![Universiteit Leiden - The Netherlands logo]

The zoonotic potential of Oesophagostomum bifurcum in Ghana. Epidemiological, morphological and genetic studies

Gruijter, J.M. de

CHAPTER 6

HIGH RESOLUTION DNA FINGERPRINTING BY AFLP TO STUDY THE GENETIC VARIATION AMONG *OESOPHAGOSTOMUM BIFURCUM* (NEMATODA) FROM HUMAN AND NON-HUMAN PRIMATES FROM GHANA

J. M. de Gruijter, R. B. Gasser, A. M.Polderman, V. Asigri and L. Dijkshoorn

Abstract

An AFLP approach was established to investigate genetic diversity within *Oesophagostomum bifurcum* (order Strongylida) from human and non-human primates. Evaluation of different combinations of restriction enzymes (n = 8) and primers (n = 29) demonstrated that the use of *Hind*III/*Bgl*II digested templates and primers with the selective nucleotides +AG/+AC, respectively, was the most effective for the analysis of *O. bifurcum* DNA. A total of 63 *O. bifurcum* adults from human, Patas monkey, Mona monkey and Olive baboon hosts from different geographical regions in Ghana were subjected to analysis using this method. Cluster analysis revealed four genetically distinct groups, namely *O. bifurcum* from the Patas monkey (I), from the Mona monkey (II), from the Olive baboon (III) and from humans (IV). These findings were concordant with those achieved previously using RAPD analysis and supports population genetic substructuring within *O. bifurcum* according to host species. The results demonstrated the effectiveness of the present AFLP method for establishing genetic variation within *O. bifurcum*, and indicate its applicability to other parasitic nematodes of human and/or veterinary health importance.

Introduction

In the last decades, human infection with *Oesophagostomum bifurcum* (Nematoda: Strongylida) has emerged as an important infection in northern Togo and Ghana.[25,44] Currently, it is estimated that 250 000 people are infected with this geo-helminth, and at least one million people are at risk.[23,26] The pathological effects caused by *O. bifurcum* can result in a uni-nodular disease, called the 'Dapaong tumour', which presents as a granulomatous mass in the abdominal wall or within the abdominal cavity and is usually associated with fever.[24,26,154,155] Alternatively, a multi-nodular disease can occur, which is characterized by the presence of hundreds of pea-sized nodules in the wall of the large intestine, giving rise to peritonitis, bowel obstruction and/or, in some cases, cutaneous lesions.[24,26,37] In spite of the public health importance of *O. bifurcum* in northern Togo and Ghana, there are serious gaps in our knowledge of the epidemiology and transmission of human oesophagostomiasis.

It has been suggested that non-human primates could represent reservoir hosts for human oesophagostomiasis.[39] However, in northern Ghana, there is a significant difference in the geographical distribution of the infection between humans and non-human primates (van Lieshout, unpublished observations). In addition, morphological studies have shown that there can be significant variation in the morphology (i.e., parasite length, width and length of the oesophagus) of the adults of *O. bifurcum* among species of primate hosts (chapter 3).[20] These observations have suggested the existence of population variation within *O. bifurcum* from human and non-human primates from Ghana.

While previous investigations of ribosomal and mitochondrial DNA did not reveal clear evidence of genetic substructuring within *O. bifurcum* from humans and different species of non-human primates from Ghana,[21,140] a recent study using random amplification of polymorphic DNA (RAPD) analysis provided support for the existence of different genetic groups of *O. bifurcum* according to host species.[156] These findings supported the hypothesis that humans harbour a host-affiliated variant of *O. bifurcum* compared with non-human primate species. In this study, we established and employed an AFLP approach to provide an independent data set to further test this hypothesis.

AFLP[TM] is a high stringency DNA fingerprinting method, based on the selective amplification of restriction fragments produced from genomic DNA and their subsequent display by denaturing gel electrophoretic analysis.[119] While widely applied to a range of organismal groups, including vertebrate animals, plants, fungi and bacteria,[120,157-160] to our knowledge there are but two studies describing its use for the fingerprinting of parasitic nematodes of veterinary importance, namely the abomasal nematode *Haemonchus contortus*,[122] and the bovine lungworm, *Dictyocaulus viviparus*,[121] and no investigations of parasitic nematodes of human health importance. The aim of the present study was to assess the value of

the AFLP as a tool for the analysis of genetic variation within *O. bifurcum* from human and non-human primates from different geographical regions in Ghana.

Materials and Methods

Parasites

Adult worms of *O. bifurcum* (n = 63) were obtained from humans and species of non-human primate from three different geographical regions in Ghana (see Table 1). Worms from humans were obtained from the faeces of infected patients after treatment with pyrantel pamoate, as described previously,[44] whereas worms from non-human primates were removed from the large intestine at necropsy. The worms were washed extensively in physiological saline and then stored in 70% ethanol until required for DNA isolation. Each specimen of *O. bifurcum* was identified morphologically using published keys and descriptions.[12,20,127] In addition, all individuals were identified to species by single-strand conformation polymorphism (SSCP) analysis of the second internal transcribed spacer (ITS-2) of nuclear ribosomal DNA (rDNA), as described previously.[156]

Table 1   Adult *Oesophagostomum bifurcum* specimens used for AFLP analysis.

Sample code	Host (individual)	Location (region, country)
H30, H31, H38, H39	Human (A)	Upper East Region, Ghana
H32, H33, H49, H50	Human (B)	Upper East Region, Ghana
H34-H37	Human (C)	Upper East Region, Ghana
H40-H43	Human (D)	Upper East Region, Ghana
H45-H48	Human (E)	Upper East Region, Ghana
H51-H53	Human (J)	Upper East Region, Ghana
H54-H57, H62	Human (K)	Upper East Region, Ghana
H58-H61	Human (L)	Upper East Region, Ghana
PMN2, PMN6, PMN8, PMN9	Patas monkey (F)	Upper East Region, Ghana
PMD1-3, PMD7, PMD15	Patas monkey (G)	Northern region
B2-B5	Olive baboon (H)	Northern region
B6-B17	Olive baboon (N)	Northern region
M4-M8, M10	Mona monkey (I)	Brong Ahafo region, Ghana

Isolation of genomic DNA

Genomic DNA was isolated from *O. bifurcum* individuals by sodium dodecyl-sulphate/proteinase K digestion,[134] purified over spin columns (Wizard™ DNA Clean-Up, Promega, Madison, WI, USA) and then eluted into 30 µl of H_2O. Isolation and purification of DNA from the large intestinal content from non-infected hosts (i.e., control-DNA samples) were carried out as described previously.[46] All DNA samples were treated with RNAse (0.25 mg/ml) (Boehringer, Almere, The Netherlands).

Digestion of genomic DNA and ligation of adapters

Genomic DNA (~10 ng) was double-digested with restriction enzymes for 4 h at 37°C in a volume of 20 µl using 10 U of each restriction enzyme (Roche, Basel, Switzerland) and 1 mg/ml of bovine serum albumin (BSA). The digestion product was precipitated with ethanol and resuspended in 5 µl of H_2O. Subsequently, 5 µl of a mixture containing T4 ligation buffer (50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol (DTT), 1 mM ATP, 25 µg/ml BSA), 10 U of T4 DNA-ligase (New England Biolabs, Beverly, MA, USA) and 10 pmol of adapter (100 pmol for *Mse*I or *Taq*I adapters) were added, followed by incubation for 16 h at 15°C. The adapter sequences were as described previously.[119,161,162] After ligation the samples were stored at 4°C until required for AFLP analysis.

AFLP analysis

AFLP analysis was carried out as recommended for organisms with complex genomes.[119] Pre-selective amplification was performed in 20 µl volumes using 30 ng of each 'core primer' (Table 2), 2.5 mM dNTPs, 1.5 mM $MgCl_2$, 1.25 U *Taq* polymerase (Promega) and 10 µl of ligation product. Amplification was performed in an iCycler (Biorad, Hercules, CA, USA) using the following cycling profile: 2 min at 94°C, 30 cycles of 30 sec at 94°C, 1 min at 56°C and 1 min at 72°C. Pre-selective amplicons were resolved in 1.5% agarose-TBE (65 mM Tris-HCl, 27 mM boric acid and 1 mM EDTA, pH 9; Bio-Rad, Hercules, CA) gels prior to selective amplification. For the selective amplification, 10 µl of a 1/100 dilution of the pre-selective amplicon were used in a total volume of 20 µl. AFLP reactions were performed as described for the pre-selective amplification, with the exception that 5 ng of a Cy5-labelled and 30 ng of unlabeled primer were used, and that 1 or 2 selective nucleotides was/were added to the 3'-end of each primer (Table 2). The selective amplification consisted of 12 cycles of 30 sec at 94°C, 30 sec, commencing at 65°C and subsequently reducing the temperature each cycle by 0.7°C, and 1 min at 72°C, followed by 23 cycles of 30 sec at 94°C, 30 sec at 56°C and 1 min at 72°C. For electrophoresis, an aliquot (10 µl) of each PCR product was mixed with 3 µl of formamide loading dye (Amersham Pharmacia Biotech, Roosendaal, The Netherlands), heated for 5 min at 94°C, and snap cooled on ice. Then, 4 µl of each product were loaded on to a denaturing polyacrylamide gel (ReproGel High Resolution; Amersham Pharmacia Biotech) and subjected

to electrophoresis in TBE buffer (0.1 M Tris, 83 mM boric acid, 1 mM EDTA, pH 8) for 500 min at 30 W at 55°C using an ALFexpress I DNA analysis system (Amersham Pharmacia Biotech). The Cy5-labeled ALFexpress sizer 50-500 (Amersham Pharmacia Biotech) was used as a molecular weight marker. The resultant peak patterns were converted to TIFF files and analysed employing the Bionumerics 2.0 software package (Applied Maths, Sint-Martens-Latem, Belgium). The optimal settings for position tolerance (1.0) and optimization (0.8) were calculated automatically. Similarities between banding patterns were determined using the Pearson product moment coefficient (r), and a dendrogram was constructed based on the analysis of data using the unweighted pair group method using arithmetic averages (UPGMA). The reliability of the clusters in the dendrogram was determined by calculating the cophenetic values, which represent the correlation between the calculated similarities and the dendrogram-derived similarities.[163]

Results

Optimisation of AFLP for O. bifurcum DNA

To achieve useful polymorphic and reproducible AFLP banding patterns, different restriction enzyme and primer combinations were tested and PCR conditions were optimised for *O. bifurcum* DNA.

Evaluation of different restriction enzyme and primer combinations

Eight different restriction enzyme combinations (*Eco*RI/*Mse*I, *Eco*RI/*Taq*I, *Eco*RI/*Apa*I, *Eco*RI/*Bgl*II, EcoRI/*Hind*III, *Hind*III/*Taq*I, *Hind*III/*Apa*I or *Hind*III/*Bgl*II) and 29 primer sets (Table 2) were evaluated for their ability to give reproducible, polymorphic AFLP banding patterns among individual *O. bifurcum* adults. Initially, the restriction enzyme combination *Eco*RI/*Mse*I was investigated because it had been used successfully for the AFLP analysis of a range of organisms,[164] including individual adults of the gastric nematode *Haemonchus contortus* (order Strongylida).[122] However, AFLP analysis of *O. bifurcum* DNA digested with this restriction enzyme combination resulted in patterns consisting of only small (~100-150 bp), unevenly distributed bands, with a low level of polymorphism among individuals. Similar results were obtained using each of the other combinations of 4-6 base cutters (i.e., *Eco*RI/*Taq*I and *Hind*III/*Taq*I) tested. Thus, these restriction enzyme combinations were ineffective for *O. bifurcum* DNA. Of the combinations of the 6-6 base cutters tested, those with AT-rich recognition sequences (i.e., *Eco*RI/*Bgl*II, *Eco*RI/*Hind*III and *Hind*III/*Bgl*II) resulted in patterns, which were more polymorphic compared with those with GC-rich recognition sequences (i.e., *Eco*RI/*Apa*I and *Hind*III/*Apa*I). The addition of selective nucleotides to the core primers led to a reduction in the number of amplicons obtained via selective amplification. While the AFLP

banding patterns obtained using primer extensions of one and/or two nucleotides (+1/+1; +1/+2) were difficult to analyse because of a great number of unevenly distributed bands, the use of primers which were both extended by two selective nucleotides (+2/+2) gave patterns which were well suited for analysis. Banding patterns were more polymorphic when one or more G's or C's were used to extend each primer. Based on the evaluation, the *Hind*III/*Bgl*II restriction enzyme combination, using the primer extension +AG/+AC, respectively (i.e., *Hind*III+AG/*Bgl*II+AC), was considered to be most effective. Banding patterns produced under these conditions consisted of ~60-70 discrete bands, ranging in size from 50-500 bp, and were shown to be reproducible on consecutive days using the same amplicons and using amplicons produced on different days.

Template DNA

The amount of genomic DNA obtained from individual *O. bifurcum* adults after purification was limited (~100 ng). Only part of this DNA could be used for analysis, such that a sufficient

Table 2 Restriction enzyme and selective nucleotide combinations tested in this study. A, T, C and G are the selective nucleotides added at the 3' end of the core primers for the selective amplification.

	*Mse*I	*Taq*I	*Apa*I	*Bgl*II	*Hind*III
*Eco*RI	A/C	AC/GC	AA/GA	A/A	**A/A**
	A/CC	AC/AT	AA/AT	A/AC	A/AC
	AA/CC			AA/AC	AA/AT
	AC/CT			AC/GA	GC/AC
	AC/CC			GC/AC	
*Hind*III		AG/GC	AG/GA	A/A	
		AG/AT		A/AC	
		GC/AT		AG/AG	
				CG/AC	

Sequences of the core primers used in the pre-selective amplification:
*Eco*RI	5'-Cy5-GAC TGC GTA CCA ATT C-'3
*Mse*I	5'-Cy5-GAT GAG TCC TGA GTA A-'3
*Taq*I	5'-Cy5-GAT GAG TCC TGA CCG A-'3
*Apa*I	5'-Cy5-GAC TGC GTA CAG GCC C-'3
*Bgl*II	5'-Cy5-GAG TAC ACT GTC GAT CT-'3
*Hind*III	5'-Cy5-GAC TGC GTA CCA GCT T-'3

amount of DNA remained to assess the reproducibility of the results. For these reasons, the minimum amount of *O. bifurcum* template DNA with which clear, reproducible AFLP banding patterns were produced was assessed by titration. After purification, genomic DNA of *O. bifurcum* individuals was eluted in a volume of 30 µl. Subsequently, 1, 3, 5, 10 and 15 µl of this DNA solution were used for the AFLP analysis of *O. bifurcum* DNA using the restriction enzyme and primer combination *Hind*III+AG/*Bgl*II+AC. The minimum amount of DNA with which reproducible AFLP banding patterns were obtained was 3 µl (~10 ng), and this amount was chosen for analysis.

PCR conditions

The PCR conditions for the AFLP analysis of *O. bifurcum* DNA differed from those described previously for DNA from organisms with complex genomes[119] in that 30 cycles (instead of 20) were necessary to produce sufficient pre-selective amplicon for further analysis. Pre-selective amplification performed for 20 or 25 cycles resulted in amplicons, which could not be detected on 1.5% agarose gels, and resulted in irreproducible amplicons upon selective amplification.

The pre-selective amplicons were diluted before being subjected to selective amplification to reduce 'background' on electrophoretic gels. Several dilutions were tested, and the most effective one was chosen for further analyses. Pre-selective amplicons were diluted 1/10, 1/20, 1/50, 1/100, 1/150 or 1/200 in H_2O and then tested in the selective amplification. A 1/100 dilution was demonstrated to be most effective and resulted in reproducible AFLP banding patterns with minimum background.

AFLP analysis of O. bifurcum individuals from human and non-human primates

DNA samples from 63 individual adults of *O. bifurcum* (32 from human, 16 from the Olive baboon, 9 from the Patas monkey and 6 from the Mona monkey) (Table 1) were subjected to AFLP analysis using the restriction enzyme/primer combination *Hind*III+AG/*Bgl*II+AC. The AFLP banding patterns showed substantial polymorphism among individuals from each of the four host species (see Fig. 1). Some amplicons were produced from the no-DNA control samples and samples containing DNA isolated from large intestinal contents from non-infected hosts (i.e., human, Mona monkey, Patas monkey or Olive baboon), but none of them were the same in their position on gels as those produced from any of the *O. bifurcum* individuals examined. Banding patterns consisting of ~60-70 discrete bands (ranging in size from 50-500 bp) were produced for each of the 63 *O. bifurcum* adults analysed. Analysis of each sample on a different gel on a different day established the reproducibility of each profile. Bands in the range of 100-450 bp had a reproducibility of 100%, and thus were used for cluster analysis. After normalisation, background subtraction and optimisation of the converted TIFF-files, cluster analysis was performed and a dendrogram constructed to depict the genetic diversity
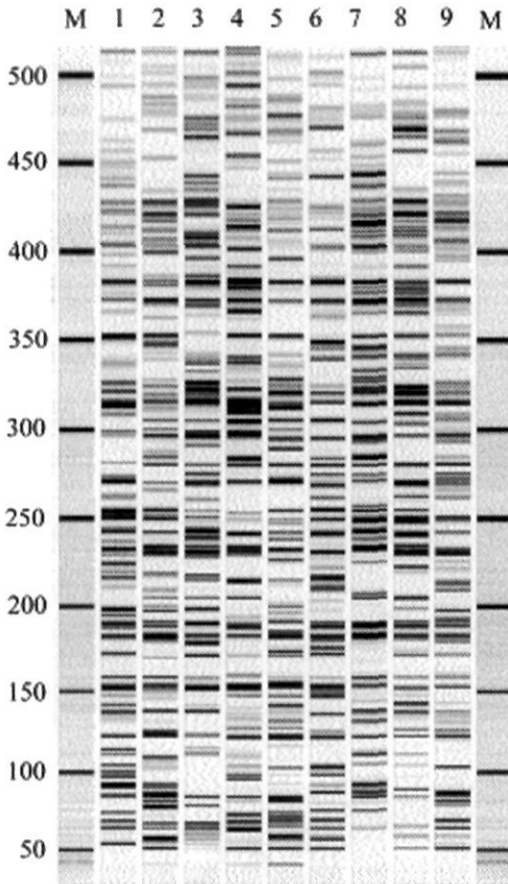
Figure 1 Example of the AFLP banding patterns obtained for *Oesophagostomum bifurcum* from humans from Ghana (lanes 1-9) using the restriction enzyme and primer combination *Hind*III+AG / *Bgl*II+AC. M represents the Alfexpress sizer 50-500 molecular weight marker. The numbers on the left indicate the fragment size in bp.

among all 63 samples analysed (Fig. 2). The Pearson product moment coefficient (*r*) values among them, as derived from the similarity matrix (not shown), ranged from 35-85%. The dendrogram consisted of four main clusters (I, II, III and IV), which were supported by cophenetic values ranging from 72-92%. Cluster I contained all 9 *O. bifurcum* individuals from the Patas monkey, cluster II comprised all 6 specimens from the Mona monkey, cluster III included all 16 from the Olive baboon and cluster IV contained all 32 derived from humans. There was no correlation between the clusters of *O. bifurcum* and the geographical origin (Table 1) of the hosts.

Discussion

The present study demonstrated clearly that the optimum conditions for AFLP analysis of *O. bifurcum* differed from those used by other workers for other invertebrates[120,159,161] and even from those described previously for other parasitic nematodes, such as *Haemonchus contortus*,[122] which is within the same order as *O. bifurcum*. Of the different restriction enzyme and primer combinations tested, *Hind*III+AG/*Bgl*II+AC was shown to be most effective for *O. bifurcum* based on the high number of evenly distributed bands and the high level of polymorphism detected among individuals. Also, the number of cycles (n = 30) in the pre-selective amplification had to be modified and pre-selective amplicons diluted (1/100) before being subjected to the selective amplification to obtain reproducible banding patterns with minimum background on electrophoretic gels.

Although there are no consensus guidelines for the evaluation of PCR-based DNA fingerprinting methods in parasitology, the criteria by which such methods are usually evaluated in microbiology relate to 'typeability' (i.e., the proportion of tested samples which produce a fingerprint), reproducibility, discriminating power and 'typing system concordance' (i.e., comparison of the results with independent data sets obtained by using other typing methods).[165,166] Also, the convenience (i.e., simplicity and rapidity) of the method is of significance.[165,166] These criteria were all considered in the present study.

Informative, polymorphic banding patterns were obtained for all 63 *O. bifurcum* individuals analysed, achieving a typeability of 100%. The reproducibility of results was 100% for all bands of 100-450 bp. The production of a unique AFLP pattern for each sample showed the high discrimination power of the method. Cluster analysis of the data resulted in a dendogram consisting of four main clusters, allowing the discrimination among *O. bifurcum* from Patas monkey (cluster I), Mona monkey (cluster II), Olive baboon (cluster III) and human (cluster IV). While the dendrogram provided strong support for the clustering of *O. bifurcum* according to host, it did not reveal a relationship between the clustering and geographical origin (Table 1) of the host species. The latter finding was supported by the following observations. (i) *O. bifurcum* specimens from Patas monkeys from two different geographical regions (i.e., Upper East and Northern region) grouped together (ii) *O. bifurcum* from human from the Upper East region grouped into a distinct cluster compared with *O. bifurcum* from a Patas monkey from the same region, and (iii) *O. bifurcum* from the Olive baboon and the Patas monkey, both from the Northern region, were each represented in a different cluster. This lack of association with geographical origin is not surprising, given that the distances between the different regions are relatively short (i.e., 500-1000 km) and that these regions are not isolated geographically from one other.
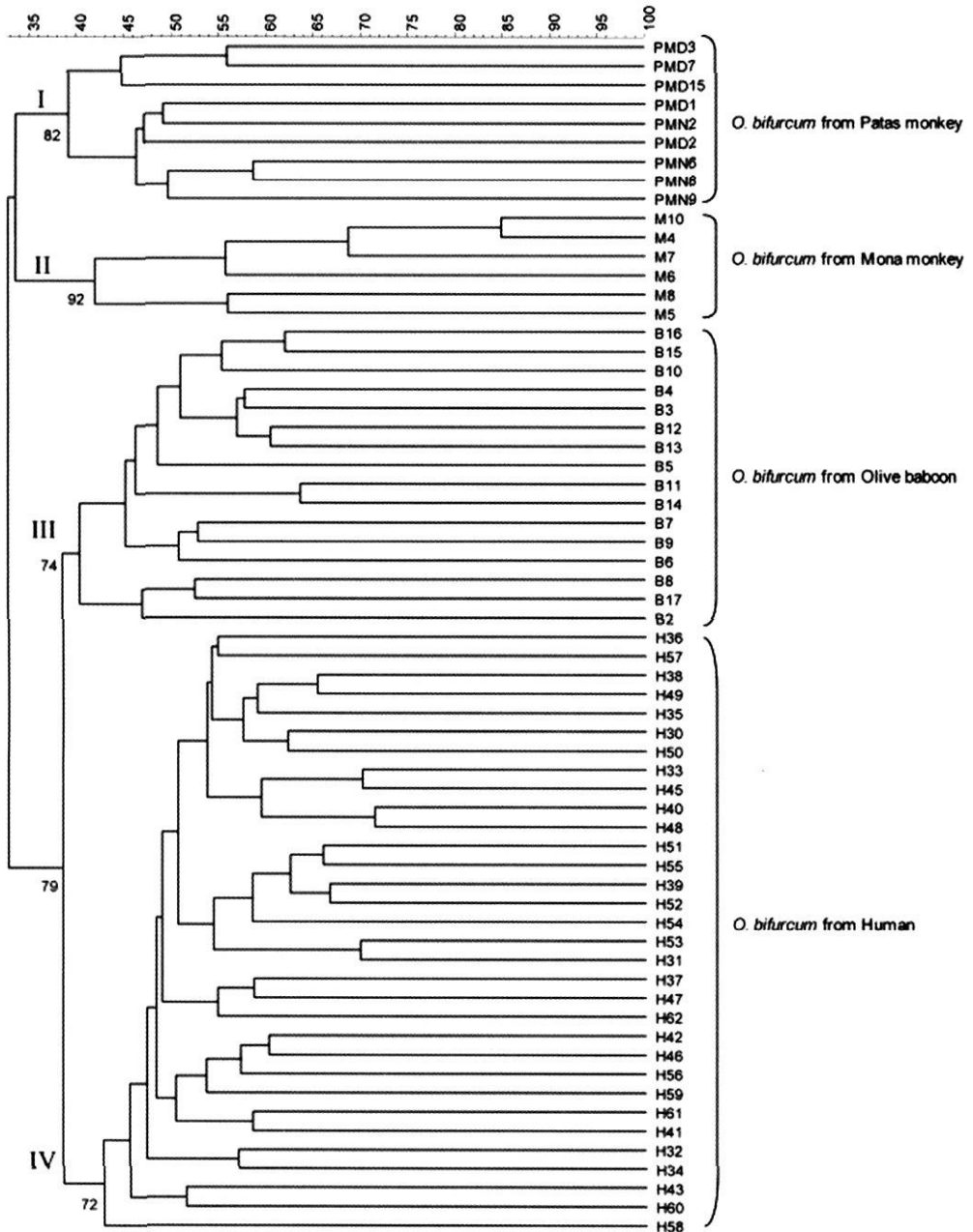
Figure 2 Dendrogram based on AFLP analysis of 63 *Oesophagostomum bifurcum* from human, Patas monkey, Mona monkey, and Olive baboon using *Hind*III + AC/*Bgl*II + AC. The scale represents the Pearson's product-moment correlation coefficient (*r*) in percentage. The numbers represent cophenetic values.

The results of the cluster analysis were concordant with those found in a previous investigation using RAPD data for a set of 41 individuals of *O. bifurcum* from the same host species in Ghana.[156] However, a difference between the dendrograms based on AFLP and that on RAPD data was that the parasite from humans and from the Olive baboon grouped to the exclusion of the parasite from the Mona monkey and Patas monkey. This difference in topology between the two dendrograms could be associated with the fact that each method screens different parts of the *O. bifurcum* genome (which may have evolved at different rates and thus contain distinct levels of genetic variation) and/or may relate to the number of *O. bifurcum* individuals used in each study. Nonetheless, for both studies, all *O. bifurcum* individuals from each of the four host species clearly represented a distinct genetic cluster.

From a technical perspective, AFLP analysis is relatively simple, and the procedure (i.e., digestion, ligation, amplification and electrophoresis) can be performed within two to three days. The use of relatively long primers (16-19 mer) and high annealing temperatures (56-65°C) in the PCR makes this method more robust and reliable compared with some other PCR-based fingerprinting methods, such as low-stringency RAPD. Moreover, the results of this study showed that the AFLP banding patterns were of a high quality. This attribute makes them suitable for computerized comparison, which is of importance when large numbers of samples from different laboratories or from different gels are compared and for the formation of databases.

The present findings demonstrate that the AFLP is well-suited for population genetic and molecular epidemiological studies of *O. bifurcum*. For instance, it may be used to investigate host-parasite co-evolution. The present AFLP dendogram showed that *O. bifurcum* from Patas monkey (cluster I) and *O. bifurcum* from Mona monkey (cluster II), two species of primate belonging to the same tribe within the subfamily Cercopithecinae of the family Cercopithecidae,[167] were genetically more similar to one another than they were to that of human or Olive baboon. This finding is concordant with a previous analysis of RAPD data[156] and suggests that the parasite evolves with the species of primate host. The present AFLP approach could be used to screen large numbers of *O. bifurcum* specimens from a wide range of closely and distantly related primate host species to test this proposal. Furthermore, AFLP may be a useful tool to identify sources of *O. bifurcum* infection.

Given that the presence of infective *O. bifurcum* larvae in grassland in Ghana has been demonstrated recently (Polderman, unpublished observations), comparative AFLP analysis of large numbers of individual *O. bifurcum* larvae from the environment (i.e., grassland) and those obtained from humans could provide improved insight into modes of transmission, and would have implications for prevention and control. While the focus of the present study was on *O. bifurcum*, the present AFLP approach should also be useful for investigating important questions regarding the population genetics and systematics of other key bursate nematodes of human health significance, such as hookworms. For instance, the approach would be

particularly applicable to test the hypothesis, based on nuclear and mitochondrial DNA data sets,[65,168-170] that *N. americanus* represents a complex of multiple species.

Together with the previous RAPD data,[156] the present AFLP data provide strong support that *O. bifurcum* from human and different species of non-human primates represent genetically distinct groups. This information could have important public health implications for regions in Ghana where human and non-human primate species (i.e., Patas monkey, Mona monkey and Olive baboon) live in sympatry. However, while it is suggested that *O. bifurcum* from non-human primates may not have a high infectivity to humans, it cannot yet be excluded that cross-infection does take place. Biological studies to test the ability of the different genetic genotypes to infect different species of primates have to be conducted. Previously, such an experimental study[138] revealed that *O. bifurcum* from non-human primates appeared to be poorly susceptible to infection with *O. bifurcum* from humans. However, the species of primate (*Macaca mulatta*) used in that study was distinct from those included in the present study and present in Togo and Ghana. The molecular evidence (i.e., the RAPD and AFLP data) that different genotypes of *O. bifurcum* infect different species of primate host stimulates the need for cross-infection experiments using the species of non-human primates studied herein. This could elucidate the role that non-human primates play in the transmission of human oesophagostomiasis.