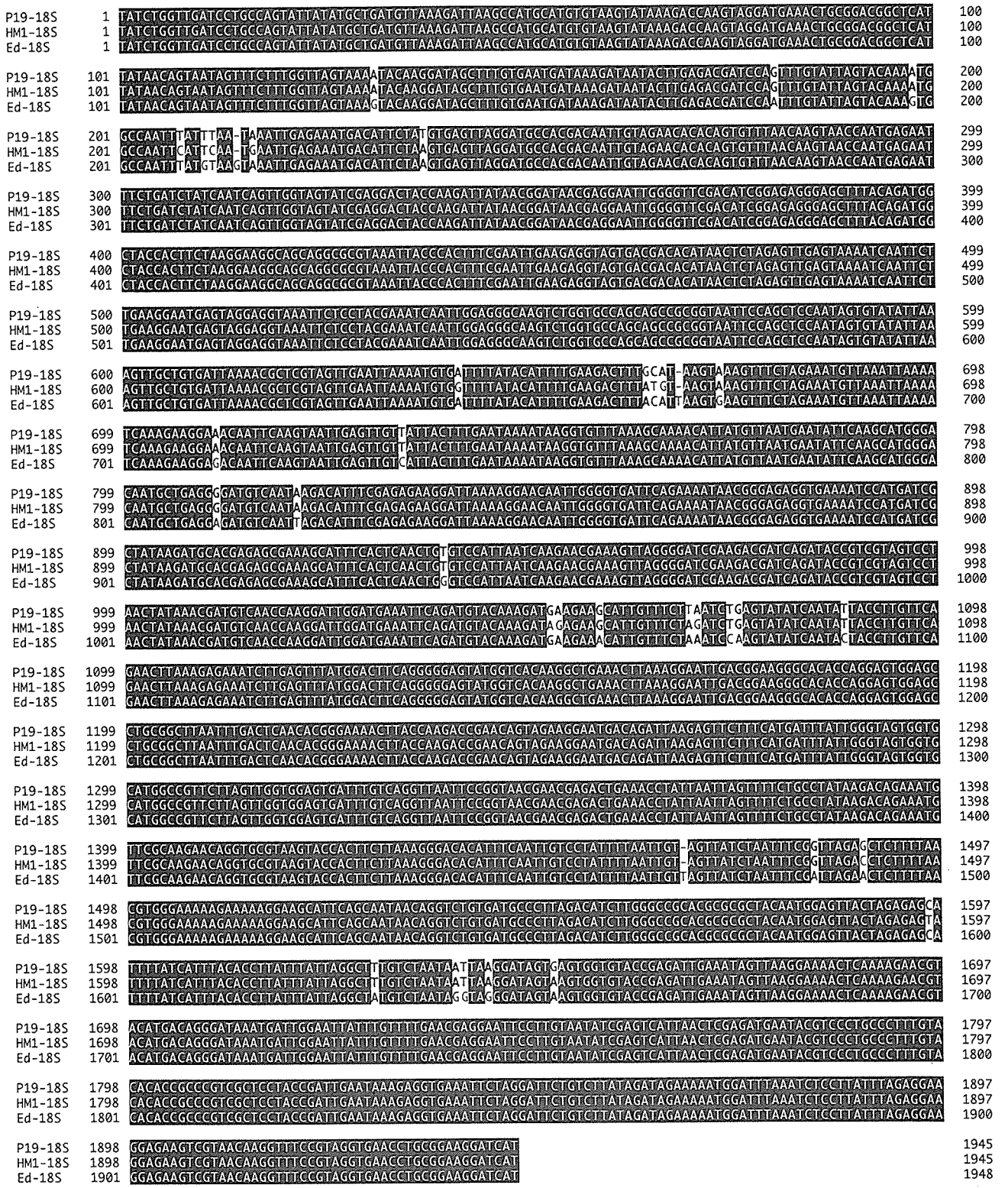
P19-18S 1 TATCTGGTTGATCCTGCCAGTATTATATGCTGATGTTAAAGATTAAGCCATGCATGTGTAAGTATAAAGACCAAGTAGGATGAAACTGCGGACGGCTCAT 100
HM1-18S 1 TATCTGGTTGATCCTGCCAGTATTATATGCTGATGTTAAAGATTAAGCCATGCATGTGTAAGTATAAAGACCAAGTAGGATGAAACTGCGGACGGCTCAT 100
Ed-18S 1 TATCTGGTTGATCCTGCCAGTATTATATGCTGATGTTAAAGATTAAGCCATGCATGTGTAAGTATAAAGACCAAGTAGGATGAAACTGCGGACGGCTCAT 100
P19-18S 101 TATAACAGTAATAGTTTCTTTGGTTAGTAAAATACAAGGATAGCTTTGTGAATGATAAAGATAATACTTGAGACGATCCAGTTTGTATTAGTACAAAATG 200
HM1-18S 101 TATAACAGTAATAGTTTCTTTGGTTAGTAAAATACAAGGATAGCTTTGTGAATGATAAAGATAATACTTGAGACGATCCAGTTTGTATTAGTACAAAATG 200
Ed-18S 101 TATAACAGTAATAGTTTCTTTGGTTAGTAAAGTACAAGGATAGCTTTGTGAATGATAAAGATAATACTTGAGACGATCCAATTTGTATTAGTACAAAGTG 200
P19-18S 201 GCCAATTTATTTAA-TAAATTGAGAAATGACATTCTATGTGAGTTAGGATGCCACGACAATTGTAGAACACACAGTGTTTAACAAGTAACCAATGAGAAT 299
HM1-18S 201 GCCAATTCATTCAA-TGAATTGAGAAATGACATTCTAAGTGAGTTAGGATGCCACGACAATTGTAGAACACACAGTGTTTAACAAGTAACCAATGAGAAT 299
Ed-18S 201 GCCAATTTATGTAAGTAAATTGAGAAATGACATTCTAAGTGAGTTAGGATGCCACGACAATTGTAGAACACACAGTGTTTAACAAGTAACCAATGAGAAT 300
P19-18S 300 TTCTGATCTATCAATCAGTTGGTAGTATCGAGGACTACCAAGATTATAACGGATAACGAGGAATTGGGGTTCGACATCGGAGAGGGAGCTTTACAGATGG 399
HM1-18S 300 TTCTGATCTATCAATCAGTTGGTAGTATCGAGGACTACCAAGATTATAACGGATAACGAGGAATTGGGGTTCGACATCGGAGAGGGAGCTTTACAGATGG 399
Ed-18S 301 TTCTGATCTATCAATCAGTTGGTAGTATCGAGGACTACCAAGATTATAACGGATAACGAGGAATTGGGGTTCGACATCGGAGAGGGAGCTTTACAGATGG 400
P19-18S 400 CTACCACTTCTAAGGAAGGCAGCAGGCGCGTAAATTACCCACTTTCGAATTGAAGAGGTAGTGACGACACATAACTCTAGAGTTGAGTAAAATCAATTCT 499
HM1-18S 400 CTACCACTTCTAAGGAAGGCAGCAGGCGCGTAAATTACCCACTTTCGAATTGAAGAGGTAGTGACGACACATAACTCTAGAGTTGAGTAAAATCAATTCT 499
Ed-18S 401 CTACCACTTCTAAGGAAGGCAGCAGGCGCGTAAATTACCCACTTTCGAATTGAAGAGGTAGTGACGACACATAACTCTAGAGTTGAGTAAAATCAATTCT 500
P19-18S 500 TGAAGGAATGAGTAGGAGGTAAATTCTCCTACGAAATCAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGTGTATATTAA 599
HM1-18S 500 TGAAGGAATGAGTAGGAGGTAAATTCTCCTACGAAATCAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGTGTATATTAA 599
Ed-18S 501 TGAAGGAATGAGTAGGAGGTAAATTCTCCTACGAAATCAATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGTGTATATTAA 600
P19-18S 600 AGTTGCTGTGATTAAAACGCTCGTAGTTGAATTAAAATGTGATTTTATACATTTTGAAGACTTTGCAT-AAGTAAAGTTTCTAGAAATGTTAAATTAAAA 698
HM1-18S 600 AGTTGCTGTGATTAAAACGCTCGTAGTTGAATTAAAATGTGGTTTTATACATTTTGAAGACTTTATGT-AAGTAAAGTTTCTAGAAATGTTAAATTAAAA 698
Ed-18S 601 AGTTGCTGTGATTAAAACGCTCGTAGTTGAATTAAAATGTGATTTTATACATTTTGAAGACTTTACATTAAGTGAAGTTTCTAGAAATGTTAAATTAAAA 700
P19-18S 699 TCAAAGAAGGAAACAATTCAAGTAATTGAGTTGTTATTACTTTGAATAAAATAAGGTGTTTAAAGCAAAACATTATGTTAATGAATATTCAAGCATGGGA 798
HM1-18S 699 TCAAAGAAGGAAACAATTCAAGTAATTGAGTTGTTATTACTTTGAATAAAATAAGGTGTTTAAAGCAAAACATTATGTTAATGAATATTCAAGCATGGGA 798
Ed-18S 701 TCAAAGAAGGAGACAATTCAAGTAATTGAGTTGTCATTACTTTGAATAAAATAAGGTGTTTAAAGCAAAACATTATGTTAATGAATATTCAAGCATGGGA 800
P19-18S 799 CAATGCTGAGGGGATGTCAATAAGACATTTCGAGAGAAGGATTAAAAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTGAAAATCCATGATCG 898
HM1-18S 799 CAATGCTGAGGGGATGTCAATAAGACATTTCGAGAGAAGGATTAAAAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTGAAAATCCATGATCG 898
Ed-18S 801 CAATGCTGAGGAGATGTCAATTAGACATTTCGAGAGAAGGATTAAAAGGAACAATTGGGGTGATTCAGAAAATAACGGGAGAGGTGAAAATCCATGATCG 900
P19-18S 899 CTATAAGATGCACGAGAGCGAAAGCATTTCACTCAACTGTGTCCATTAATCAAGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCT 998
HM1-18S 899 CTATAAGATGCACGAGAGCGAAAGCATTTCACTCAACTGTGTCCATTAATCAAGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCT 998
Ed-18S 901 CTATAAGATGCACGAGAGCGAAAGCATTTCACTCAACTGGGTCCATTAATCAAGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCCT 1000
P19-18S 999 AACTATAAACGATGTCAACCAAGGATTGGATGAAATTCAGATGTACAAAGATGAAGAAGCATTGTTTCTTAATCTGAGTATATCAATATTACCTTGTTCA 1098
HM1-18S 999 AACTATAAACGATGTCAACCAAGGATTGGATGAAATTCAGATGTACAAAGATAGAGAAGCATTGTTTCTAGATCTGAGTATATCAATATTACCTTGTTCA 1098
Ed-18S 1001 AACTATAAACGATGTCAACCAAGGATTGGATGAAATTCAGATGTACAAAGATGAAGAAACATTGTTTCTAAATCCAAGTATATCAATACTACCTTGTTCA 1100
P19-18S 1099 GAACTTAAAGAGAAATCTTGAGTTTATGGACTTCAGGGGGAGTATGGTCACAAGGCTGAAACTTAAAGGAATTGACGGAAGGGCACACCAGGAGTGGAGC 1198
HM1-18S 1099 GAACTTAAAGAGAAATCTTGAGTTTATGGACTTCAGGGGGAGTATGGTCACAAGGCTGAAACTTAAAGGAATTGACGGAAGGGCACACCAGGAGTGGAGC 1198
Ed-18S 1101 GAACTTAAAGAGAAATCTTGAGTTTATGGACTTCAGGGGGAGTATGGTCACAAGGCTGAAACTTAAAGGAATTGACGGAAGGGCACACCAGGAGTGGAGC 1200
P19-18S 1199 CTGCGGCTTAATTTGACTCAACACGGGAAAACTTACCAAGACCGAACAGTAGAAGGAATGACAGATTAAGAGTTCTTTCATGATTTATTGGGTAGTGGTG 1298
HM1-18S 1199 CTGCGGCTTAATTTGACTCAACACGGGAAAACTTACCAAGACCGAACAGTAGAAGGAATGACAGATTAAGAGTTCTTTCATGATTTATTGGGTAGTGGTG 1298
Ed-18S 1201 CTGCGGCTTAATTTGACTCAACACGGGAAAACTTACCAAGACCGAACAGTAGAAGGAATGACAGATTAAGAGTTCTTTCATGATTTATTGGGTAGTGGTG 1300
P19-18S 1299 CATGGCCGTTCTTAGTTGGTGGAGTGATTTGTCAGGTTAATTCCGGTAACGAACGAGACTGAAACCTATTAATTAGTTTTCTGCCTATAAGACAGAAATG 1398
HM1-18S 1299 CATGGCCGTTCTTAGTTGGTGGAGTGATTTGTCAGGTTAATTCCGGTAACGAACGAGACTGAAACCTATTAATTAGTTTTCTGCCTATAAGACAGAAATG 1398
Ed-18S 1301 CATGGCCGTTCTTAGTTGGTGGAGTGATTTGTCAGGTTAATTCCGGTAACGAACGAGACTGAAACCTATTAATTAGTTTTCTGCCTATAAGACAGAAATG 1400
P19-18S 1399 TTCGCAAGAACAGGTGCGTAAGTACCACTTCTTAAAGGGACACATTTCAATTGTCCTATTTTAATTGT-AGTTATCTAATTTCGGTTAGAGCTCTTTTAA 1497
HM1-18S 1399 TTCGCAAGAACAGGTGCGTAAGTACCACTTCTTAAAGGGACACATTTCAATTGTCCTATTTTAATTGT-AGTTATCTAATTTCGGTTAGACCTCTTTTAA 1497
Ed-18S 1401 TTCGCAAGAACAGGTGCGTAAGTACCACTTCTTAAAGGGACACATTTCAATTGTCCTATTTTAATTGTTAGTTATCTAATTTCGATTAGAACTCTTTTAA 1500
P19-18S 1498 CGTGGGAAAAAGAAAAAGGAAGCATTCAGCAATAACAGGTCTGTGATGCCCCTTAGACATCTTGGGCCGCACGCGCGCTACAATGGAGTTACTAGAGAGCA 1597
HM1-18S 1498 CGTGGGAAAAAGAAAAAGGAAGCATTCAGCAATAACAGGTCTGTGATGCCCCTTAGACATCTTGGGCCGCACGCGCGCTACAATGGAGTTACTAGAGAGTA 1597
Ed-18S 1501 CGTGGGAAAAAGAAAAAGGAAGCATTCAGCAATAACAGGTCTGTGATGCCCCTTAGACATCTTGGGCCGCACGCGCGCTACAATGGAGTTACTAGAGAGCA 1600
P19-18S 1598 TTTTATCATTTACACCTTATTTATTAGGCTTTGTCTAATAATTAAGGATAGTGAGTGGTGTACCGAGATTGAAATAGTTAAGGAAAACTCAAAAGAACGT 1697
HM1-18S 1598 TTTTATCATTTACACCTTATTTATTAGGCTTTGTCTAATAATTAAGGATAGTAAGTGGTGTACCGAGATTGAAATAGTTAAGGAAAACTCAAAAGAACGT 1697
Ed-18S 1601 TTTTATCATTTACACCTTATTTATTAGGCTATGTCTAATAGGTAGGGATAGTAAGTGGTGTACCGAGATTGAAATAGTTAAGGAAAACTCAAAAGAACGT 1700
P19-18S 1698 ACATGACAGGGATAAATGATTGGAATTATTTGTTTTGAACGAGGAATTCCTTGTAATATCGAGTCATTAACTCGAGATGAATACGTCCCTGCCCTTTGTA 1797
HM1-18S 1698 ACATGACAGGGATAAATGATTGGAATTATTTGTTTTGAACGAGGAATTCCTTGTAATATCGAGTCATTAACTCGAGATGAATACGTCCCTGCCCTTTGTA 1797
Ed-18S 1701 ACATGACAGGGATAAATGATTGGAATTATTTGTTTTGAACGAGGAATTCCTTGTAATATCGAGTCATTAACTCGAGATGAATACGTCCCTGCCCTTTGTA 1800
P19-18S 1798 CACACCGCCCGTCGCTCCTACCGATTGAATAAAGAGGTGAAATTCTAGGATTCTGTCTTATAGATAGAAAAATGGATTTAAATCTCCTTATTTAGAGGAA 1897
HM1-18S 1798 CACACCGCCCGTCGCTCCTACCGATTGAATAAAGAGGTGAAATTCTAGGATTCTGTCTTATAGATAGAAAAATGGATTTAAATCTCCTTATTTAGAGGAA 1897
Ed-18S 1801 CACACCGCCCGTCGCTCCTACCGATTGAATAAAGAGGTGAAATTCTAGGATTCTGTCTTATAGATAGAAAAATGGATTTAAATCTCCTTATTTAGAGGAA 1900
P19-18S 1898 GGAGAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCAT 1945
HM1-18S 1898 GGAGAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCAT 1945
Ed-18S 1901 GGAGAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCAT 1948

Fig. 2

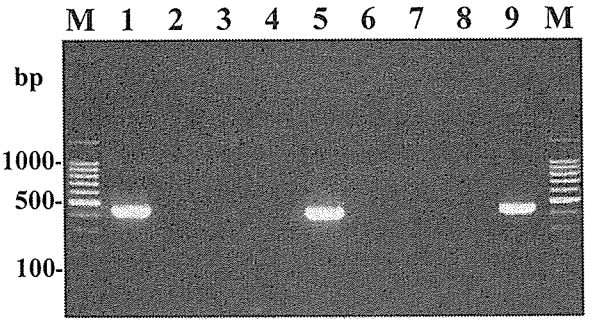

Fig. 3

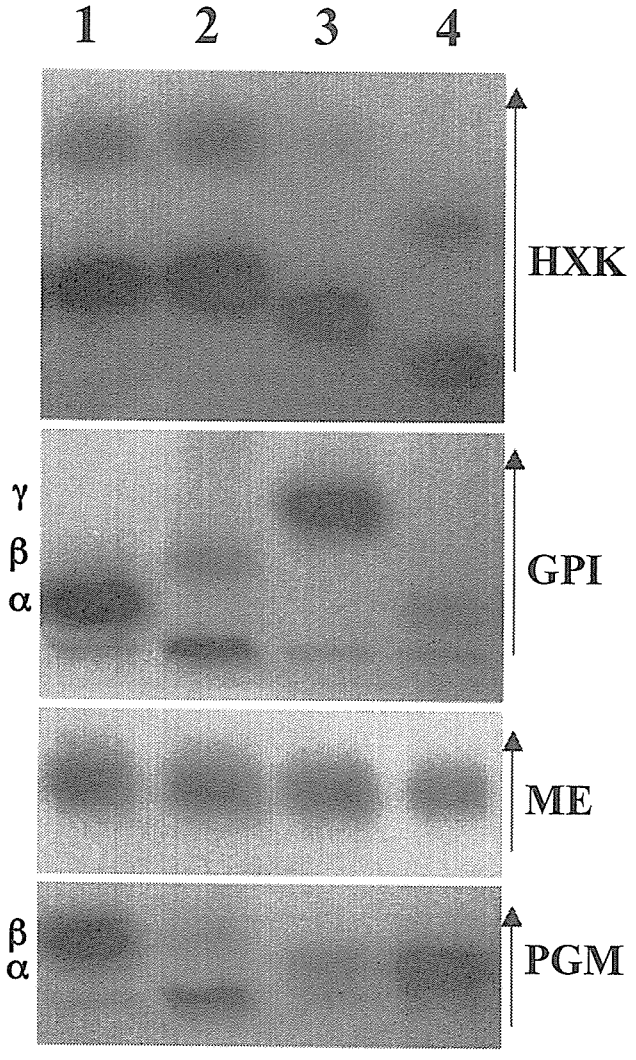

Fig. 4

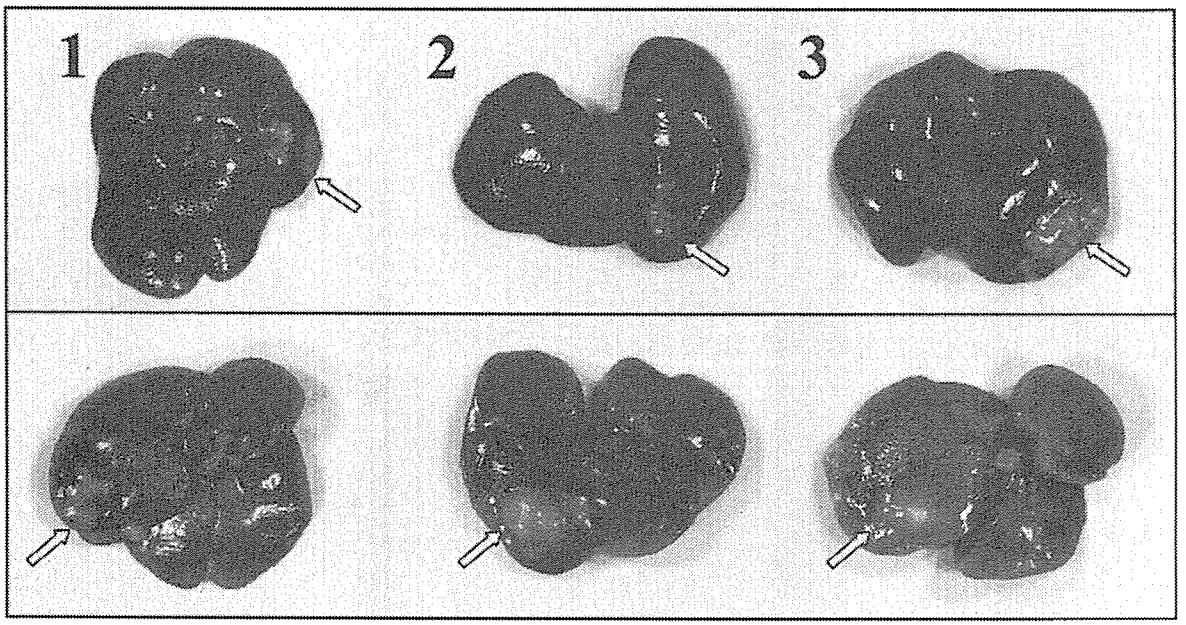

Fig. 5

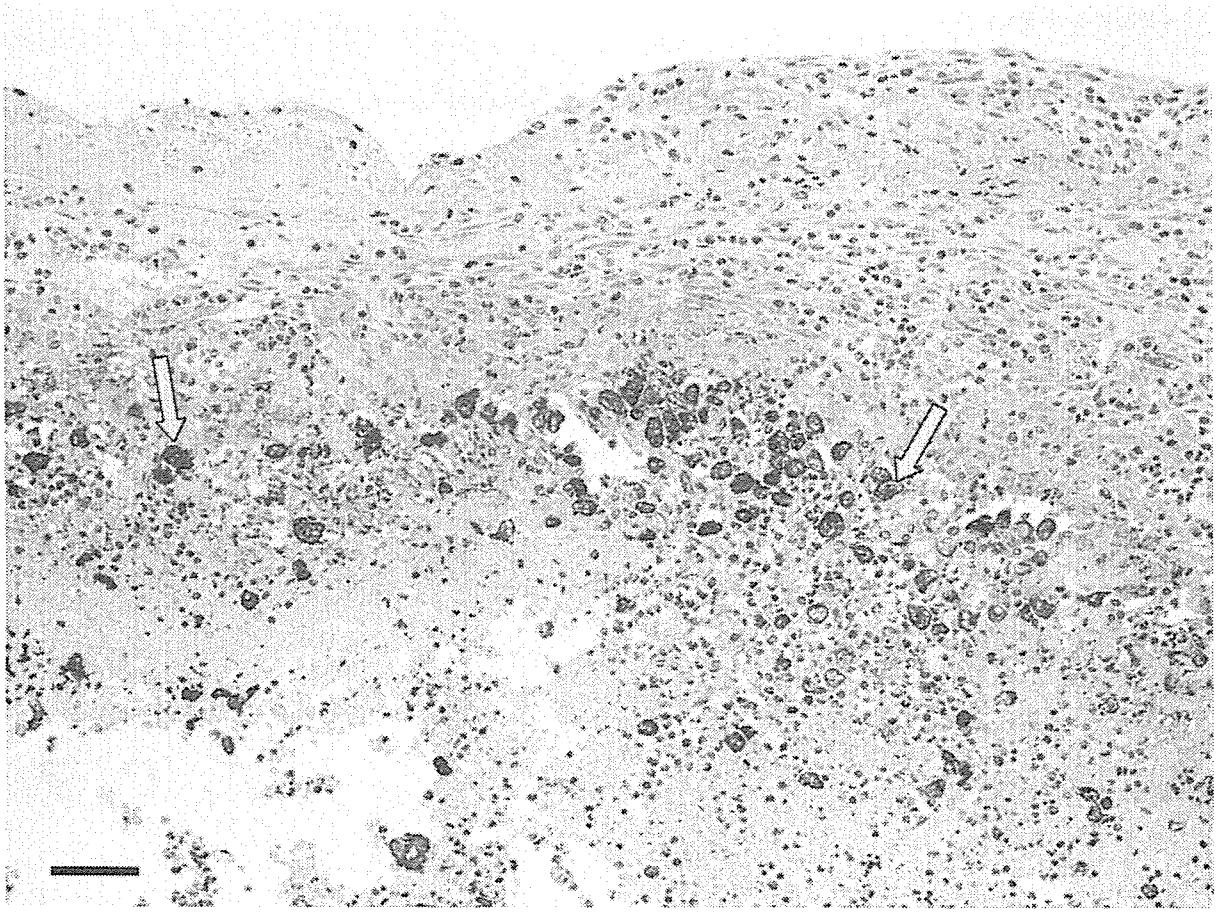

Fig. 6

Parasitology Research, in press

Original Paper

Comparison of *Entamoeba histolytica* DNA isolated from a cynomolgus monkey with human isolates

Jun-ichiro Takano · Toyoko Narita · Hiroshi Tachibana · Keiji Terao · Koji Fujimoto

Jun-ichiro Takano (Ä) · Toyoko Narita · Koji Fujimoto

The Corporation for Production and Research of Laboratory Primates

1 Hachimandai, Tsukuba-shi, Ibaraki 305-0843, Japan

E-mail: takano@primate.or.jp

Tel.: +81-29-8372121

Fax: +81-29-8372299

Hiroshi Tachibana

Department of Infectious Diseases, Tokai University School of Medicine

Isehara-shi, Kanagawa 259-1193, Japan

Keiji Terao

Tsukuba Primate Research Center, National Institute of Biomedical Innovation

1 Hachimandai, Tsukuba-shi, Ibaraki 305-0843, Japan

Abstract

Three protein-coding loci in DNA of an *Entamoeba histolytica* strain (EHMfas1) isolated from cynomolgus monkey (*Macaca fascicularis*) were sequenced; these loci corresponded to the genes for chitinase, the serine-rich *E. histolytica* protein (SREHP), and the 16S-like small subunit ribosomal RNA (16S-like SSUrRNA). The nucleotide and deduced amino-acid sequences of chitinase and SREHP were compared with sequences from human isolates. EHMfas1 had several specific mutations in units in the polymorphic regions of the chitinase and SREHP loci, with some repetition of these mutated units. The sequence of the 16S-like SSUrRNA gene was compared with other *Entamoeba* species. In phylogenetic analysis, EHMfas1 was not categorized in the *E. histolytica* cluster, but between *E. histolytica* and *E. dispar*. To our knowledge, this is the first molecular characterization of *E. histolytica* isolated from cynomolgus monkey, and our results indicate that EHMfas1 may be a subspecies of *E. histolytica* that infects cynomolgus monkey.

Introduction

Amoebiasis is caused by infection with *Entamoeba histolytica* and is one of the most important parasitic diseases in humans. In non-human primates, amoebiasis has been reported in captive and in wild-trapped non-human primates (Amyx et al. 1978; Beaver et al. 1988), but most of these reports did not distinguish *E. histolytica* from *E. dispar*, which is morphologically similar to *E. histolytica* but non-pathogenic. One case of *E. histolytica* infection in Japanese macaque (*Macaca fuscata*) was determined by isoenzyme analysis (Tachibana et al 1990), and PCR tests have been used to identity six cases of infections in three species of old world monkeys and three species of new world monkeys (Verweij et al. 2003) and in one case in cynomolgus monkey (*Macaca fascicularis*) (Takano et al. 2005). However, it is unclear whether the *E. histolytica* strains in non-human primates originated from humans or are species-specific strains.

The genes for chitinase (de la Vega et al. 1997; Ghosh et al. 2000; Haghighi et al. 2002; Haghighi et al. 2003) and serine-rich *E. histolytica* protein (SREHP) (also known as K2) (Li et al. 1992; Clark and Diamond 1993; Kohler and Tannich 1993; Stanley et al. 1990; Ayeh-Kumi et al. 2001; Haghighi et al. 2002; Haghighi et al. 2003) have highly polymorphic DNA loci, including the protein-coding sequences. These polymorphic loci have been used to study the molecular epidemiology and geographic diversity within human isolates of *E. histolytica*. Genes for the 16S-like small subunit ribosomal RNA (16S-like SSUrDNA) (Medlin et al. 1988) have been used for identification of species (Clark and Diamond 1991; Novati et al. 1996) and phylogenetic analysis of the genus *Entamoeba* has been performed using riboprinting (Clark and Diamond 1997) and nucleotide sequencing (Silberman et al. 1999; Clark et al. 2006).

In this study, we determined the DNA sequences of the chitinase and SREHP loci of the *Entamoeba histolytica* strain EHMfas1 and compared these sequences with those

from human isolates. We also examined the 16S-like SSUrDNA gene sequence to compare EHMfas1 with other *Entamoeba* species.

Materials and Methods

E. histolytica from cynomolgus monkey

The *E. histolytica* strain EHMfas1 was isolated from a healthy cynomolgus monkey imported into Japan from China (Takano et al. 2005).

DNA preparation and sequencing

Total genomic DNA from trophozoites was extracted using a QIAamp DNA Stool Mini Kit (Qiagen) according to the manufacturer's directions. Chitinase and SREHP were amplified by PCR using previously described primer sets: 5′-GGA ACA CCA GGT AAA TGT ATA-3′ and 5′-TCT GTA TTG TGC CCA ATT-3′ for the chitinase forward and reverse primers, respectively; and 5′-GCT AGT CCT GAA AAG CTT GAA GAA GCT G-3′ and 5′-GGA CTT GAT GCA GCA TCA AGG T-3′ for the SREHP forward and reverse primers, respectively (Ghosh et al. 2000). PCR was carried out in a 50 μl reaction mixture containing 2 μl of extracted DNA and 0.1 μg of bovine serum albumin per μl using Takara Ex *Taq* (Takara). A total of 35 cycles of PCR were performed as follows: denaturation at 94°C for 30 s, annealing at 54°C for 30 s, and extension at 72°C for 45 s. An initial denaturation step at 94°C for 5 min and a final extension step at 72°C for 7 min were also performed.

16S-like SSUrRNA was also amplified using previously described primer sets (RD5 and RD3; Clark and Diamond, 1997), with some modification, and newly designed internal primer sets to obtain accurate nucleotide sequencing: RD5, 5′-ATC TGG TTG ATC CTG CCA GT-3′ and 1r, 5′-TCG TCA CTA CCT CTT CAA TTC

G-3′; 2f, 5′-CAC TTC TAA GGA AGG CAG CAG-3′ and 2r, 5′-AGC GAT CAT GGA TTT TCA CC-3′; 3f, 5′-AAA AGG AAC AAT TGG GGT GA-3′ and 3r, 5′-CCA TGC ACC ACT ACC CAA TA-3′; 4f, 5′-TGG TCA CAA GGC TGA AAC TT-3′ and 4r, 5′-CCA AGA TGT CTA AGG GCA TCA-3′; 5f, 5′-GGG AAA AAG AAA AAG GAA GCA-3′ and RD3, 5′-ATC CTT CCG CAG GTT CAC CT-3′). A total of 30 cycles of PCR were performed as follows: denaturation at 94°C for 1 min, annealing at 58°C for 30 s, and extension at 72°C for 1 min. An initial denaturation step at 94°C for 5 min and a final extension step at 72 °C for 7 min were also performed.

Each PCR product of EHMfas1 was directly sequenced with an ABI PRISM BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) on an ABI PRISM 3100-Avant Genetic Analyzer (Applied Biosystems), according to the manufacturer's directions. Each DNA sequence was analyzed several times using independent DNA preparations.

DNA sequences of chitinase and SREHP were aligned using GENETYX ver. 7.0.11 (Genetyx). The GenBank accession numbers of the sequences used for comparison with the chitinase and SREHP genes are shown in Figures 1 and 2.

Phylogenetic Analysis

Analysis and multiple alignment of DNA sequences of 16S-like SSUrDNA genes were performed with ClustalX (Thompson et al. 1997), and the phylogenetic tree was constructed using the neighbor-joining (NJ) method (Saitou et al. 1987). The GenBank accession numbers and length of the sequences used in the phylogenetic analysis were as follows: X56991 (*E. histolytica*: 1911 bp), X64142 (*E. histolytica* strain HM-1:IMSS: 1910 bp), X61116 (*E. histolytica* strain SD4: 1907 bp) , Z49256 (*E.*

dispar: 1913 bp), AF149906 (*E. moshkovskii*: 1911 bp), AF149910 (*E. terrapinae*: 1910 bp), AF149909 (*E. insolita*: 1909 bp), AF149907 (*E. hartmanni*: 1927 bp), D28490 (*E. gingivalis*: 1910 bp), AF149905 (*E. invadens*: 1932 bp), AF149908 (*E. ranarum*: 1901 bp), AF149912 (*E. chattoni*: 1830 bp), AF149913 (*E. polecki*: 1825 bp), AF149914 (*E. coli*: 2068 bp), DQ286371 (*E. equi*: 1886 bp), DQ286372 (*E. suis*: 1970), DQ286373 (*E. ecuadoriensis*: 1908 bp) and AJ566411 (*E. struthionis*: 1830 bp). Comparisons were made over 1911 bp of EHMfas1. However, since the submitted *E. suis* sequence lacked some homologous base pairs at the 5′ and 3′ termini, the GenBank sequences were aligned after deleting some terminal base pairs to allow a direct alignment.

Nucleotide sequence accession numbers

The nucleotide sequence data reported here have been submitted to the GenBank/EMBL/DDBJ database under accession numbers AB282755 (chitinase), AB197935 (SREHP) and AB197936 (16S-like SSUrDNA).

Results

Polymorphism in the chitinase and SREHP loci of EHMfas1

The chitinase and SREHP genes of EHMfas1 were sequenced and compared with human isolates, which were genotyped and classified by the combination pattern of repeating units in the polymorphic region (Haghighi et al. 2002, 2003) (Figures 1 and 2).

The 5′-conserved region, the repeat-containing polymorphic region, and the 3′-conserved region of the EHMfas1 chitinase gene were compared with those of human isolates. A number was tentatively assigned to each unit in the nucleotide and deduced amino-acid sequences (Figure 1). In human isolates, the polymorphic region

is constructed from combinations of one type of non-repeating unit, three types of repeating units, and two conserved regions (Haghighi et al. 2002). The nucleotide sequence of EHMfas1 consists of two types of non-repeating units (CN1 and CN7), three types of repeating units (CN2, CN3 and CN5), and the 5′- and 3′-conserved regions (CN5′C and CN3′C1) (Figure 1A). The CN1 and CN2 units were common to EHMfas1 and the human isolates, but CN3, CN5, CN7 and CN3′C2 were EHMfas1-specific mutated units; these units corresponded to the CN2, CN4, CN6 and CN3′C1 units, respectively, in human isolates, with a single-nucleotide substitution in each unit (Figure 1A). In the deduced amino-acid sequence, only the CA3′C2 had a single amino-acid substitution compared to the corresponding human units (Figure 1B).

The 5′-conserved region, the repeat-containing polymorphic region, and the 3′-conserved region of the EHMfas1 SREHP gene were also compared with those of human isolates. As for chitinase, a number was tentatively assigned to each unit in the nucleotide and deduced amino-acid sequences (Figure 2). In human isolates, the polymorphic region of the SREHP gene contains combinations of four types of non-repeating units, five types of repeating units, and two conserved regions; some of the human isolates also have strain-specific mutated units (Haghighi et al. 2002, 2003). The nucleotide sequence of EHMfas1 consists of six types of non-repeating units (SN1, SN2, SN5, SN16, SN17 and SN20); four types of repeating units (SN3, SN9, SN10 and SN11), and the 5′- and 3′-conserved regions (SN5′C and SN3′C) (Figure 2A). The SN1, SN3, SN10, SN11 and SN16 units were common to EHMfas1 and the human isolates, whereas SN2, SN5, SN9, SN17, SN20 and SN3′C2 were EHMfas1-specific mutated units; these units corresponded to the SN1, SN4, SN8, SN16, SN19 and SN3'C1 units, respectively, in human isolates, with a single- or double-nucleotide substitution in each unit (Figure 2A). Furthermore, three EHMfas1-specific mutations

were observed as block insertions in the repeat-containing region: in the third unit in polymorphic region GAGGAA (EE), the fifth unit in GATGAAGAA (DEE), and the seventh unit in GAGAAT (EN) (Figures 2A and 2B). In the deduced amino-acid sequence, two types of non-repeating units (SA2 and SA5) and SA3′C2 had a single amino-acid substitution compared to the corresponding human units (Figure 2B). EHMfas1 lacked the SA6 unit found in human isolates.

Phylogenetic relationship of EHMfas1 among *Entamoeba* species

The 16S-like SSUrDNA of EHMfas1 was sequenced directly and the phylogenetic relationship among *Entamoeba* species was reconstructed with inclusion of EHMfas1 (Figure 3). In this phylogenetic tree, EHMfas1 was not categorized in the *E. histolytica* cluster, but between the *E. histolytica* cluster and the *E. dispar* branch.

Discussion

Three studies of *E. histolytica* infection in non-human primates have been performed (Tachibana et al. 1990; Verweij et al. 2003; Takano et al. 2005), using isoenzyme analysis, monoclonal antibody tests, and PCR. However, it is not clear if these infections were natural to the non-human primates or zoonotic infections from humans; i.e., whether *E. histolytica* is host-specific or a zoonotic agent. In this study, we sequenced three loci of EHMfas1 isolated from a cynomolgus monkey imported from China into Japan for use in medical research (Takano et al. 2005), and compared these data with human isolates.

The sequence of the chitinase gene was compared with those from human isolates to determine whether EHMfas1 could be categorized into any of the seven genotypes known to be present in humans (Haghighi et al. 2002, 2003). Chitinase genes from

human isolates have been genotyped based on the combination pattern of units in the polymorphic region, but our results showed that the nucleotide sequence of the chitinase gene of EHMfas1 and the deduced amino-acid sequence did not fit with any of the known genotypes (Figures 1A and 1B). Furthermore, the chitinase gene of EHMfas1 comprised several new units containing nucleotide substitutions from the corresponding units in human isolates; these mutated units suggest that EHMfas1 may be a subspecies of *E. histolytica,* rather than a new genotype.

The sequence of the SREHP gene was also compared with those in human isolates to confirm the conclusions drawn from the chitinase sequence, since the SREHP gene is more diverse and shows more variation in the human isolates (Haghighi et al. 2002, 2003). Similarly to the chitinase gene, the SREHP gene of EHMfas1 could not be categorized into any known human genotype at the nucleotide and amino-acid levels (Figures 2A and 2B), and this gene also included some new units containing nucleotide substitutions, compared to human sequences, and also included three block insertions and deletion of a common human unit. Therefore, comparison of the SREHP gene in EHMfas1 with those in human isolates also suggests that EHMfas1 is a subspecies of *E. histolytica.*

Some strain-specific mutated units have been observed in the SREHP gene in human isolates (SN7 for type C and type N, SN12 for type B, SN13 for type D, SN14 for type II, SN15 for type G, and SN18 for type VI); these strain-specific units were either non-repeated (SN7, SN12, SN14, SN15) or tandemly repeated (SN13 and SN18). From an evolutionary perspective, it is thought that non-repeated mutative units were derived from a common unit quite recently due to nucleotide substitutions, and that tandemly repeated units were replicated after nucleotide substitution, as a result of geographical differentiation of strains. In contrast to the strain-specific units in human

isolates, SN9 in EHMfas1 was dispersely repeated six times. This result indicates that SN9 divided from SN8 or from a common ancestral unit before unit amplification; in other words, EHMfas1 might have divided from human isolates before the effect of geographical differentiation in human isolates. In the same way, the CN3 and CN5 units in the chitinase gene, which were EHMfas1-specific mutated units, were dispersely repeated two and three times, respectively. Furthermore, insertions of three EHMfas1-specific blocks and deletion of the common unit SA6 support the hypothesis that EHMfas1 is a cynomolgus monkey-specific strain of *E. histolytica*. The results also suggest that EHMfas1 is a subspecies of *E. histolytica*, although it is possible that it is a new species. At present, the chitinase and SREHP genes have been sequenced from only a small number of isolates and are highly polymorphic between the various isolates. However, we believe the differences between codon usage in these two genes are useful to distinguish EHMfas1, which was detected as *E. histolytica*, using species-specific PCR and antigen-capture ELISA (Takano et al. 2005), together with 16S-like SSUrDNA analysis.

The 16S-like SSUrDNA was sequenced to investigate the relationship between EHMfas1 and the *Entamoeba* species, because 16S-like SSUrDNA genes have been used for previous phylogenetic analysis of *Entamoeba* (Clark and Diamond 1997; Silberman et al. 1999; Clark et al. 2006). Analysis of 16S-like SSUrDNA also suggested that EHMfas1 is a subspecies or a closely related new species of *E. histolytica*, since in the phylogenetic tree EHMfas1 was categorized as intermediate between the *E. histolytica* cluster and the *E. dispar* branch (Figure 3). In addition, isoenzyme analysis of EHMfas1 was also performed to compare EHMfas1 with human isolates. The zymodeme pattern of EHMfas1 (i.e., PGM: *E. dispar* type, ME: *E. histolytica*/*E. dispar* type, GPI: gamma band only, and HK: new type in which the faster

running band indicated the *E. histolytica* type, whereas the slower running band was intermediate between the *E. histolytica* and *E. dispar* types) was not classified into any known patterns. Isoenzyme analysis also supports our hypothesis that EHMfas1 is a subspecies or a closely-related new species of *E. histolytica* and *E. dispar*.

In the current paper, we have described one amoeba strain isolated from a cynomolgus monkey. *E. histolytica* infection is unusual in non-human primates, (Smith and Meerovitch 1985; Jackson et al. 1990; Rivera and Kanbara 1999; Tachibana et al. 2000; Tachibana et al. 2001; Takano et al. 2005) and only a small number of isolates have been studied. Therefore, it will be necessary to examine more *E. histolytica* strains isolated from non-human primates to determine the true relationships among non-human primates and human isolates.

Acknowledgement

We are grateful to Dr. Seiki Kobayashi for analyze of zymodeme pattern of EHMfas1. This work was supported by a Grant-in-Aid for Scientific Research from the Japanese Society for the Promotion of Science and a grant for Research on Emerging and Re-emerging Infectious Diseases from the Ministry of Health, Labour and Welfare of Japan.

References

Amyx HL, Asher DM, Nash TE, Gibbs CJ, Gajdusek DC (1978) Hepatic amebiasis in spider monkeys. Am J Trop Med Hyg 27: 888-891

Ayeh-Kumi PF, Ali IM, Lockhart LA, Gilchrist CA, Petri WA Jr, Haque R (2001) *Entamoeba histolytica*: genetic diversity of clinical isolates from Bangladesh as

demonstrated by polymorphisms in the serine-rich gene. Exp Parasitol 99: 80-88

Beaver PC, Blanchard JL, Seibold HR (1988) Invasive amebiasis in naturally infected New World and Old World monkeys with and without clinical disease. Am J Trop Med Hyg 39: 343-352

Clark CG, Diamond LS (1991) Ribosomal RNA genes of 'pathogenic' and 'nonpathogenic' *Entamoeba histolytica* are distinct. Mol Biochem Parasitol 49: 297-302

Clark CG, Diamond LS (1993) *Entamoeba histolytica*: a method for isolate identification. Exp Parasitol 77: 450-455

Clark CG, Diamond LS (1997) Inraspecific variation and phylogenetic relationships in the genus *Entamoeba* as revealed by riboprinting. J Euk Microbiol 44: 142-154

Clark CG, Kaffashian F, Tawari B, Windsor JJ, Twigg-Flesner A, Davies-Morel MC, Blessmann J, Ebert F, Peschel B, Le Van A, Jackson CJ, Macfarlane L, Tannich E (2006) New insights into the phylogeny of Entamoeba species provided by analysis of four new small-subunit rRNA genes. Int J Syst Evol Microbiol 56: 2235-2239

de la Vega H, Specht CA, Semio CE, Robbins PW, Eichinger D, Caplivski D, Ghosh S, Samuelson J (1997) Cloning and expression of chitinases of *Entamoebae*. Mol Biochem Parasitol 85: 139-147

Ghosh S, Frisardi M, Ramirez-Avila, Descoteaux S, Sturm-Ramirez K, Newton Sanchez OA, Sntos-Preciado JI, Ganguly C, Lohia A, Reed S, Samuelson J (2000) Molecular epidemiology of *Entamoeba* spp.: Evidence of a bottleneck (Demographic Sweep) and transcontinental spread of diploid parasites. J Clin Microbiol 38: 3815-3821

Haghighi A, Kobayashi S, Takeuchi T, Masuda G, Nozaki T (2002) Remarkable genetic polymorphism among *Entamoeba histolytica* isolates from a limited geographic area. J Clin Microbiol 40: 4081-4090

Haghighi A, Kobayashi S, Takeuchi T, Thammapalerd N, Nozaki T (2003) Geographic diversity among genotypes of *Entamoeba histolytica* field isolates. J Clin Microbiol 41: 3748-3756

Jackson TF, Sargeaunt PG, Visser PS, Gathiram V, Suparsad S, Anderson CB (1990) *Entamoeba histolytica*: naturally occurring infections in baboons. Arch Invest Med (Mex) 21 Suppl 1: 153-156

Kohler S, Tannich E (1993) A family of transcripts (K2) of *Entamoeba histolytica* contains polymorphic repetitive regions with highly conserved elements. Mol Biochem Parasitol 59: 49-58

Li E, Kunz-Jenkins C, Stanley SL Jr (1992) Isolation and characterization of genomic clones encoding a serine-rich *Entamoeba histolytica* protein. Mol Biochem Parasitol 50: 355-358

Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotic 16S-like SSUrRNA-coding regions. Gene 71: 491-499

Novati S, Sironi M, Granata S, Bruno A, Gatti S, Scaglia M, Bandi C (1996) Direct sequencing of the PCR amplified SSUrRNA gene of *Entamoeba dispar* and the design of primers for rapid differentiation from *Entamoeba histolytica*. Parasitology 112: 363-369

Rivera WL, Kanbara H (1999) Detection of *Entamoeba dispar* DNA in macaque feces by polymerase chain reaction. Parasitol Res 85: 493-495

Robinson GL (1968) The laboratory diagnosis of human parasitic amoebae. Trans R Soc Trop Med Hyg 62: 285-294

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406-425

Silberman JD, Clark CG, Diamond LS, Sogin ML (1999) Phylogeny of the genera *Entamoeba* and *Endolimax* as deduced from small-subunit ribosomal RNA sequence. Mol Biol Evol 16: 1740-1751

Smith JM, Meerovitch E (1985) Primates as a source of *Entamoeba histolytica*, their zymodeme status and zoonotic potential. J Parasitol 71: 751-756

Stanley SL Jr, Becker A, Kunz-Jenkins C, Foster L, Li E (1990) Cloning and expression of a membrane antigen of *Entamoeba histolytica* possessing multiple tandem repeats. Proc Natl Acad Sci USA 87: 4976-4980

Tachibana H, Kobayashi S, Kato Y, Nagakura K, Kaneda Y, Takeuchi T (1990) Identification of a pathogenic isolate-specific 30,000-Mr antigen of *Entamoeba histolytica* by using a monoclonal antibody. Infect Immun 58: 955-960

Tachibana H, Cheng XJ, Kobayashi S, Fujita Y, Udono T (2000) *Entamoeba dispar*, but not *E. histolytica*, detected in a colony of chimpanzees in Japan. Parasitol Res 86: 537-541

Tachibana H, Cheng XJ, Kobayashi S, Matsubayashi N, Gotoh S, Matsubayashi K (2001) High prevalence of infection with *Entamoeba dispar*, but not *E. histolytica*, in captive macaques. Parasitol Res 87: 14-17

Takano J, Narita T, Tachibana H, Shimizu T, Komatsubara H, Terao K, Fujimoto K (2005) *Entamoeba histolytica* and *Entamoeba dispar* infections in cynomolgus monkeys imported into Japan for research. Parasitol Res 97: 255-257

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25: 4876-4882