.8% agarose was run at 80 V for 4 h and blotted on Genescreen.

(1) P_1 (2) PB_1 (3) UPB_6 (4) UPB_8

Sizes given (kb) are based on the BRL 1 kb ladder.


1
2
3
4
kb
5.1
3.4
1.7

3. Comparison of 15 isolates in the AF_1 family

3.1 Restriction endonuclease profiles

Genomic DNA from 15 isolates was digested with 5 enzymes and restriction profiles obtained after electrophoresis. Figure 10 shows results of one enzyme digestion. Although the UPB_3 profile in lane 15 is not clear it can also be seen in Figure 21B, lane 4. Analysis of these profiles allows this group of 15 to be narrowed to 9 that are significantly different from each other.

AF_1 and AF_2 have identical restriction patterns (Figure 14B). Except for minor differences obtained after digestion with ClaI , AF_3 and AF_4 are also identical to each other (Figure 15A). These differences lie between 1.4 kb and 1.6 kb on the profile. The profiles of martin isolates demonstrate that M_2, M_3, and M_5 have a similar pattern while M_{289} and MSIL have a second pattern (Figure 15B, lanes 1-5). Both fisher isolates, F_1, and F_{424}, belong to one group (Figure 15B, lanes 6-7). SL, TC, PF, and UPB_3 each have unique profiles (Figure 14A, lanes 5, 13-15). Therefore they each represent a separate group. Representatives of each of these groups were chosen based on the availabiltiy of isolate DNA. The representatives chosen for further analysis were AF_1, M_3, MSIL, F_{424}, SL, TC, AF_3, PF, and UPB_3.

The 9 representative isolates were digested with 7 enzymes and electrophoresed. A schematic summary of the profiles is presented (Figures 16-22). Although each profile falls within the basic AF pattern there are many

subtle differences. The similarities and differences have been summarized in Tables as described in Results section 1.1 (Tables 6,7). It should be noted that the negatives of the pictures in Figures 16-22 produced clearer profiles than the positives and were used for comparing and counting bands. Differences in the degree of band intensity were not included in the analysis.

The 6 groups (representing 11 isolates) with the greatest similarity to AF_1 are categorized as the AF_1 core group (Table 6). The F values are very high ranging from 0.95-0.99. The 3 groups (representing 4 isolates) showing the greatest differences from AF_1 are AF_3, PF, and UPB_3 (Table 7). The F values range from 0.56 for the AF_1-UPB_3 comparison to 0.89 for the AF_3-PF one. These values can be compared with a background F value of about 0.20 for isolates from different strains (Table 2).

3.2 Presence of the 1.7 kb repetitive family in the AF_1 family isolates

Genomic DNA from 8 AF_1 family isolates was spotted onto nitrocellulose. It was then probed with labelled pPRA in order to determine whether the 1.7 kb repetitive family is represented in these isolates. No hybridization was detected at 65°C or at the less stringent temperature of 55°C (Figure 9A, lane 1c, 1d, 2a-d, 3c, 3d and Figure 9B, lane 3c, 3d). It can be concluded that the 1.7 kb repetitive family sequence is not present in the AF_1 isolates or that it exists in very low copy numbers.

Figure 14.

A - Restriction endonuclease XbaI profiles.

Electrophoresis was performed in 1% agarose for 16 h at 30 V.

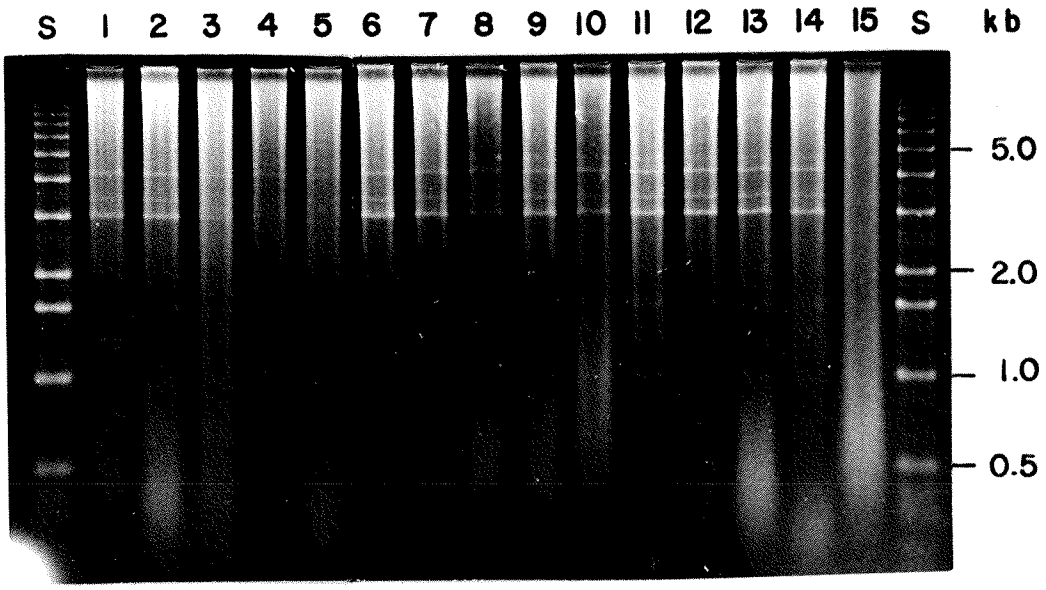
(1) AF_1	(2) AF_2	(3) AF_3
(4) AF_4	(5) PF	(6) M_2
(7) M_3	(8) M_5	(9) M_{289}
(10) MSIL	(11) F_1	(12) F_{424}
(13) SL	(14) TC	(15) UPB_3

B - Restriction endonuclease profiles using 4 enzymes as shown. Electrophoresis was performed in 1% agarose for 4 h at 80 V.

(1) AF_1 (2) AF_2

The standard used in Figures A and B is the BRL 1 kb ladder.

A


S 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 S kb
5.0
2.0
1.0
0.5

B

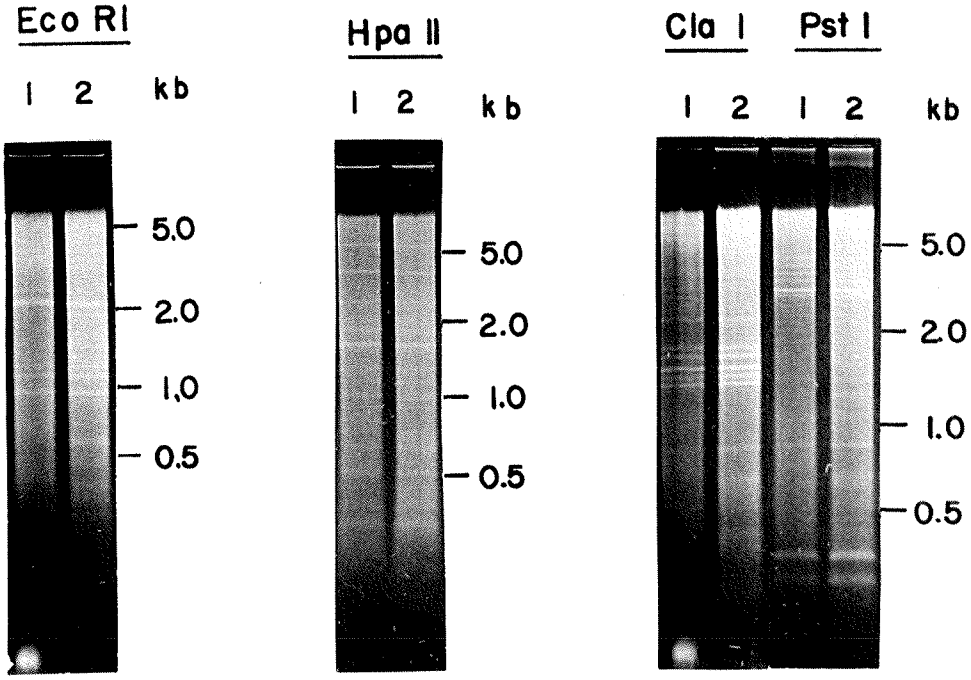

Eco RI
1 2 kb
5.0
2.0
1.0
0.5
Hpa II
1 2 kb
5.0
2.0
1.0
0.5
Cla I Pst I
1 2 1 2 kb
5.0
2.0
1.0
0.5

Figure 15.

A - Restriction endonuclease profiles of DNA from T. spiralis isolates using 5 enzymes as shown. Electrophoresis was performed in 1% agarose for 4 h at 80 V.
E=EcoRI, C=ClaI, H=HindIII, R=RsaI, X=XbaI.

(1) AF_3 (2) AF_4

B - Restriction endonuclease ClaI profiles.
Electrophoresis was performed in 1% agarose for 16 h at 30 V.

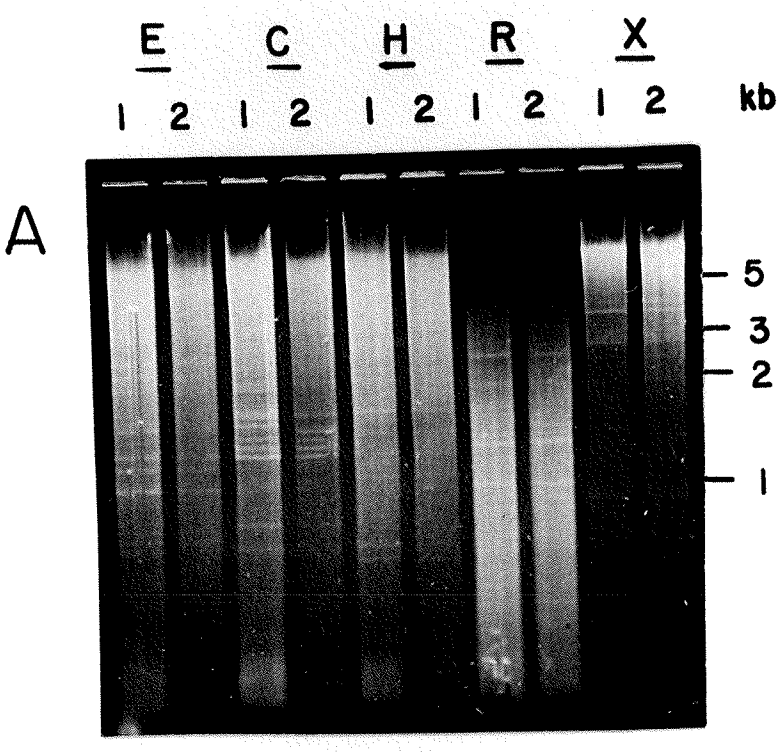
(1) M_2 (2) M_3

(3) M_5 (4) M_{289}

(5) MSIL (6) F_1

(7) F_{424}

Sizes given are based on the BRL 1 kb ladder.


E
C
H
R
X
1 2 1 2 1 2 1 2 1 2
kb
A
5
3
2
1

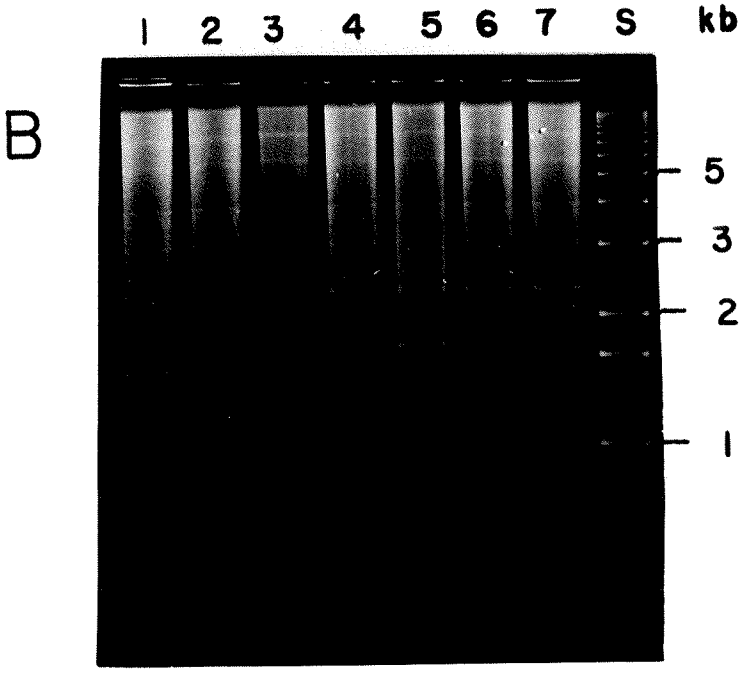

1 2 3 4 5 6 7 S
kb
B
5
3
2
1

Figure 16.

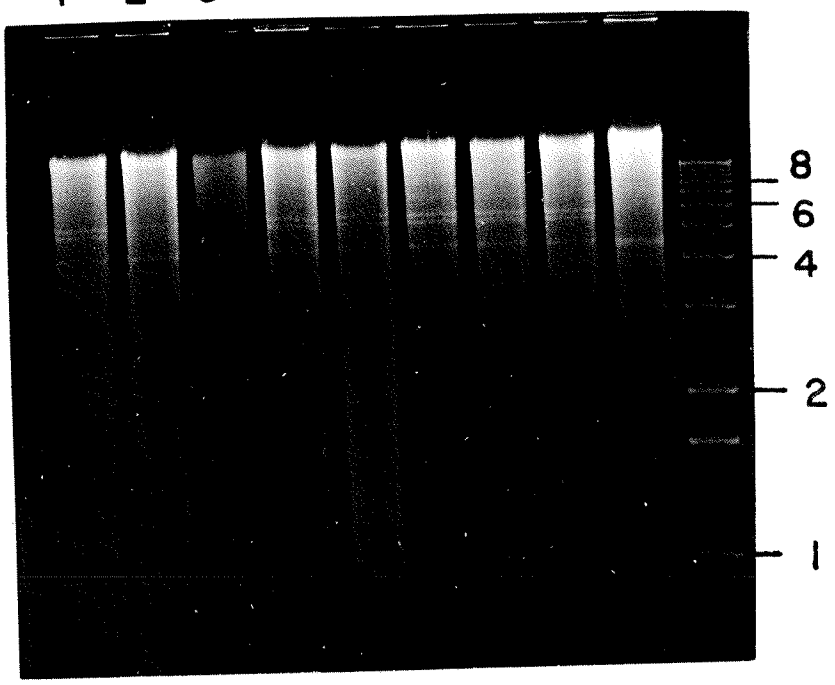
A - Restriction endonuclease BglII profiles.
Electrophoresis was performed in 1% agarose for 4 h at 80 V.
The standard (S) used is the BRL 1 kb ladder.

(1) AF_1	(2) AF_3	(3) PF
(4) M_3	(5) MSIL	(6) F_{424}
(7) SL	(8) TC	(9) UPB_3

B - Diagrammatic summary of bands in profiles in Figure A.

A=AF_1, 2=AF_3, 3=PF, 4=UPB_3.

A



B

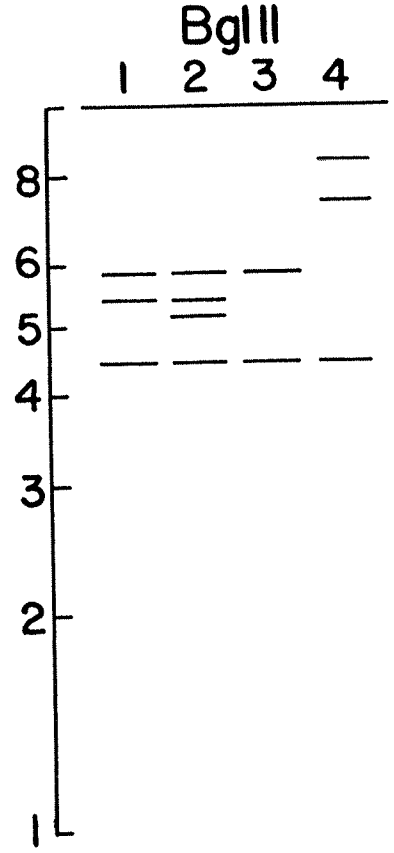


Figure 17.

A - Restriction endonuclease EcoRI profiles. Electrophoresis was performed in 1% agarose for 16 h at 30 V. The standard (S) used is the BRL 1 kb ladder.

(1) AF_3 (2) PF (3) M_3

(4) MSIL (5) F_{424} (6) SL

(7) TC (8) UPB_3

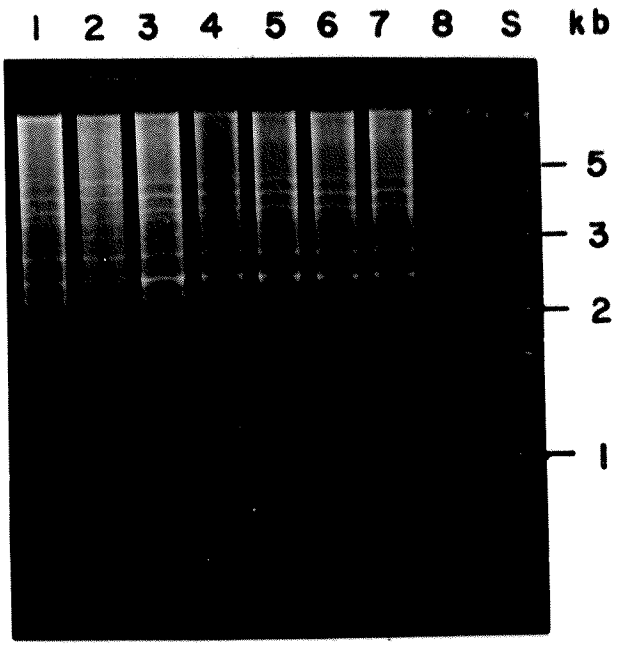
B - Restriction endonuclease EcoRI profiles. Electrophoresis was performed in 1% agarose for 4 h at 80 V. The standard (S) used is the BRL 1 kb ladder.

(1) AF_1 (2) UPB_3

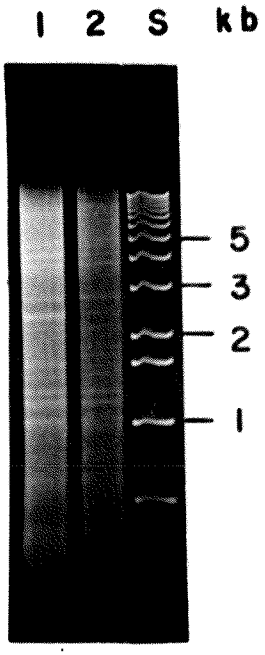
C - Diagrammatic summary of bands in profiles in Figures A and B.

1=AF_1, 2=AF_3, 3=PF, 4=UPB_3.

A



B



C

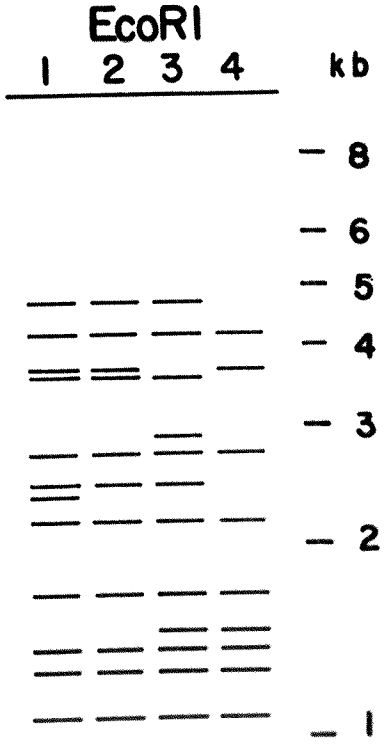


Figure 18.

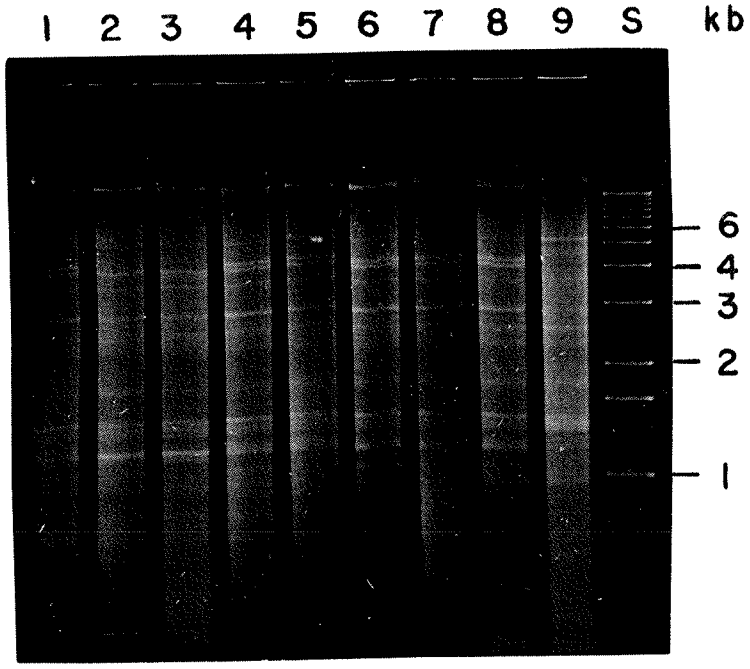
A - Restriction endonuclease HaeIII profiles. Electrophoresis was performed in 1% agarose for 4 h at 80 V. The standard (S) used is the BRL 1 kb ladder.

(1) AF_1	(2) AF_3	(3) PF
(4) M_3	(5) MSIL	(6) F_{424}
(7) SL	(8) TC	(9) UPB_3

B - Diagrammatic summary of bands in profiles in Figure A.

1=AF_1, 2=AF_3, 3=PF, 4=UPB_3.

A



B

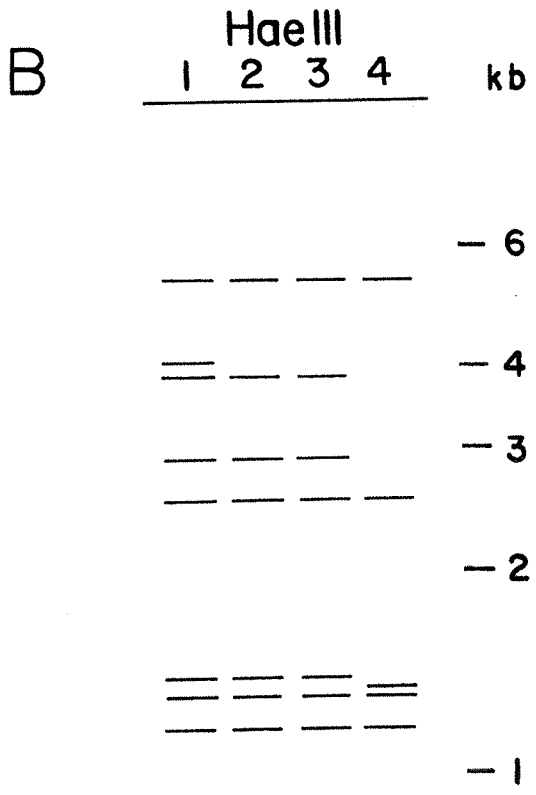


Figure 19.

A - Restriction endonuclease Hind III profiles.

Electrophoresis was performed in 1% agarose for 16 h at 25 V.

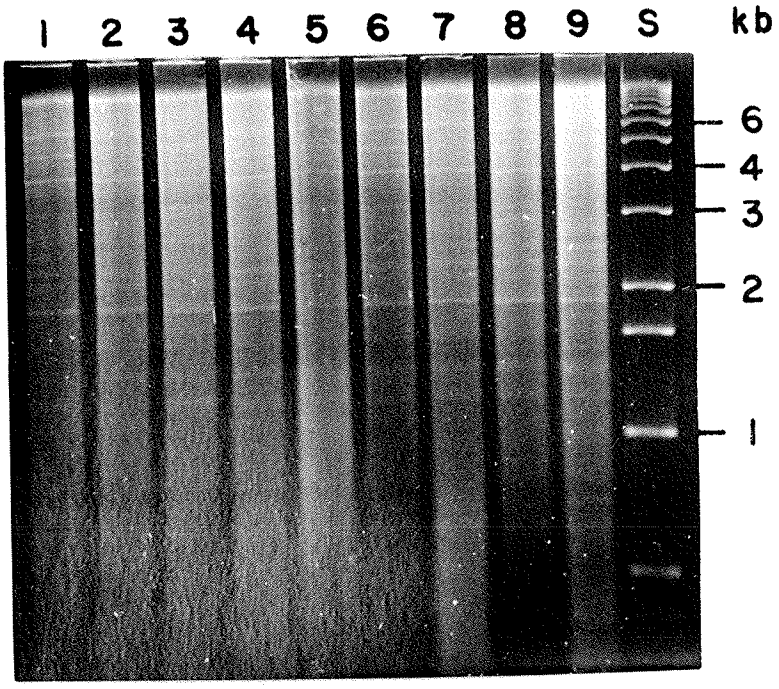
The standard (S) used is the BRL 1 kb ladder.

(1) AF_1	(2) AF_3	(3) PF
(4) M_3	(5) MSIL	(6) F_{424}
(7) SL	(8) TC	(9) UPB_3

B - Diagrammatic summary of bands in profiles in Figure A.

1=AF_1, 2=AF_3, 3=PF, 4=UPB_3.

A



B

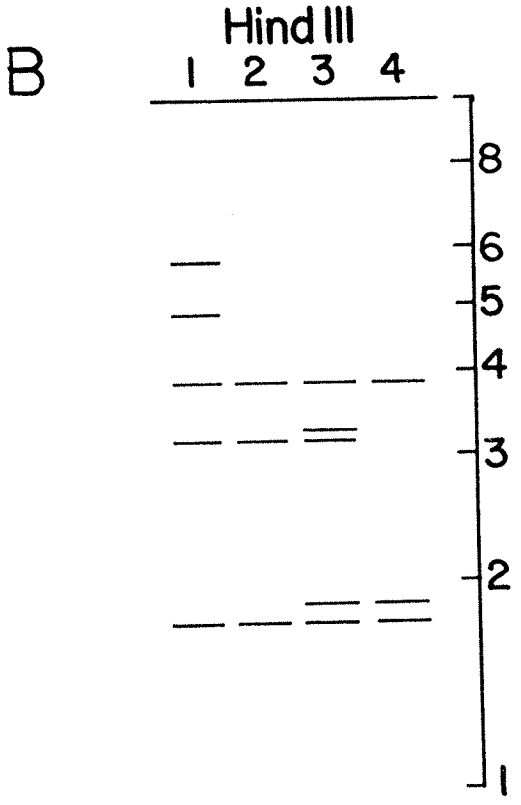


Figure 20.

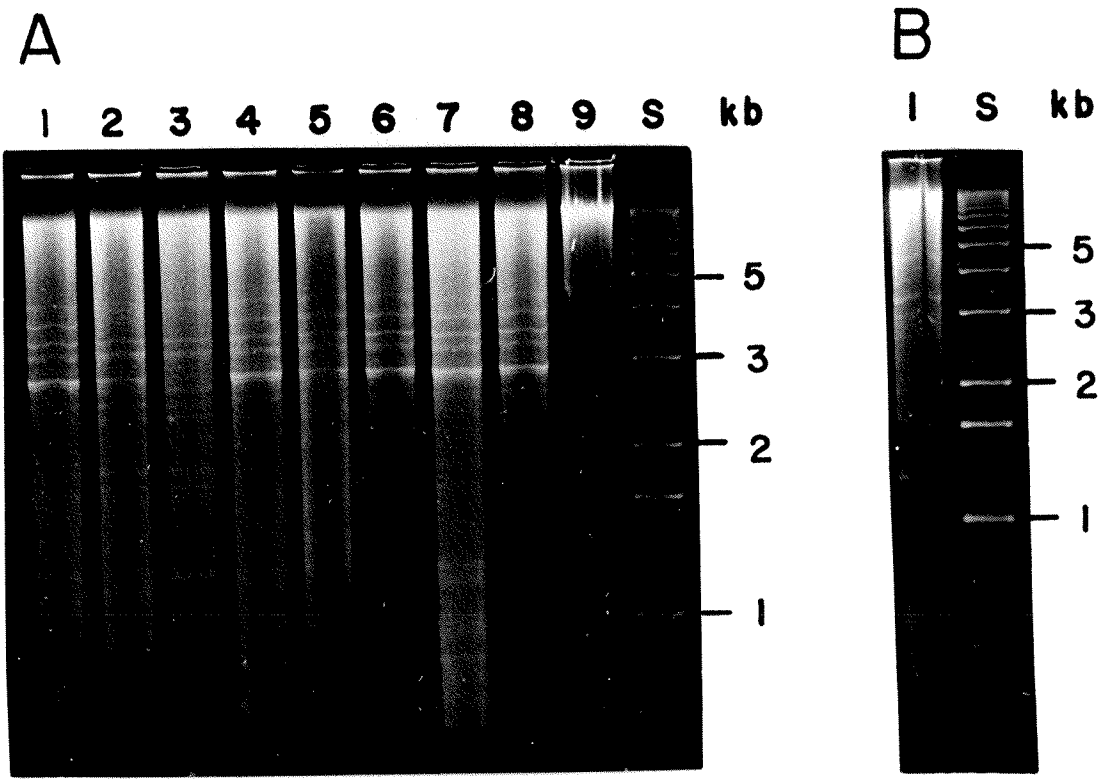
A - Restriction endonuclease PstI profiles.
Electorphoresis was performed in 1% agarose for 16 h at 30 V.
The standard (S) used is the BRL 1 kb ladder.

(1) AF_1	(2) AF_3	(3) PF
(4) M_3	(5) MSIL	(6) F_{424}
(7) SL	(8) TC	(9) UPB_3

B - Restriction endonuclease PstI profiles.
Electrophoresis was performed in 1% agarose for 17 h at 25 V.
The standard (S) used is the BRL 1 kb ladder. (1) UPB_3

C - Diagrammatic summary of bands in profiles in Figures A and B.

1=AF_1, 2=AF_3, 3=PF, 4=UPB_3.


A
1 2 3 4 5 6 7 8 9 S kb
5
3
2
1
B
1 S kb
5
3
2
1

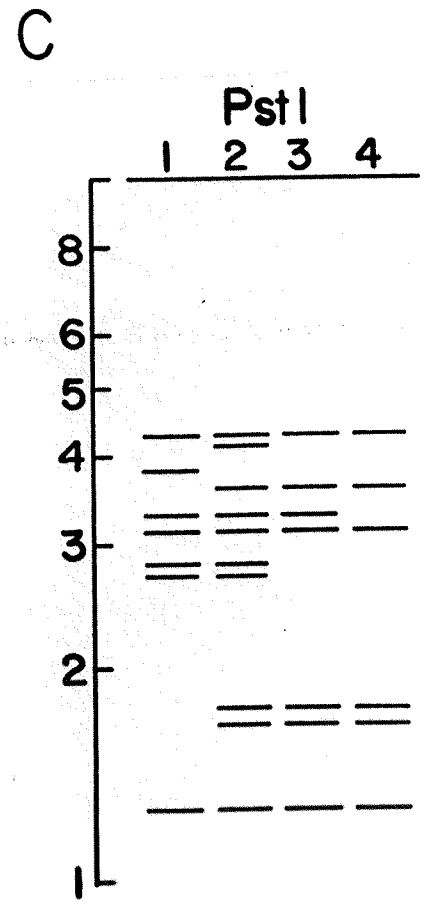

C
PstI
1 2 3 4
8
6
5
4
3
2
1

Figure 21.

A - Restriction endonuclease XbaI profiles.
Electrophoresis was performed in 1% agarose for 4 h at 80 V.
The standard (S) used is the BRL 1 kb ladder.

(1) AF_1 (2) AF_3 (3) PF
(4) M_3 (5) MSIL (6) F_{424}
(7) SL (8) TC (9) UPB_3

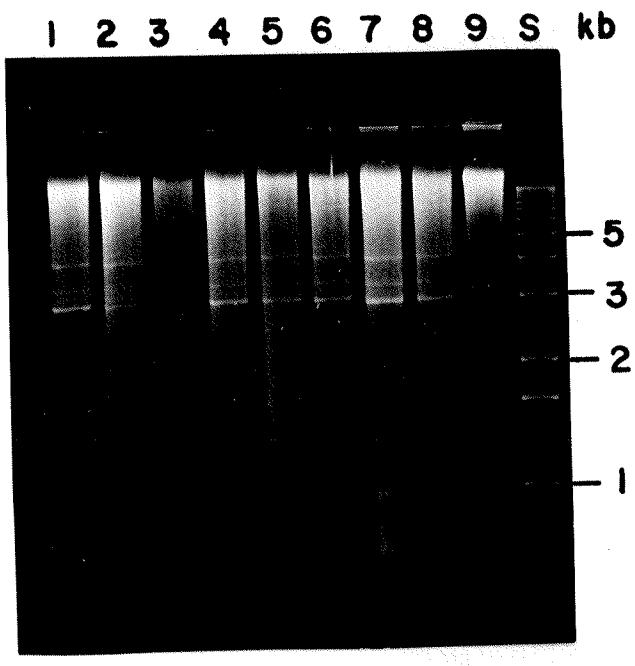
B - Restriction endonuclease XbaI profiles.
Electrophoresis was performed in 1% agarose for 4 h at 80 V.
The standard (S) used in Figure B is the BRL 1 kb ladder.

(1) AF_1 (2) AF_2
(3) AF_3 (4) PF

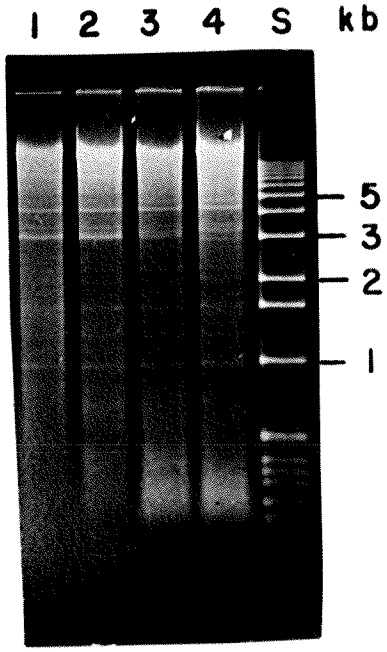
C - Diagrammatic summary of bands in profiles in Figures A and B.

1=AF_1, 2=AF_2, 3=PF, 4=UPB_3.

A



B



C

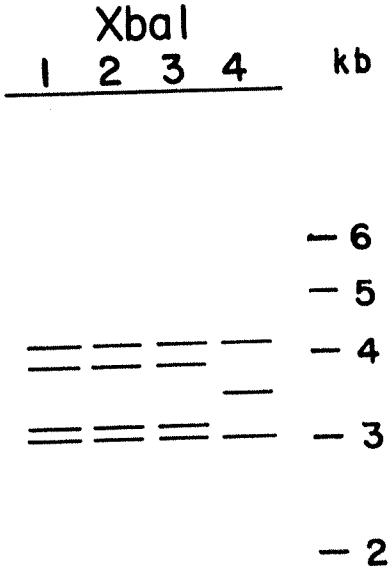


Figure 22.

A - Restriction endonuclease ClaI profiles. Electrophoresis was performed in 1% agarose for 16 h at 40 V. The standard (S) used is the BRL 1 kb ladder.

(1) AF_1 (2) AF_2 (3) AF_3 (4) AF_4 (5) PF

(6) M_2 (7) M_3 (8) M_5 (9) M_{289} (10) MSIL

(11) F_1 (12) F_{424} (13) SL (14) TC (15) UPB_3

B - Restriction endonuclease ClaI profiles. Electrophoresis was performed in 1.5% agarose for 20 h at 35 V. The standard (S) used is the BRL 1 kb ladder.

(1) AF_1 (2) AF_3 (3) PF (4) M_3 (5) MSIL

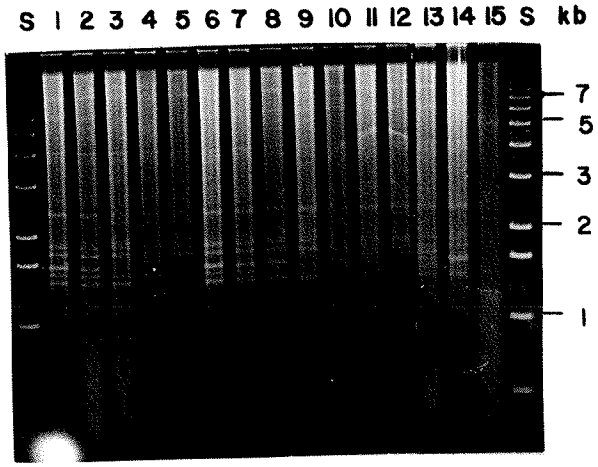
(6) F_{424} (7) SL (8) TC (9) UPB_3

C - Diagrammatic summary of bands in profiles in Figures A and B.

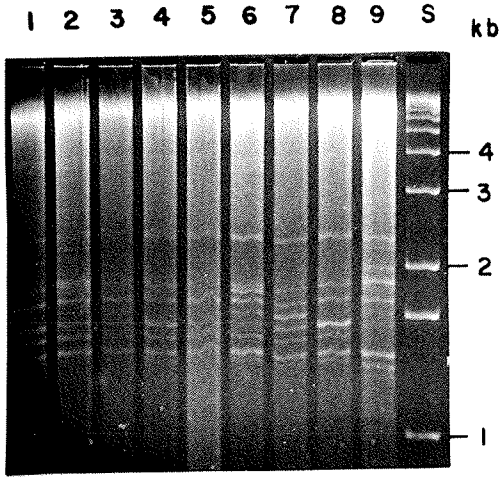
1=AF_1, 2=AF_3, 3=PF, 4=UPB_3, 5=M_3,

6=MSIL, 7=F_{424}, 8=SL, 9=TC

A


S 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 S kb
7
5
3
2
1

B


1 2 3 4 5 6 7 8 9 S kb
4
3
2
1

C

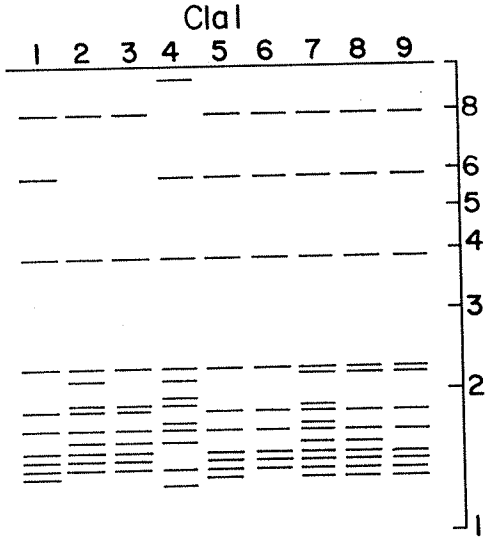

ClaI
1 2 3 4 5 6 7 8 9
8
6
5
4
3
2
1

Table 6. Relationships between repetitive DNA families within the AF_1 core group of T. spiralis isolates.

Isolates	N_x	N_y	N_{xy}	F
AF_1 - MSIL	49	48	48	0.99
AF_1 - F_1	49	53	49	0.96
AF_1 - SL	49	51	49	0.98
AF_1 - TC	49	50	49	0.99
MSIL - F_1	48	53	48	0.95
MSIL - SL	48	51	48	0.97
MSIL - TC	48	50	48	0.98
F_1 - SL	53	51	51	0.98
F_1 - TC	53	50	50	0.97
SL - TC	51	50	50	0.99

See Table 2 for definition of values

$F = 2N_{xy}/N_x+N_y$ (Nei and Li, 1979)

Table 7. Relationships between repetitive DNA families of the AF_1 strain of T. spiralis isolates.

Isolates	N_x	N_y	N_{xy}	F
AF_1 - AF_3	49	50	42	0.85
AF_1 - PF	49	47	38	0.79
AF_1 - UPB_3	49	41	25	0.56
AF_3 - PF	50	47	43	0.89
AF_3 - UPB_3	50	41	30	0.66
PF - UPB_3	47	41	30	0.68

See Table 2 for definition of values

$F = 2N_{xy}/N_x+N_y$ (Nei and Li, 1979)

DISCUSSION

DISCUSSION

Restriction endonuclease profiles of 21 T. spiralis isolates have been obtained. These profiles were compared with respect to restriction fragment length differences in their repetitive DNA. It should be noted that the differences in band patterns between isolates may be due to nonhomology between repetitive families of different isolates or to alternate restriction site spacing within homologous families. The 21 isolates can be broken down into three distinct strains. P_1, AF_1, and TP are representative of these strains.

When profiles of the three representative isolates are studied it is found that there are examples of coincident bands but that the profiles in general are very distinctive. Profiles were compared quantitatively by calculating F, the fraction of restriction fragments shared between two sets of profiles (Nei and Li, 1979). The low F values obtained for AF_1-P_1 and AF_1-TP pairs indicate that these three strains are distinct from each other (Table 2).

The fact that P_1 and AF_1 are relatively close to each other (F=0.27) while both of them are equally distant from TP (F=0.15 and F=0.14 respectively) may be significant in supporting the idea that TP is the most divergent of all T. spiralis isolates (Chambers et al, 1986).

The 5 isolates that comprise the P_1 strain (P_1, P_2, PB, UPB_6, and UPB_8) have identical profiles. This shows that

although this strain is geographically widely dispersed and is found in domestic and wild hosts, the strain has developed little variability. This may reflect a relatively recent introduction of the worm into North America.

Fifteen isolates from wild hosts comprise the AF_1 family strain. Since some of the isolates have nearly identical profiles the group can be narrowed down to 9 representative isolates. Six of these isolates are very similar and make up the "AF_1 core group". Minor differences are seen in the Cla I profiles. A comparison of these 6 isolates shows a range of F values from 0.95 to 0.99 for the 7 restriction enzymes used (Table 6).

The 3 isolates not in the core group are AF_1, PF_1, and UPB_3. (AF_3 and AF_4 are virtually identical and are represented by AF_3). These isolates have significantly lower F values (0.58 to 0.86) when compared with AF_1 and with each other (Table 7) than those within the AF_1 core group. UPB_3 is the most divergent with an average F value of 0.56. In order to determine whether these isolates should be considered part of the same strain as AF_1, F values obtained from comparing P_1 and TP with AF_1 were studied. As shown, low values of 0.27 and 0.14 respectively were obtained (Table 2). This indicates that the 4 divergent isolates are significantly more like AF_1 than are P_1 and TP.

The genetic distinctions between the 21 isolates studied may now be correlated with host species and geographical location. Since the sample size is so small the conclusions

are tentative and preliminary. It appears that the host species does not determine the genetic identity of the infecting worm. UPB_3, UPB_6, and UPB_8 were isolated from black bears in Pennsylvania. However UPB_3 is a member of the AF_1 strain while UPB_6 and UPB_8 are part of the P_1 strain. Similarily AF_1, AF_2, AF_3, and AF_4, were isolated from arctic foxes in the same region of the Northwest Territories. These two distinct types of worms have an F value of 0.85 indicating that they are not genetically identical. It appears that T. spiralis does not develop strong host specificity. This characteristic may account for the widespread distribution of the parasite.

There are no definite correlations between geographical proximity and genetic identity of the isolates. However, the 3 strains were found in different parts of the world. The P_1-like isolates come from various locales in the United States, the AF_1-like isolates range between Pennsylvania and the arctic, and TP was isolated in the Soviet Union. The remote site of isolation of the TP isolate does seem to indicate that its separation from the other strains might be correlated to its distinct genetic structure. However closer observation of the geographic locations of the isolates within the P_1 strain and the AF_1 strain indicate that genetic diversity is independent of geographic location.

The P_1 family isolates come from various parts of North America, yet they all belong to the same strain. This implies that this strain may be endemic to domestic pigs and

that the strain was likely brought to America by a small group of domesticated pigs. The occurrence of this strain of T. spiralis in wild black bears in Pennsylvania indicates that the strain of T. spiralis endemic to domestic pigs has been transmitted to black bears in Pennsylvania.

Within the AF_1 strain the most divergent isolates occur in the extreme north of the range ($AF_{3/4}$) and in the extreme south (PF_1, UPB_3). The genetic divergence of these isolates may be due to unusual isolating or hybridizing events in these locations. In the arctic, such an event would be taking place in the presence of the more conventional type of isolates ($AF_{1/2}$). In Pennsylvania, two separate events would be needed to explain the differences between the 2 divergent isolates, PF_1 and UPB_3 (F=0.68).

The comparison of restriction fragment length differences in the repetitive DNA portion of the genome appears to be an extremely sensitive method for the study of population structure, and for genetic fingerprinting of isolates but correlation of genetic distinctions with geographical location or host species is proving to be complex. Biological variation, such as that seen for infectivity (Dick et al, 1983a), may be more directly related to the genetic variation seen in this study. By continuing to analyze more isolates it may be possible to discover other correlating parameters with which to determine the strain evolution of T. spiralis.

The 1.7 kb Eco R1 repetitive family in T. spiralis P_1 bulk DNA has been partially characterized. A typical member of the 1.7 kb family (pPRA) has been cloned. By using it as a probe, the family arrangement, size, and homogeneity of the family has been investigated.

The insert in the recombinant plasmid (pPRA) was mapped with 9 restriction enzymes. Fifteen different restriction sites were found. Since there are no signs of reiteration within the segment, the 1.7 kb sequence can be taken as the monomer form of the repeated family.

When pPRA was used to probe restriction profiles of P_1 bulk DNA digestions, only bands belonging to the 1.7 kb repetitive family hybridized. This indicates that the insert in the recombinant plasmid is a member of the repetitive family. Further confirmation was obtained from restriction profiles of bulk P_1 DNA cut with 4 restriction enzymes that have a single site in the insert of pPRA (Pst I, XbaI, Hae III, and AccI). In all 4 cases a 1.7 kb band, similar to the one observed after an EcoRI digestion, was obtained. These results indicate that the cloned sequence containing these 4 sites is a member of a highly repetitive family that is 1.7 kb in length.

The experiments investigating the arrangement of the 1.7 kb family indicate that the members are arranged in arrays of direct tandem repeats. Probing a blot of partial digestions of P_1 DNA with pPRA resulted in a ladder of bands. This ladder represents multimers of the monomer 1.7 kb

sequence. The members of the family must lie adjacent to each other in order to produce dimers, trimers and other multimers from a partial digestion. Therefore the family must be tandemly arranged.

The tandem arrangement of the family is confirmed by the 1.7 kb band obtained after P_1 digestion by enzymes that have single cutting sites in the repetitive sequence. If an enzyme cuts only once in the monomer sequence and the sequence is 1.7 kb in length, a 1.7 kb band would be expected when the family is tandemly arranged.

When a P_1 genomic library was screened with pPRA, approximately 1% of the plaques contained homologous DNA. In order to interpret this data it is assumed that an average EMBL3 insert size is 17 kb and that the library is representative for repetitive sequences. If the sequences are minimally dispersed, the 17 kb insert in each phage hybridizing to the probe would consist primarily of the 1.7 kb repeated sequences. In this case, the proportion of plaques hybridizing to the 1.7 kb probe should be roughly the same as the proportion of 1.7 kb sequences in the genome. Since the obtained data indicates that about 2% of the genome consists of the 1.7 kb repetitive family (Results section 1.2.e.), approximately 2% of the plaques should hybridize to pPRA.

Since it has been demonstrated that the 1.7 kb repetitive family is tandemly arranged there must be at least two copies of the repeat adjacent to each other in any

sequence. Therefore, in the case of maximum dispersion with only two copies of the repeat next to each other, each hybridizing plaque must contain 3.4 kb of the repetitive sequence. This implies that each hybridizing plaque would contain only one-fifth of the repetitive DNA (3.4 kb) that a plaque would contain in the case of minimum dispersion (17 kb). Therefore up to 5 times as many plaques might contain the repetitive sequence if the family is maximally dispersed. In that case, the proportion of hybridizing plaques could reach 10%.

The results obtained indicate that approximately 1% of the plaques hybridize to pPRA. Therefore it is proposed that the 1.7 kb repetitive family members are minimally dispersed in a few large arrays throughout the genome.

Theoretical analysis would indicate that approximately 2% of the plaques should hybridize if the family members are minimally dispersed. This discrepancy is likely due to underrepresentation of the the 1.7 kb family in the library. This problem has been reported with tandemly repeated DNA in other genomic libraries (Botchan et al, 1974; Maniatis et al, 1982). If the P_1 library used for this experiment is severely affected by this bias the assessment of array dispersion may not be accurate.

The 1.7 kb repetitive family has a high degree of homogeneity. However there are at least 2 types of heterogeneity within the family. Even after exhaustive digestion with EcoRI , a significant number of dimers and

trimers remain. These multimers exist because the Eco site is not present in every member of the 1.7 kb family. There is also a significant degree of hybridization between the 1.7 kb family representative and a 1.9 kb band. The 1.9 kb family is likely another version of the 1.7 kb family. These heterogeneities may be the result of subarrays present in the genome. These subarrays could have been produced during successive rounds of amplification (Singer, 1982).

The results of the dot-blot experiment indicate that there are approximately 2800 copies of the 1.7 kb repeat within a haploid genome. This estimate is based on 2 assumptions. Firstly, it is assumed that there is a high level of homogeneity within the family. Each member of the family would then be recognized by the probe to the same degree. The majority of family members share precise fragment length and at least 8 restriction sites. Therefore it is assumed that the copy number estimation was not significantly affected by family heterogeneity. Secondly, it must be assumed that the library contains a fair representation of tandemly repeated DNA. It has been shown that some genomic libraries have a bias against tandem repeats (Horz et al, 1974; Maniatis et al, 1982). If this was the case in the P_1 library, then the copy number could be greater than 2800.

The estimation obtained from the densitometer reading indicates that the 1.7 kb repetitive family represents approximately 4.3% of the P_1 genome. This is a very crude

measurement and it is not seen as a contradiction of the 2% estimate provided by the more accurate dot-blot experiment.

With the characterization of pPRA it is possible to determine to what extent the 1.7 kb repetitive family is present in the T. spiralis isolates studied. The results of the hybridization experiments indicate that many copies of the 1.7 kb family exist in the P_1 family isolates. However, the sequence is not detectable in the AF_1 family and must be present in the TP isolate in a low copy number.

The restriction profiles of the pig isolates and the black bear isolates have been shown to be identical. These isolates (P_1, P_2, PB, UPB_6, UPB_8) are considered to represent a distinct strain of T. spiralis. Identical hybridization patterns are obtained when these restriction profiles are probed with pPRA (Figure 13). Therefore the 1.7 kb probe can be used as a strain-specific probe at high stringency and normal autoradiogram times.

The P_1 family members all contain the 1.7 kb repetitive family although the isolates were obtained from hosts in various parts of the United States. This confirms the premise that the host species and the geographical location of the isolates are not correlated to genetic identity.

The 1.7 kb repetitive family representative did not hybridize to any of the AF_1 family isolates. If the 1.7 kb sequence is present in these isolates its copy number must be very low. Restriction profiles have shown that the

members of this family have profiles which are similar to each other but quite distinct from other strains. The lack of hybridization between pPRA and the AF_1 family members confirms that these isolates are part of a strain distinct from the pig strain from which the insert in pPRA was cloned.

Analysis of the restriction profiles places the TP isolate in a third strain of T. spiralis. This isolate also shows a different hybridization response to pPRA. This is not unexpected since TP was isolated in the Soviet Union, a region far removed from either of the other two sources of isolates. Very faint hybridization of the probe to the 1.7 kb band in the Eco profile was observed. This degree of hybridization is not sufficient to undermine the usefulness of pPRA for strain-specific diagnosis. However it appears that the 1.7 kb sequence has been conserved in these 2 Trichinella strains and that some functional constraint may be operating on the sequence. The large contrast in copy number in the 2 strains indicates that the sequence has been amplified in the pig strain after the two strains diverged. This conclusion would be consistent with the library hyopthesis for the generation of satellite DNA families (Fry and Salser, 1977).

The pPRA strain-specific probe has already been used to identify several T. spiralis isolates from Germany and France. Through restriction profile comparisons and pPRA hybridization is was rapidly determined that these 2 isolates belong to the P_1 family (T.A. Dick, personal

communication). This information can be used to determine the source of the T. spiralis infection in the host animals.

Strain-specific probes are becoming vital for early disease identification of several parasitic infections. DNA probes specific for Plasmodium falciparum (Barker et al, 1986), Leishmania (Lopes and Wirth, 1986), filarial worms (Sim et al, 1986), and onchocerciasis (Wirth, 1986) have been prepared. The successful identification of these parasites with these strain-specific probes allows for a more accurate analysis of the epidemiology of the disease as well as allowing for a better course of treatment.

A variety of isolates should be examined in order to substantiate the proposed strain delineations. It would be useful to look at more isolates from overlapping geographical locations as well as to study the extent to which the P_1 strain has been transferred on to a variety of wild hosts in contact with the domesticated animals. Restriction profiles can be compared and the strain-specific P_1 probe can be used to detect the degree of homology between isolates. In this way it might be possible to answer the questions regarding T. spiralis evolution and speciation.

REFERENCES

REFERENCES

Arnason, U. (1982) Southern blot hybridization in cetaceans, using killer whale restriction fragment as a probe. Hereditas 97:47-49.

Arnason, U., Purdom, I.F., and Jones, K.W. (1982) Cetacean molecular hybridization using balenopterid satellite DNA cRNAs as probes. Hereditas 97:33-36.

Arnason, U. and Widegren, B. (1984) Different rates of divergence in highly repetitive DNA of cetaceans. Hereditas 101:171-177.

Arnason, U. and Widegren, B. (1986) Pinniped phylogeny enlightened by molecular hybridizations using highly repetitive DNA. Mol. Biol. Evol. 3:356-365.

Barker, R.H., Jr., Suebsaeng, L. Rooney, W., Alecrim, G.C., Dourado, H.V., and Wirth, D.F. (1986) Specific DNA probe for the diagnosis of Plasmodium falciparum malaria. Science 231:1434-1436.

Beauchamp, R.S., Pasternak, J. and Straus, N.A. (1979) Characterization of the genome of the free-living nematode Panagrellus silusiae: Absence of short period interspersion. Biochemistry 18:245-251.

* Belosevic, M. and Dick, T.A. (1979) Trichinella spiralis: Comparison of stages in host intestine with those of an arctic Trichinella sp. Expl. Parasit. 48:432-446.

Benton, W.D. and Davis, R.W. (1977) Screening λ gt recombinant clones by hybridization to single plaques in situ. Science 196:180-182.

Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513-1523.

* Boev, S.N., Britov, V.A., and Orlov, I.V. (1979) Species composition of trichinellae. Wiad. parazyt. 25:495-503.

Bostock, C.J. (1980) A function for satellite DNA? Trends Biochem. Sci. 5:117-119.

Botchan, M., McKenna, G., and Sharp, P.A. (1974) Cleavage of mouse DNA by a restriction enzyme as a clue to the arrangement of genes. Cold Spring Harbor Symp. Quant. Biol. 38:383-395.

* Britov, V.A., and Boev, S.N. (1972) Taxonomic rank of various strains of Trichinella and their circulation in nature. Vestn. Akad. Nauk. SSSR 28:27-32.

Britten, R.J. and Kohne, D.E. (1968) Repeated sequences in DNA. Science 161:529-540.

Chadee, K., Dick, T.A., and Faubert, G.M. (1983) Sensitivity of Trichinella sp. isolates to thiabendazole. Can. J. Zool. 61:139-146.

Chambers, A.E., Almond, N.M., Knight, M., Simpson, A.J.G., and Parkhouse, R.M.E. (1986) Repetitive DNA as a tool for the identification and comparison of nematode variants: application to Trichinella isolates. Mol. Biochem. Parasitol. 21:113-120.

Curran, J., Baillie, D.L., and Webster, J.M. (1985) Use of genomic DNA restriction fragment length differences to identify nematode species. Parasitology 90:137-144.

Dick, T.A. (1983a) Species, and infraspecific variation. In: Trichinella and Trichinosis (Campbell, W.C., ed.), pp. 31-73, Plenum Press, New York.

Dick, T. (1983b) The species problem in Trichinella. In: Concepts in Nematode Systematics. (Stone, A.R., Platt, H.M., Khalil, L.F., eds.), pp. 351-360, Academic Press, New York.

Dick, T.A. and Chadee, K. (1983) Interbreeding and gene flow in the genus Trichinella. J. Parasitol. 69:176-180.

Dick, T.A., Currans, J., and Klassen, G. (1985) Genetics and molecular biology of Trichinella. In: Sixth International Conference on Trichinellosis (Kim, C.S., ed.), State University of New York Press, Albany, N.Y. pp. 118-128.

Emmons, S.W., Rosenzweig, B., and Hirsh, D. (1980) Arrangement of repeated sequences in the DNA of the nematode Caenorhabditis elegans. J. Mol. Biol. 144:481-500.

Files, J.G. and Hirsh, D. (1981) Ribosomal DNA of Caenorhabditis elegans. J. Mol. Biol. 149:223-240.

Flavell, R. (1982) Sequence amplification, deletion and rearrangement: major sources of variation during species divergence. In: Genome Evolution. (Dover, G.A., and Flavell, R., eds.), pp. 301-323, Academic Press, New York.

Flavell, R.B., O'Dell, M., and Smith, D.B. (1979) Repeated sequence DNA comparisons between Triticum and Aegilops species. Heredity 42:309-322.

Flavell, R.B., Rimpau, J., and Smith, D.B. (1977) Repeated sequence DNA relationships in four cereal genomes. Chromosoma (Berl.) 63:205-222.

Fry, K., and Salser, W. (1977) Nucleotide sequences of HS-α satellite DNA from Kangaroo Rat Dipodomys ordii and characterization of similar sequences in other rodents. Cell 12:1069-1084.

* Garkavi, B.L. (1972) Species of Trichinella isolated from wild animals. Veterianariya 10:90-91.

Hardman, N. (1986) Structure and function of repetitive DNA in eukaryotes. Biochem. J. 234:1-11.

Horz, W., Hess, I., and Zachau, H.G. (1974) Highly regular arrangement of a restriction-nuclease-sensitive site in rodent satellite DNAs. Eur. J. Biochem. 45:501-512.

Horz, W. and Zachau, H.G. (1977) Characterization of distinct segments in mouse satellite DNA by restriction nucleases. Eur. J. Biochem. 73:383-392.

* Hulinska, D., and Shaikenov, B. (1980) Scanning electron microscopic studies on developmental adult stages of four Trichinella species. Angew. Parasitol. 21:150-158.

Kafatos, F., Jones, C., and Efstratiadis, A. (1979) Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucleic Acids Res. 7:1541-1552.

Klassen, G.R., Thiessen, J.P., and Dick, T.A. (1986) Restriction endonuclease analysis of repetitive sequences in the Trichinella genome: Three strain-specific patterns J. Parasit. 72:772-775.

Klassen, G.R., Thiessen, J.P., and Dick, T.A. (1986) A 1.7 kb DNA segment in Trichinella spiralis is highly repetitive in isolates from pigs but not in isolates from wild hosts. Mol. Biochem. Parasitol. 21:227-233.

Lewin, R. (1982) Repeated DNA still in search of a function. Science 217:621-623.

Lima-de-Faria, A., Arnason, U., Widegren, B., Essen-Moller, J., Isaksson, M., Olsson, E., and Jaworska, H. (1984) Conservation of repetitive DNA sequences in deer species studied by southern blot transfer. J. Mol. Evol. 20:17-24.

Lopes, U.G.L. and D.F. Wirth (1986) DNA probes specific for visceral leishmaniasis. Mol. Biochem. Parasitol. (in press).

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

McCutchan, T.F., Simpson, A.J.G., Mullins, J.A., Sher, A., Nash, T.E., Lewis, F., and Richards, C. (1984) Differentiation of schistosomes by species, strain, and sex by using cloned DNA markers. Proc. Natl. Acad. Sci. USA 81:889-893.

McReynolds, L.A., DeSimone, S.M., and Williams, S.A. (1986) Cloning and comparison of repeated DNA sequences from the human filarial parasite Brugia malayi and the animal parasite Brugia pahangi. Proc. Natl. Acad. Sci. USA 83:797-801.

Moritz, K.B. and Roth, G.E. (1976) Complexity of germline and somatic DNA in Ascaris. Nature 259:55-57.

* Mutafova, T., and Komandarev, S. (1976) On the karyotype of a laboratory Trichinella strain from Bulgaria. Z. Parasitenkd. 48:247-250.

Mydynsky, L.J. and Dick, T.A. (1985) The use of enzyme polymorphisms to identify genetic differences in the genus Trichinella. J. Parasitol. 71:671-677.

Nei, M., and Li, W-H. (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76:5269-5273.

Nelson, D.W. and Honda, B.M. (1985) Genes coding for 5S ribosomal RNA of the nematode Caenorhabditis elegans Gene. 38:245-251.

* Penkova, R.A., and Romanenko, L.N. (1973) The investigation of Trichinella chromosomes. Proc. All-Union K.I. Skryabin Inst. Helminthol. 20:133-142.

Richter, H. and Ennis, H.L. (1985) Characterization of a new repetitive sequence in the Dictyostelium discoideum genome. Arch. Biochem. Biophys. 242:16-22.

Rigby, P.W.J., Dieckman, M., Rhodes, C., and Berg, P. (1977) Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113:237-251.

Rimpau, J., Smith, D., and Flavell, R. (1978) Sequence organisation analysis of the wheat and rye genomes by interspecies DNA/DNA hybridisation. J. Mol. Biol. 123:327-359.

Rimpau, J., Smith, D.B., and Flavell, R.B. (1980) Sequence organisation in barley and oats chromosomes revealed by interspecies DNA/DNA hybridisation. Heredity 44:131-149.

Rose, A.M., Baillie, D.L., Candido, E.P.M., Beckenbach, K.A. and Nelson, D. (1982) The linkage mapping in cloned restriction fragment length differences in Caenorhabditis elegans. Mol. Gen. Genet. 188:286-291.

Rose, M.R. and Doolittle, W.F. (1983) Molecular biological mechanisms of speciation. Science 220:157-162.

Roth, G.E. and Moritz, K.B. (1981) Restriction enzyme analysis of the germ line limited DNA of Ascaris suum. Chromosoma 83:169-190.

Searcy, D.G. and MacInnis, A.J. (1970) Measurement by DNA renaturation of the genetic basis of parasite reduction. Evolution 24:796-806.

Shaw, D.R., Khandekar, P., Siddiqui, M.A.Q. and Ennis, H.L. (1986) The 3'-noncoding region of the chick myosin light-chain gene hybridizes to a family of repetitive sequences in the slime mold Dictyostelium discoideum. Arch. Biochem Biophys. 246:829-837.

Sim, B.K.L., Piessens, W.F., and Wirth, D.F. (1986) DNA probes specific for Brugia malayi. Mol. Biochem. Parasitol. (in press).

Singer, M.F. (1982) Highly repeated sequences in mammalian genomes. Int. Rev. Cytol. 76:67-112.

Skinner, D.M. and Triplett, L.L. (1967). The selective loss of DNA satellites on deproteinization with phenol. Biochem. Biophys. Res. Commun. 28:892-897.

Smith, G.P. (1973) Unequal crossing over and the evolution of multigene families. Cold Spring Harbor Symp. Quant. Biol. 38:507-513.

Smith, D.B. and Flavell, R.B. (1974) The relatedness and evolution of repeated nucleotide sequences in the genomes of some gramineae species. Biochem. Genet. 12:243-256.

Streek, R.E., Moritz, K.B., and Beer, K. (1982) Chromatin diminution in Ascaris suum: nucleotide sequence of the eliminated satellite DNA. Nucleic Acids Res. 10:3495-3502

Sulston, J.E. and Brenner, S. (1974) The DNA of Caenorhabditis elegans. Genetics. 77:95-104.

Teschke, C. (1985) Unusual cleavage pattern of the germ line limited satellite DNA of Parascaris equorum univalens. Naturwissenschaften 72:207-209.

Tobler, H., Muller, F., Back, E., and Aeby, P. (1985) Germ line - soma differentiation in Ascaris: A molecular approach. Experientia 41:1311-1319.

Widegren, B., Arnason, U., and Akusjarvi, G. (1985) Characteristics of a conserved 1,579-bp highly repetitive component in the killer whale, Orcinus orca. Mol. Biol. Evol. 2:411-419.

Wirth, D.F. (1986) DNA probes in the detection of parasitic infections. In: DNA probes - Applications in genetic and infectious disease and cancer. (Lerman, L.S., ed.), pp. 185-188, Cold Spring Harbor Laboratory, New York.

* Not consulted