LONDON
SCHOOL *of*
HYGIENE
&TROPICAL
MEDICINE

1    <u>Title</u>

2    Limited intra-genetic diversity in *Dientamoeba fragilis* housekeeping genes.

3

4    <u>Authors</u>

5    Christen Rune Stensvold[a,*], C Graham Clark[b], Dennis Röser[a]

6

7    <u>Affiliation</u>

8    [a]Laboratory of Parasitology, Department of Microbiology and Infection Control, Statens Serum Institut,

9    Artillerivej 5, DK-2300 Copenhagen S, Denmark

10   [b]Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street,

11   WC1E 7HT London, UK.

12

13   <u>Corresponding author</u>

14   Christen Rune Stensvold, Laboratory of Parasitology, Department of Microbiology and Infection Control,

15   Statens Serum Institut, Artillerivej 5, DK-2300 Copenhagen S, Denmark. Tel: +45 3268 8224; email: <u>run@ssi.dk</u>

16

17

18

19

20

21

22

23

24    Abstract

25    *Dientamoeba fragilis* is a common parasite of unsettled clinical significance. Differences in clinical outcome of

26    intestinal parasitic infections may reflect parasite genetic diversity, and so tools to study intra-genetic diversity

27    that could potentially reflect differences in clinical phenotypes are warranted. Here, we show that genetic

28    analysis of three *Dientamoeba fragilis* housekeeping genes enables clear distinction between two genotypes,

29    but that integration of housekeeping genes in multi-locus sequencing tools for *D. fragilis* may have limited

30    epidemiological and clinical value.

31

32    Keywords

33    *Dientamoeba*, genetic diversity, actin, elongation factor, multilocus sequencing, molecular epidemiology

34

35    Highlights

36    Carriers of *Dientamoeba fragilis* may or may not experience symptoms → intragenetic diversity may be

37    associated with clinical outcome →SSU rDNA analysis enables the distinction of two genotypes → analysis of

38    two additional *D. fragilis* genes did not add further genetic resolution → analysis of *D. fragilis* housekeeping

39    genes may have limited epidemiological value.

40

41

42

43

44

45

46

47    <u>Manuscript</u>

48

49    **1. Introduction**

50

51    *Dientamoeba fragilis* is an intestinal parasite of unsettled clinical significance and possibly transmitted by

52    pinworm (Johnson et al., 2004; Röser et al., 2013a; Stensvold et al., 2007a). In our laboratory 43% of

53    approximately 22,000 faecal DNAs from patients with intestinal symptoms and analyzed by real-time PCR were

54    positive, with a range in positive proportion from 10—70% depending on age group (Röser et al., 2013b). The

55    parasite is common in individuals both with and without intestinal symptoms (Stensvold et al., 2009), and

56    similar to the situation for various other intestinal parasites, identification of tools to study intra-genetic

57    diversity that could potentially reflect differences linked to clinical outcome of infection and facilitate

58    epidemiological studies appears relevant.

59

60    RFLP analysis of SSU rDNA PCR products enables distinction between the two genotypes currently known

61    (genotypes 1 and 2); the sequences differ by at least 2% (Johnson and Clark, 2000; Peek et al., 2004; Stark et

62    al., 2005). Genotyping has also been performed by SSU rDNA SNP analysis using PCR and pyrosequencing

63    (Stensvold et al., 2007b). The value of sequencing the Internal Transcribed Spacer (ITS) region for typing studies

64    of *D. fragilis* is limited due to intra-strain genetic heterogeneity (Windsor et al., 2006). C-profiling was

65    developed as a means of extracting useful data from sequenced ITS clones (Bart et al., 2008), but the method

66    has only been employed in a minor case report (Stark et al., 2009), and so little is known on its applicability and

67    epidemiological relevance on a broader scale.

68

69    Studies of other housekeeping genes may prove useful in terms of obtaining higher resolution than can be

70    obtained by studies of SSU rRNA genes alone, as in the case of other metamonads such as *Giardia* and

71    *Trichomonas* (Cornelius et al., 2012; Feng and Xiao, 2011). Two *D. fragilis* genotype 1 housekeeping genes,

72    namely actin and elongation factor 1 alpha (EF-1α) were recently sequenced (Noda et al., 2012), and this study

73    aimed to characterize these two genes in *D. fragilis* genotype 2 and in *D. fragilis*-positive patient samples sent

74    for parasitological analysis in our clinical microbiology laboratory.

75

76    **2. Materials and methods**

77

78    A total of 40 faecal DNAs were chosen randomly among those testing positive for *D. fragilis* by a *D. fragilis*-

79    specific real-time PCR (Verweij et al., 2007) in our clinical microbiology laboratory. DNAs had been extracted

80    directly from fresh faecal specimens from patients with gastrointestinal complaints in the absence of viral or

81    bacterial pathogens, using the automated NucliSENS® easyMag® protocol (Andersen et al., 2013). Each DNA

82    was submitted to single round conventional PCRs targeting actin and EF-1α genes, but also SSU rRNA genes for

83    confirmation of the real-time PCR result and for genotyping. Primers for SSU rDNA amplification by

84    conventional PCR and sequencing were those used by Röser et al. (2013a) (Table 1), while primers for

85    amplification of actin and EF-1α genes were designed based on GenBank accession nos. AB468093 and

86    AB468119, respectively. In cases where virtually complete genes (>95%) could not be obtained, primers

87    targeting a minor fragment of the genes were used (Table 1).

88

89    Since actin and EF-1α gene sequences were available only for genotype 1 (Noda et al., 2012) and not had been

90    characterized for genotype 2, these genes were amplified from DNA from the Bi/PA strain (kindly provided by

91    Dr Graham Clark) and sequenced bidirectionally; sequences were submitted to GenBank (Accession nos.

92   KC967121-KC967122). As a control measure, the SSU rRNA gene was amplified from the Bi/PA strains as well,

93   and the 364 bp SSU rDNA sequence obtained in the present study showed 100% identity to the Bi/PA strain

94   sequence present in GenBank (acc. no. U37461).

95

96   Virtually complete actin and EF-1α sequences (>95% gene coverage) representing the Bi/PA strain were

97   translated, concatenated, aligned with translated and concatenated reference sequences (Noda et al., 2012)

98   including *D. fragilis* genotype 1 (DfA3 and DfE3C clones), and submitted to phylogenetic analysis, including

99   distance-based (Neighbor-Joining (NJ)) and Maximum Likelihood (ML) analysis, using Molecular Evolutionary

100  Genetics Analysis version 5 (MEGA 5) (Tamura et al., 2011); ModelTest  (Posada and Crandall, 1998) was

101  performed and the WAG + Γ model selected. Statistical support for distance-based and ML trees was evaluated

102  using bootstrapping (1,000 replicates). Phylogenetic analysis of each individual translated gene (actin and EF-

103  1α) was also performed; for ML analysis, the WAG + Γ model was selected for analysis of actin proteins, while

104  the rtRev + Γ model was chosen for EF-1α proteins. Since these models are not available for NJ analysis, NJ

105  analysis used JTT + Γ, and the gamma value (given in the ModelTest output) was 0.41 and 0.5 for the actin and

106  EF-1α, respectively.

107

108  All data were anonymised prior to analysis, and so no personally identifiable data were included in the study.

109

110  **3. Results and Discussion**

111

112  Using the faecal DNA templates from patient samples, the SSU rRNA, actin, and EF-1α genes could be amplified

113  and unambiguously sequenced in 32/40, 29/40 and 21/40 cases, respectively. As seen, EF-1α genes could be

114  successfully amplified and sequenced in only 53% of the cases, which could be explained by the fact that Ct-

115 values obtained by real-time PCR (SSU rRNA gene) were significantly lower for DNAs from which EF-1α genes

116 could be amplified and unambiguously sequenced than for the DNAs where either no amplification was

117 obtained or where (often faint) PCR products gave rise to unclear sequence traces (p<0.001; Student's T-test

118 for comparison of means (data not shown)).

119

120 Sequences were aligned and interpreted manually. One patient sample (1/32, 3%) (T14157) was found to

121 belong to genotype 2, while the remainder of the samples (31/32, 97%) for which SSU rDNAs were available

122 belonged to genotype 1; these data are in line with previous reports on the relative prevalence of the two

123 genotypes (Johnson and Clark, 2000; Peek et al., 2004; Windsor et al., 2006). T14157 and Bi/PA were 100%

124 identical across all three genes (data not shown). T14157 was from a 62 year old male with persistent intestinal

125 symptoms, who had submitted multiple faecal samples for traditional clinical microbiology analyses with no

126 evidence of enteric viruses, enteropathogenic bacteria or other intestinal parasites except for *Blastocystis*;  this

127 patient was the oldest patient in the study group (n=40; median age: 16.5 years [IQR 6.0—42.0]).

128

129 The two genotypes differed by 29 unambiguous SNPs scattered across the actin gene, (Supplementary Fig. 1),

130 of which 4 were non-synonymous substitutions. Likewise, across the EF-1α gene (Supplementary Fig. 2), 25

131 scattered unambiguous SNPs were identified, of which 4 were non-synonymous substitutions. In comparison,

132 SSU rRNA genes from the two genotypes differ by at least 2% and hence, the amount of genetic variation seen

133 across the actin and EF-1α genes, which are both in the size range of 800-850 bp, is comparable to the amount

134 of variation seen in the SSU rRNA gene, if only a little higher (about 3%).

135

136    No strain-unique SNPs were detected across any of the two genes among the genotype 1 samples. However,

137    there were several positions in each sequence exhibiting consistent allelic heterozygosity, although difficult to

138    discern in some of the trace files, and representing synonymous substitutions only.

139

140    Phylogenetic analysis of concatenated actin and EF-1α proteins using translated sequence data and reference

141    sequences from the alignment given by (Noda et al., 2012)) consolidated the existence of two genotypes

142    clustering with maximum bootstrap support, and sharing a most recent common ancestor with *Histomonas*

143    (Figure 1); individual trees produced for each translated gene consolidated these phylogenetic inferences

144    (Supplementary Fig. 3).

145

146    Although the study is limited by the fact that *D. fragilis* from healthy individuals was not included, the present

147    data suggest a high degree of conservation in *D. fragilis* housekeeping genes.

148

149    The data show that analysis of intra-genetic diversity in house-keeping genes may have limited epidemiological

150    and clinical usefulness in studies of *D. fragilis* in humans. However, pigs and gorillas have been identified as

151    natural hosts of *D. fragilis* (Cacciò et al., 2012; Lankester et al., 2010; Stark et al., 2008), and while SSU rDNA

152    data point towards the probability that pigs are natural hosts of genotype 1 (Cacciò et al., 2012), it remains to

153    be seen whether analysis of non-SSU rRNA genes in isolates from non-human hosts identify intra-genetic

154    variation, thereby enabling studies of transmission and further exploration of zoonotic potential.

155    As yet, *D. fragilis* genome sequences have not been published, but steadily decreasing costs related to genome

156    sequencing using high-throughput platforms and identification of ways to obtain genomic data from small

157    amounts of DNA should prompt the initiative of complete sequencing of mitochondrial or even nuclear

158    genomes in future efforts to screen isolates from symptomatic and asymptomatic carriers for genetic variation.

159    Finally, the prevalence and clinical significance of genotype 2 should be studied and compared to genotype 1.

160

161    <u>Acknowledgements</u>

162

163    Lis Lykke Wassmann is thanked for excellent laboratory assistance.

164

165    <u>Figure legends</u>

166

167    Figure 1. Phylogenetic analysis of translated and concatenated actin and EF-1α sequences representing the

168    Bi/PA and the DfA3 strains along with reference organisms from the publication by (Noda et al., 2012); ML tree

169    is shown with the support values in the order ML/NJ . Values less than 50% with both methods are either not

170    shown or marked by an asterisk. *Df = D. fragilis*.

171

172    Supplementary Figure 1: Alignment of actin gene sequences for genotype 1 (DfA3 clone; AB468093) and

173    genotype 2 (Bi/PA strain; KC967121).

174

175    Supplementary Figure 2: Alignment of EF-1α gene sequences for genotype 1 (DfE3C clone; AB468119) and

176    genotype 2 (Bi/PA strain; KC967122).

177

178    Supplementary Figure 3: Phylogenetic analysis of translated EF-1α and actin gene sequences.  The ML tree is

179    shown with the support values in the order ML/NJ. Values less than 50% with both methods are either not

180    shown or marked by an asterisk. *Df = D. fragilis.*

181

**References**

183

184    Andersen, L.O., Röser, D., Nejsum, P., Nielsen, H.V., Stensvold, C.R., 2013. Is supplementary bead beating for

185    DNA extraction from nematode eggs using the NucliSENS(R) easyMag(R) protocol necessary? J Clin Microbiol

186    51, 1345-1347.

187

188    Bart, A., van der Heijden, H.M., Greve, S., Speijer, D., Landman, W.J., van Gool, T., 2008. Intragenomic variation

189    in the internal transcribed spacer 1 region of *Dientamoeba fragilis* as a molecular epidemiological marker. J Clin

190    Microbiol 46, 3270-3275.

191

192    Cacciò, S.M., Sannella, A.R., Manuali, E., Tosini, F., Sensi, M., Crotti, D., Pozio, E., 2012. Pigs as natural hosts of

193    *Dientamoeba fragilis* genotypes found in humans. Emerg Infect Dis 18, 838-841.

194

195    Cornelius, D.C., Robinson, D.A., Muzny, C.A., Mena, L.A., Aanensen, D.M., Lushbaugh, W.B., Meade, J.C., 2012.

196    Genetic characterization of *Trichomonas vaginalis* isolates by use of multilocus sequence typing. J Clin

197    Microbiol 50, 3293-3300.

198

199    Feng, Y., Xiao, L., 2011. Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. Clin

200    Microbiol Rev 24, 110-140.

201    Johnson, E.H., Windsor, J.J., Clark, C.G., 2004. Emerging from obscurity: biological, clinical, and diagnostic

202    aspects of *Dientamoeba fragilis*. Clin Microbiol Rev 17, 553-570.

203

204    Johnson, J.A., Clark, C.G., 2000. Cryptic genetic diversity in *Dientamoeba fragilis.* J Clin Microbiol 38, 4653-4654.

205

206    Lankester, F., Kiyang, J.A., Bailey, W., Unwin, S., 2010. *Dientamoeba fragilis*: initial evidence of pathogenicity in

207    the western lowland gorilla (*Gorilla gorilla gorilla*). J Zoo Wildl Med 41, 350-352.

208

209    Noda, S., Mantini, C., Meloni, D., Inoue, J., Kitade, O., Viscogliosi, E., Ohkuma, M., 2012. Molecular phylogeny

210    and evolution of parabasalia with improved taxon sampling and new protein markers of actin and elongation

211    factor-1α. PLoS One 7, e29938.

212

213    Peek, R., Reedeker, F.R., van Gool, T., 2004. Direct amplification and genotyping of *Dientamoeba fragilis* from

214    human stool specimens. J Clin Microbiol 42, 631-635.

215

216    Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817-

217    818.

218

219    Röser, D., Nejsum, P., Carlsgart, A.J., Nielsen, H.V., Stensvold, C.R., 2013a. DNA of *Dientamoeba fragilis*

220    detected within surface-sterilized eggs of *Enterobius vermicularis*. Exp Parasitol 133, 57-61.

221

222    Röser, D., Simonsen, J., Nielsen, H.V., Stensvold, C.R., Mølbak, K., 2013b. *Dientamoeba fragilis* in Denmark:

223    epidemiological experience derived from four years of routine real-time PCR. Eur J Clin Microbiol Infect Dis. DOI

224    10.1007/s10096-013-1880-2.

225

226    Stark, D., Beebe, N., Marriott, D., Ellis, J., Harkness, J., 2005. Detection of *Dientamoeba fragilis* in fresh stool

227    specimens using PCR. Int J Parasitol 35, 57-62.

228

229    Stark, D., Phillips, O., Peckett, D., Munro, U., Marriott, D., Harkness, J., Ellis, J., 2008. Gorillas are a host for

230    *Dientamoeba fragilis*: an update on the life cycle and host distribution. Vet Parasitol 151, 21-26.

231

232    Stark, D., Barratt, J., Ellis, J., Harkness, J., Marriott, D., 2009. Repeated *Dientamoeba fragilis* infections: a case

233    report of two families from Sydney, Australia. Infect Dis Rep; 1:e4.

234

235    Stensvold, C.R., Arendrup, M.C., Mølbak, K., Nielsen, H.V., 2007a. The prevalence of *Dientamoeba fragilis* in

236    patients with suspected enteroparasitic disease in a metropolitan area in Denmark. Clin Microbiol Infect 13,

237    839-842.

238

239    Stensvold, C.R., Jespersgaard, C., Dinesen, K., Sorensen, J.F., Molbak, K., Nielsen, H.V., 2007b. Prevalence of

240    *Dientamoeba fragilis* genotypes in different Danish populations as assessed by PCR and SNP genotyping.

241    Tropical Medicine & International Health 12, 182-182.

242

243    Stensvold, C.R., Lewis, H.C., Hammerum, A.M., Porsbo, L.J., Nielsen, S.S., Olsen, K.E., Arendrup, M.C., Nielsen,

244    H.V., Mølbak, K., 2009. *Blastocystis*: unravelling potential risk factors and clinical significance of a common but

245    neglected parasite. Epidemiol Infect 137, 1655-1663.

246

247    Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary

248    genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol

249    Biol Evol 28, 2731-2739.

250

251    Verweij, J.J., Mulder, B., Poell, B., van Middelkoop, D., Brienen, E.A., van Lieshout, L., 2007. Real-time PCR for

252    the detection of *Dientamoeba fragilis* in fecal samples. Mol Cell Probes 21, 400-404.

253

254    Windsor, J.J., Macfarlane, L., Clark, C.G., 2006. Internal transcribed spacer dimorphism and diversity in

255    *Dientamoeba fragilis*. J Eukaryot Microbiol 53, 188-192.

256

257

258

259

260

261

262

263

264

265

266    Table 1. Primers used in the study (see text for details).

| Gene | Primers | Reference |
|------|---------|-----------|
| SSU rRNA (18S) | DFpn_1f<br>5'-GCC AAG GAA GCA CAC TAT GG-3'<br><br>DFpn_364r | (Röser et al., 2013a) |

| | 5'-GTA AGT TTC GCG CCT GCT-3' | |
|---|---|---|
| Actin | DF_ACTIN_3f<br>5'-CCA CAC ATT CTA CAA CGA ATT AC-3'<br><br>DF_ACTIN_157F<br>5'-GTT CTT TCA CTT TAC TCA TCA GGT C-3'<br><br>DF_ACTIN_291R<br>5'-GAC CAG CAA GGT TGA GTC TC-3'<br><br>DF_ACTIN_843r<br>5'-TGG ACC AGC TTC ATT GTA TTC-3' | Present study |
| EF-1α | DF_EF_1f<br>5'-CTC ACT TTG GAA GTT CGA ATC-3'<br><br>DF_EF_265F<br>5'-TCA AAG GCT CGT TAT GAT GAA ATC-3'<br><br>DF_EF_364R<br>5'-GAA ACC TGA GAT TGG AAC AAA C-3'<br><br>DF_EF_ 836r<br>5'-CTG TGT GGC AAT CGA AAA C-3' | Present study |

267
268
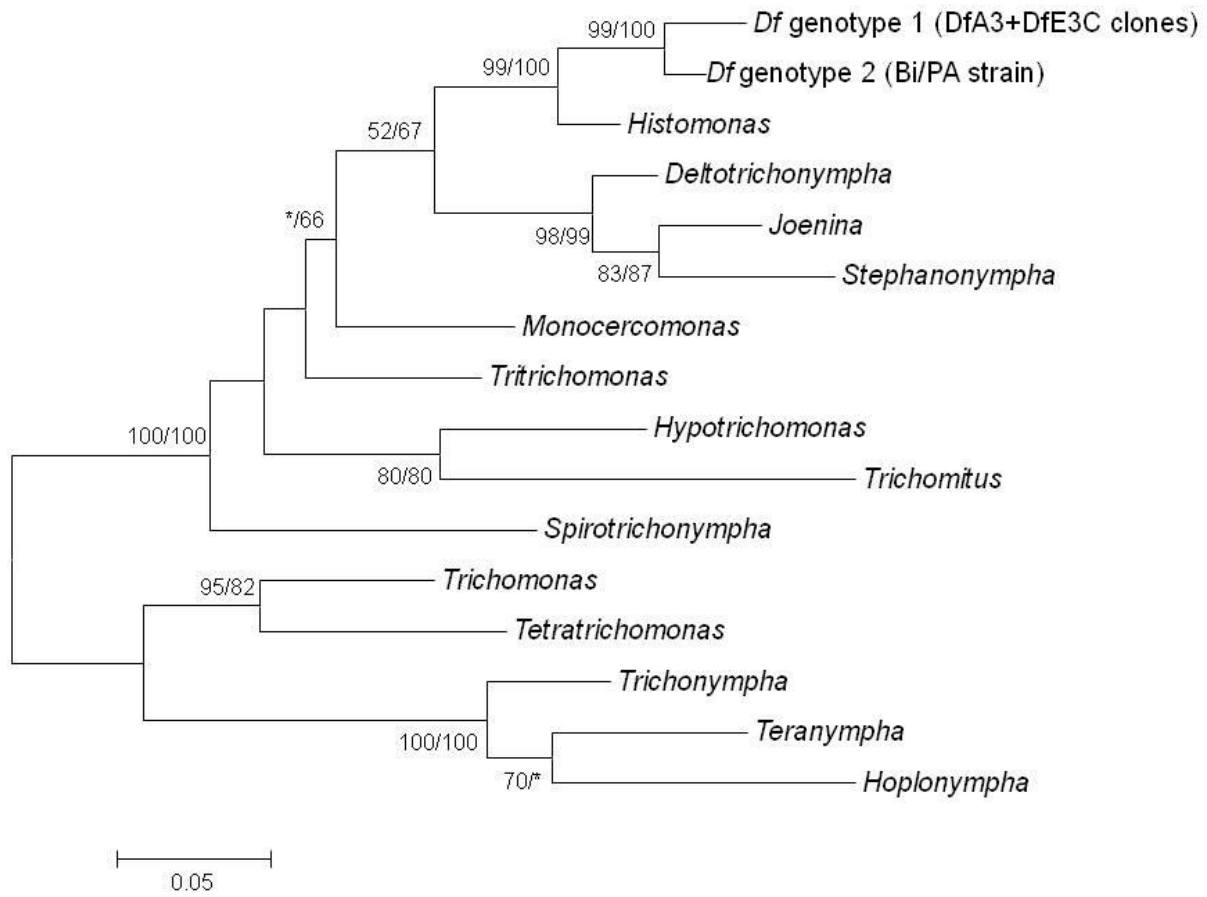
269
270      Fig 1

271



272
273    Suppl Fig 1

274

275

Suppl Fig 2 alignment:

```
              1    10    20    30    40    50    60    70    80    90    100   110   120   130
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AB468119(DfE3Cclone)  CTCACTTTGGAAGTTCGAATCACCAAAGTACTTCTTCACAATCATTGATGCTCCAGGACACAGAGACTTCATCAAGAACATGATCACAGGTACATCACAAGCTGATGCTGCTGTTCTCGTTATCGATTCA
KC967122(Bi/PAstrain)             CGAATCACCAAAGTACTTCTTCACAATCATTGATGCTCCAGGACACAGAGATTTCATCAAGAACATGATCACAGGTACATCACAAGCTGATGCTGCAGTTCTCGTTATCGATTCA
Consensus             ..............CGAATCACCAAAGTACTTCTTCACAATCATTGATGCTCCAGGACACAGAGAcTTCATCAAGAACATGATCACAGGTACATCACAAGCTGATGCTGCaGTTCTCGTTATCGATTCA

              131  140   150   160   170   180   190   200   210   220   230   240   250   260
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AB468119(DfE3Cclone)  ACACGTGGTGGTTTCAAAGCCGGTATTGCCGAACAAGGTCAAACACGCGAACACGCTCTTCTTGCCTTCACACTTGGTATCAAGCAACTCATCGTTGGTGTTAACAAGATGGATGATAACACAGTTAACT
KC967122(Bi/PAstrain) ACACGTGGTGGTTTCGAAGCCGGTATTGCCGAACAAGGTCAAACACGTGAACACGCTCTTCTTGCTTTCACACTCGGTATCAAGCAACTCATCGTTGGTGTTAACAAGATGGATGATAACACAGTTAACT
Consensus             ACACGTGGTGGTTTCaAAGCCGGTATTGCCGAACAAGGTCAAACACGcGAACACGCTCTTCTTGCcTTCACACTcGGTATCAAGCAACTCATCGTTGGTGTTAACAAGATGGATGATAACACAGTTAACT

              261  270   280   290   300   310   320   330   340   350   360   370   380   390
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AB468119(DfE3Cclone)  ACTCAAAGGCTCGTTATGATGAAATCGTTGGTGAAATGACACGTATCCTTACAAACATTGGTTTCAAGCCAGAACAATACAAGTTTGTTCCAATCTCAGGTTTCGCTGGTGATAACATGACAGAAAAGTC
KC967122(Bi/PAstrain) ACTCAAAGGCTCGTTTCGATGAAATCGTTGGTGAAATGACACGTATCCTTACAAACATTGGTTTCAAGCCAGAACAATACAAGTTYGTTCCAATMTCAGGTTTYGCTGGTGATAACATGACAGAAAAGTC
Consensus             ACTCAAAGGCTCGTTacGATGAAATCGTTGGTGAAATGACACGTATCCTTACAAACATTGGTTTCAAGCCAGAACAATACAAGTTtGTTCCAATcTCAGGTTTcGCTGGTGATAACATGACAGAAAAGTC

              391  400   410   420   430   440   450   460   470   480   490   500   510   520
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AB468119(DfE3Cclone)  ACCAAACACGCCATGGTACACAGGTGGTACACTCCTCGAAACACTCGATACACTCAACCCACCACAACGTCCATTCGACCGTCCACTCCGTCTTCCAGTTCAAGATGTTTACAAGATCTCAGGTATTGGT
KC967122(Bi/PAstrain) ACCAAACATGCCATGGTACACAGGTGGTACACTCCTCGAAACACTCGATACACTTAACCCACCACAACGTCCATATGACCGTCCACTCCGTCTTCCAGTTCAAGATGTTTACAAGATCTCAGGTATTGGT
Consensus             ACCAAACAcGCCATGGTACACAGGTGGTACACTCCTCGAAACACTCGATACACTcAACCCACCACAACGTCCATacGACCGTCCACTCCGTCTTCCAGTTCAAGATGTTTACAAGATCTCAGGTATTGGT

              521  530   540   550   560   570   580   590   600   610   620   630   640   650
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AB468119(DfE3Cclone)  ACAGTTCCAGTTGGTCGTGTTGAATCAGGTATCATGAAGCCAAACATGACAGTTATCTTCGCTCCATCATCAGTTACAACTGAAGTTAAGTCAATCGAAATGCACCACACACAACTTCCAGAAGCTGTTC
KC967122(Bi/PAstrain) ACAGTTCCAGTTGGTCGTGTTGAATCAGGTATCATGAAGCCAAACATGACAGTTGTTTTCGCTCCATCAACAGTTACAACTGAAGTTAAGTCAATCGAAATGCACCACACACAACTCCCAGAAGCCGTTC
Consensus             ACAGTTCCAGTTGGTCGTGTTGAATCAGGTATCATGAAGCCAAACATGACAGTTaTcTTCGCTCCATCAaCAGTTACAACTGAAGTTAAGTCAATCGAAATGCACCACACACAACTcCCAGAAGCcGTTC

              651  660   670   680   690   700   710   720   730   740   750   760   770   780
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AB468119(DfE3Cclone)  CAGGTGATAACATTGGTTTCAACGTTAAGTCAGTTGCCGTTTCTGATATCAAGCGTGGTTATGTTGTTGGTGAAGCTAACCGTGATCCACCAGTTCAATGCAACAGCTTCACAGCTCAAATGATCATTTC
KC967122(Bi/PAstrain) CAGGTGATAACATTGGTTTCAACGTTAAGTCAGTTGCTGTTTCTGATATCAAGCGTGGTTATGTTGTTGGTGAAGCAAACCGTGATCCACCAGTTCAATGCCTTAGCTTTACAGCTCAAATGATCATTTC
Consensus             CAGGTGATAACATTGGTTTCAACGTTAAGTCAGTTGCcGTTTCTGATATCAAGCGTGGTTATGTTGTTGGTGAAGCaAACCGTGATCCACCAGTTCAATGCaacAGCTTcACAGCTCAAATGATCATTTC

              781  790   800   810   820   830   838
              |--------+---------+---------+---------+---------+---------+------|
AB468119(DfE3Cclone)  CAACCACCCAGGTAAGATTCACGCCGGTTACCAACCAGTTTTCGATTGCCACACAGCT
KC967122(Bi/PAstrain) AAACCACCCAGGTAAGATCCACGCCGGTTACCAACCAGTT
Consensus             aAACCACCCAGGTAAGATcCACGCCGGTTACCAACCAGTT..................
```
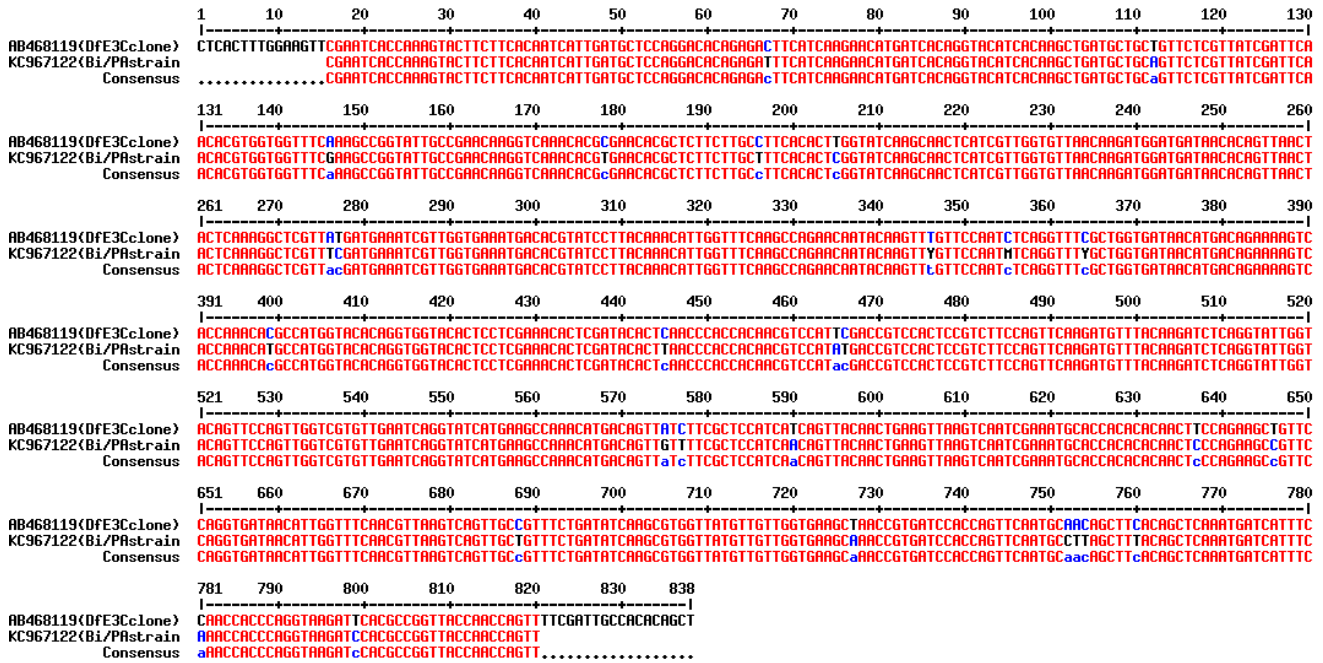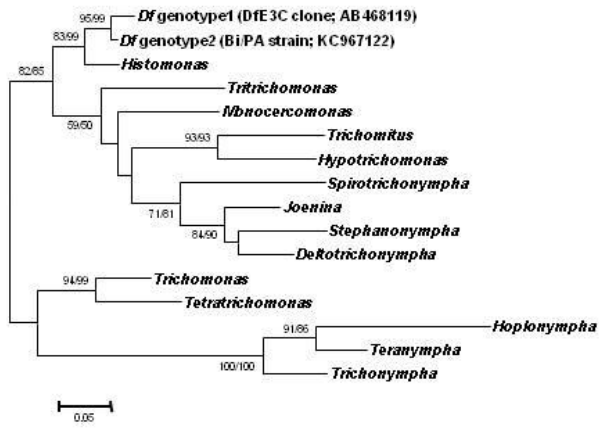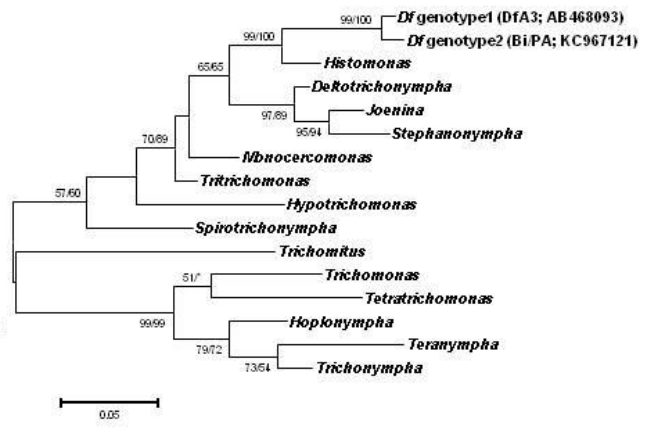
276
277  Suppl Fig 2
278

Ef1α

Actin

279
280    Suppl Fig 3