Comparative analysis of the EF-1α Intergenic Region in *Babesia divergens* isolates: Insights into TA Repeat Variation and Potential Regulatory Implications

Sezayi Ozubek[1], Alejandro Sanchez-Flores[2*], Estrella Montero[3*], Heba Alzan[4,5], Carlos E. Suarez[5], Ricardo Grande[2], Aitor Gil-Iglesias[3], Reginaldo G. Bastos[5,6], Munir Aktas[1], Luis Miguel Gonzalez[3]

[1] Department of Parasitology, Faculty of Veterinary Medicine, University of Firat, Elazig, Turkiye

[2] Unidad Universitaria de Secuenciación Masiva y Bioinformatica, Institute of Biotechnology, UNAM, Cuernavaca, México.

[3] Laboratorio de Referencia e Investigacion en Parasitologıa, Centro Nacional de Microbiología, ISCIII Majadahonda, Madrid, Spain

[4] Parasitology and Animal Diseases Department, National Research Center, Dokki, Giza, Egypt

[5] Department of Veterinary Microbiology and Pathology, College of Veterinary Medicine, Washington State University, Pullman, WA, USA

[6] Animal Disease Research Unit, USDA Agricultural Research Service, Pullman, Washington, USA

*Correspondence: alejandro.sanchez@ibt.unam.mx; estrella.montero@isciii.es.

Short title:

The EF-1α locus variation in *Babesia divergens* and other species

Abstract:

Babesia divergens, a zoonotic tick-borne pathogen, causes bovine and human babesiosis in Europe. The Elongation Factor 1 alpha (EF-1α) protein is important in many cellular processes and has emerged as a possible target for subunit vaccine development against parasitic infections, and its intergenic region (IG) is an important tool for genetic manipulation of *Babesia* parasites. While the EF-1α locus of *B. divergens* has been described, structural variation between isolates was poorly defined. In order to fill this gap, we performed a comparative analysis of the EF-1α-IG region in *B. divergens* human (Rouen 87, and Spanish sample), and bovine (Turkiye) host isolates. Our findings revealed both conserved and variable elements, particularly in TA nucleotide

repeat numbers and IG sequence length. The Spanish isolate exhibited the highest TA repeat expansion, whereas the Rouen 87 strain had the shortest IG region. Given the known role of repeat-rich promoter elements in gene regulation, these differences may influence EF-1α transcription. Additionally, these findings provide insights into the evolutionary divergence of *B. divergens* and its host adaptation mechanisms. This study establishes a foundation for future gene editing and transfection strategies, where selecting intergenic sequences with varying TA repeats could optimize transfection efficiency and explain phenotypic differences between isolates from different hosts or regions.

Keywords: *Babesia divergens*, comparative analysis, Elongation Factor 1 alpha (EF-1α), host range

Introduction

Elongation factor 1 alpha (EF-1α) is among the most abundant proteins ever reported to be expressed by eukaryotic cells and is highly conserved [1,2,3]. Traditionally, EF-1α is chiefly acknowledged for its contribution to the translation machinery, facilitating the binding of aminoacyl-tRNA to the ribosomal A site in a GTP-dependent manner [4]. Nevertheless, recent studies have revealed how it functions in various cellular processes such as microtubule severing, ubiquitin dependent proteolysis of N-terminal blocked proteins and cytoskeletal rearrangements [3,5]. In addition, EF-1α plays a crucial role in fundamental cellular events including of programmed apoptosis in which is up-regulated [3,6,7]. Further studies have also highlighted the important role of EF-1α in other apicomplexan parasites, including *Cryptosporidium parvum* and *Toxoplasma gondii*, for the entry process of these parasites into host cells. This suggest that EF-1α could be a promising target for vaccine development against cryptosporidiosis and toxoplasmosis [8,9]. Furthermore, results from studies on vaccination against avian coccidiosis

54 indicate that EF-1α is a promising candidate antigen for eliciting cross-protective immunity against
55 this disease in poultry, highlighting its potential for veterinary use [10].

56 *Babesia* are *T. gondii* and *Plasmodium*-related apicomplexan tick-borne parasites responsible for
57 acute and persistent infections in vertebrate hosts, including cattle and humans. Advancements in
58 gene editing and transfection systems are essential for unraveling the molecular intricacies of
59 *Babesia* parasites and enhancing vaccine development strategies [11,12,13]. Among the crucial
60 components, facilitating these advancements is the EF-1α, which plays a pivotal role in various
61 cellular processes. In both *Plasmodium* and *Babesia* parasites, the *EF-1α* gene exhibits a
62 distinctive organization, featuring two identical head-to-head genes, each separated by intergenic
63 regions [11,14]. The EF-1α intergenic region (EF-1α-IG) hosts robust promoters capable of driving
64 the expression of both *EF-1α* genes efficiently. This unique characteristic of the EF-1α-IG region
65 also enables it to effectively promote the expression of foreign genes in transiently and stably
66 transfected *Babesia bovis* and *Babesia bigemina* parasites [11,12,15]. The utilization of these
67 strong promoters within the EF-1α-IG region has dramatically accelerated the development of
68 transfection and gene editing systems in both *Plasmodium* and *Babesia* parasites [11,16].
69 Researchers have successfully exploited these promoters to introduce foreign genes into parasite
70 genomes, facilitating precise genetic manipulations, the development of novel vaccines, and
71 sophisticated studies on parasite biology. This breakthrough has significantly improved gene
72 function analysis, advanced our understanding of the molecular mechanisms underlying parasite
73 pathogenesis and host-parasite interactions [13,17,18], and became an important tool for
74 developing novel vectored vaccines [19,20,21,22].

75 *Babesia divergens*, one of the three major species of *Babesia* causing babesiosis in cattle, is
76 particularly significant as a zoonotic pathogen playing a role in human babesiosis in Europe, in

77 addition to its impact on cattle [23,24]. Compared to *B. bovis* and *B. bigemina*, limited information
78 is available about the infectivity and pathogenicity of *B. divergens*, leaving many questions
79 unanswered regarding important processes. Understanding the gaps of knowledge on gene
80 function and other mechanisms involved in the biology of *B. divergens* will likely require the
81 identification of novel promoters that can aid in the development of gene editing and transfection
82 systems. Thus far, only one study has drawn attention in this regard, describing a stable transfection
83 system for *B. divergens* based on the promoter of the *elongation factor Tu GTP binding domain*
84 *family protein* (*ef-tgtp*) gene [25]. Unlike other *Babesia* species found in cattle, *B. divergens* has a
85 broader host spectrum, including humans. We hypothesize that the EF-1α locus of *B. divergens,* a
86 parasite able to infect bovine and human hosts, exhibits structural and regulatory variations among
87 different isolates, which may be linked to its host adaptability, evolutionary divergence, and
88 geographic distribution. Investigating these differences will enhance our understanding of *B.*
89 *divergens* biology and provide insights into potential mechanisms of gene regulation and parasite-
90 host interactions.

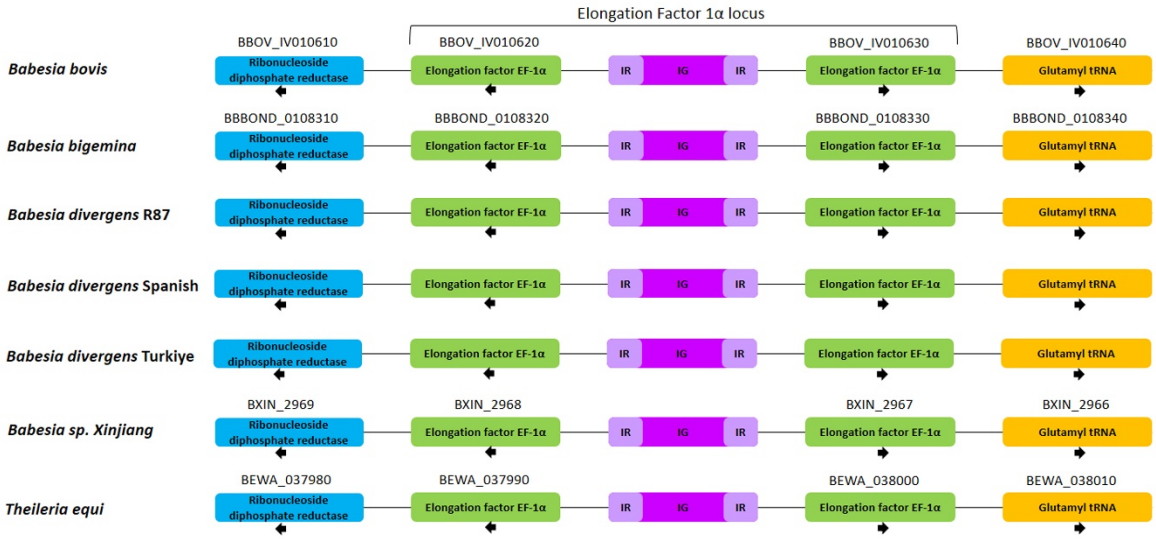
91 Here, based on existing genome data and NGS and Sanger sequencing, we focused on conducting
92 a comparative analysis of the EF-1α locus from *B. divergens* isolates obtained from humans and
93 cattle in different geographical regions, using the *B. divergens* Rouen 87 (R87) genome as a
94 reference [26,27]. Our approach focused on conducting a comparative analysis of the EF-1α locus,
95 which serves as a pivotal component in transfection investigations, among *B. divergens* isolates
96 sourced from infected bovines in Turkiye and from patients with severe babesiosis in Spain [28].
97 The data generated in this study implies the possible influence of evolutionary selective forces
98 shaping key sequence and structural differences in the EF-1α locus between these two isolates.

99

100 **Results**

101 **EF-1α locus organization and synteny in *Babesia divergens***

102 In apicomplexan parasites, the *EF-1α* gene locus predominantly resides between the
103 *Ribonucleoside diphosphate reductase* and *Glutamyl tRNA* genes as demonstrated by conserved
104 synteny between *Theileria* and *Babesia* species. Similar to *B. bovis*, *B. bigemina*, *Babesia* sp.
105 Xinjiang and *Theileria equi*, the *B. divergens* R87 and Spanish and Turkiye isolates also possess
106 two identical *EF-1α* genes located within the EF-1α locus (Figure 1). The biological significance
107 of remarkable conserved synteny among all these apicomplexan parasites remains uninvestigated.



108

109 **Figure 1:** Schematic representation of gene localization and synteny maps of the *EF-1α* gene
110 locus.

111 **EF-1α coding sequence conservation across *Babesia divergens* isolates**

112 Comparative analysis of the EF-1α locus in *B. divergens* isolates (R87, Spanish, and Turkiye
113 isolates) revealed that the *EF-1α* gene sequences (1347 bp) were 100% identical across all isolates.
114 Each of the two EF-1α ORFs from three isolates of *B. divergens* encodes identical proteins of 448

115 amino acids, with an estimated molecular mass of 49.39 kDa, which is the same size as those
116 encoded by *B. bovis* (XP_001610983.1), *B. bigemina* (XP_012766720.1), *Babesia caballi*
117 (GIX62688.1), *Babesia gibsoni* (KAK1443338.1), *Babesia duncani* (KAK2195768.1), *Babesia*
118 *ovis* (GFE54231.1), *Babesia* sp. Xinjiang (XP_028869924.1), *Theileria parva* (XP_766247.1),
119 *Theileria annulata* (XP_954051.1), and *Theileria orientalis* (UKK00547.1). The EF-1α *B.*
120 *divergens* R87, Spanish, and Turkiye ORFs show identical sequences (100% identity) and share
121 significant similarity with homologous sequences found in *B. gibsoni* (97.99% identity), *B. caballi*
122 (97.54% identity), *B. bigemina* (96.65% identity), *Babesia* sp. Xinjiang (96.43% identity), *B. bovis*
123 (94.87% identity), *B. ovis* (94.20% identity), *B. duncani* (93.97% identity), *T. orientalis* (92.41%
124 identity), *T. parva* (89.51% identity), *T. annulata* (89.06% identity), and *B. microti* (86.58%
125 identity) (Supplementary Figure 1). Notably, key regions involved in guanosine 50-triphosphate
126 (GTP) binding are conserved across these organisms, as depicted in Supplementary Figure 2. A
127 phylogenetic analysis using the EF-1α protein sequence of R87, Spanish, and Turkiye isolates,
128 along with other related piroplasma species, showed that both isolates clustered within the same
129 clade as *B. gibsoni* (Figure 2a). Similarly, in the phylogenetic analysis of *B. divergens* R87,
130 Spanish (MG944238.1), and Turkiye (PP663285.1) *18S rRNA* gene sequences, it was observed
131 that they were in the same clade with *B. divergens* isolates obtained from different hosts from
132 GenBank (Figure 2b).

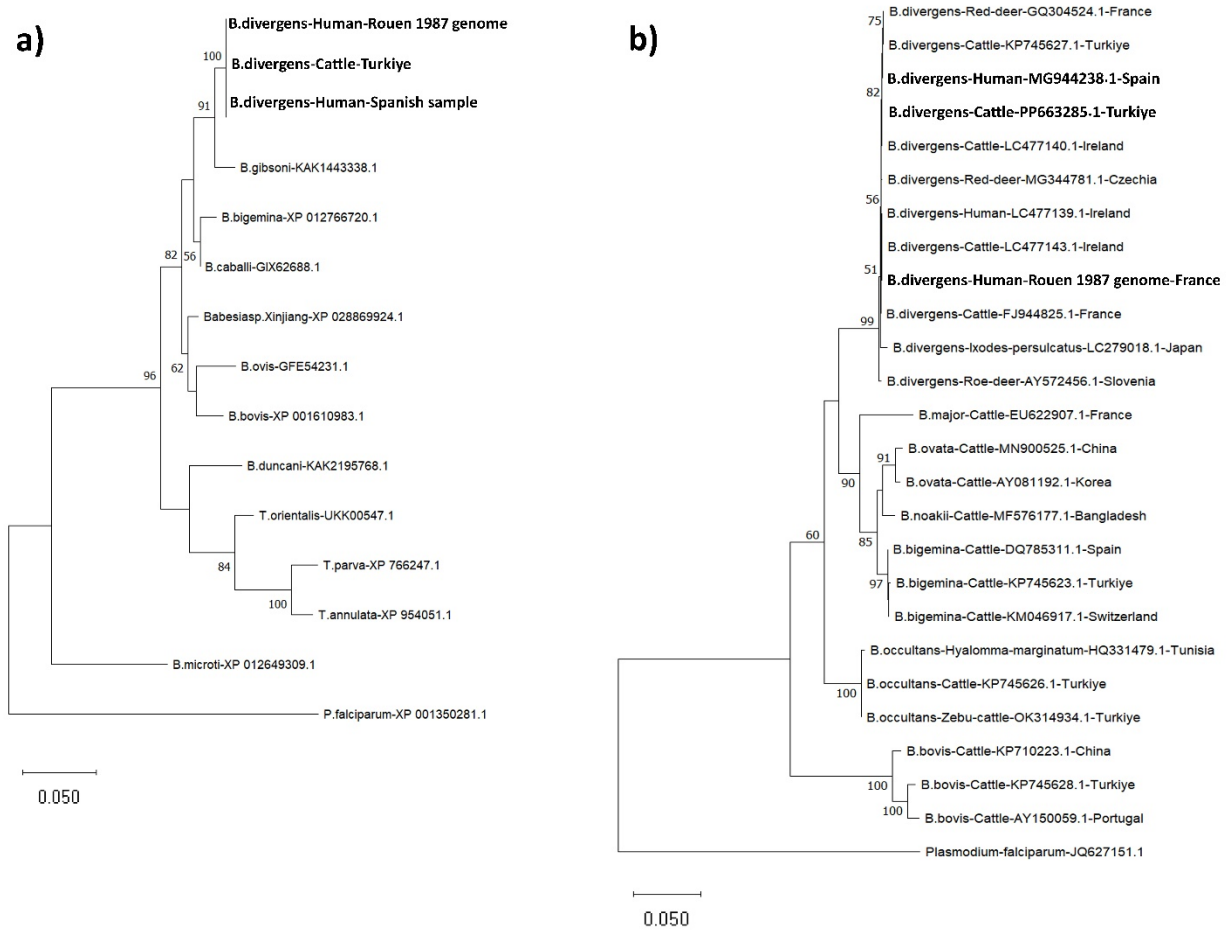
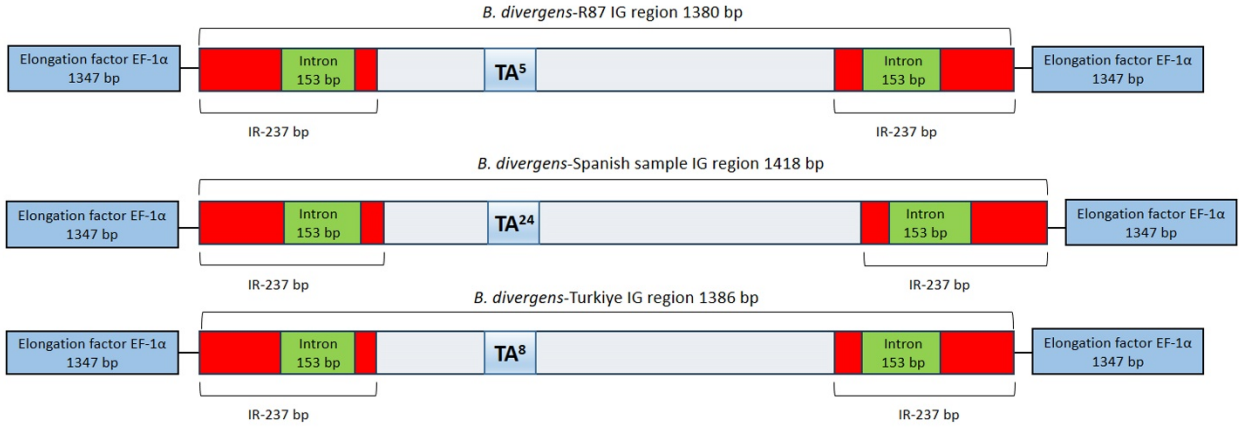


Figure 2: a) Phylogenetic analyses of piroplasm EF-1α protein sequences by maximum likelihood method based on the Le_Gascuel_2008 model [29]. Sequences were obtained from the GenBank database. The EF-1α protein sequence of *Plasmodium falciparum* was utilized as an outgroup. b) Phylogenetic analyses of piroplasm *18S rRNA* gene sequences by maximum likelihood method based on the Tamura-Nei model [30]. The *18S rRNA* gene sequence of *P. falciparum* was utilized as an outgroup. *Babesia divergens* R87, Spanish, and Turkiye isolates of EF-1α are highlighted in bold.

143 **Features, sequence alignment and nucleotide variability in IG region**

144 The EF-1α-IG regions of the three distinct *B. divergens* isolates analyzed hereby are not related in
145 sequence to the IG regions of other related *Babesia* parasites. However, the *B. divergens* EF-1α-
146 IG regions do have structural similarities with other *Babesia* parasites, including the presence of
147 terminal inverted repeat (IR) regions of 237 bp, which were conserved among all isolates. In
148 addition, similar to what was found on *B. bovis* and *B. bigemina* [11,12], the IR regions harbor a
149 153 bp intron, maintaining sequence integrity across isolates (Figure 3).

150 However, the comparative analysis of the EF-1α locus sequences in the three *B. divergens* isolates
151 (R87, Spanish, and Turkiye isolates) revealed sequence differences in the IG region. While the
152 *EF-1α* gene sequences were found to be 100% identical among all isolates, the IG region length
153 varied across the isolates, with the R87 strain containing 5 TA repeats and measuring 1380 bp, the
154 Spanish isolate harboring 24 TA repeats with an IG region of 1418 bp, and the Turkiye isolate
155 exhibiting 8 TA repeats with an IG region of 1386 bp (Figure 3). Overall, these findings indicate
156 notable differences in the IG region between *B. divergens* isolates originating from different
157 geographical regions and host sources.



158

Figure 3: Schematic representation of the EF-1α-IG region between *B. divergens* isolates. The terminal inverted repeat (IR) regions (red), intron (green, 153 bp), and TA repeats (blue) are highlighted. The IG region length varies due to differences in TA repeat numbers, with 5 TA repeats in R87 (1380 bp), 24 in the Spanish isolate (1418 bp), and 8 in the Turkiye isolate (1386 bp).

Multiple sequence alignment of the IG region across the three *B. divergens* isolates further confirmed variations in TA repeat numbers and additional nucleotide substitutions. The TA repeat sequences were located around position ~513–516 bp across all isolates but varied in number. Besides TA repeat variations, additional sequence polymorphisms were identified across the IG region (Figure 4). The results were consistent and validated using Nanopore and Sanger sequencing technologies.

```
B. divergens R87            TTTGATGAGGTCTAATAATAACTAATCTGCAAAGGGTTAGTGAGGCCAATGGAAACACAA  60      ATCTCATCTGAGAACTACAGTCTATATAACTGTGCGGGTATCTTATGTTACCTTGAGTTT  742
B. divergens Spanish sample TTTGATGAGGTCTAATAATAACTAATCTGCAAAGGGTTAGTGAGGCCAATGGAAACACAA  60      ATCTCATCTGAGAACTACAGTCTATATAACTGTGCGGGTATCTTATGTTACCTTGAGTTT  780
B. divergens Turkiye        TTTGATGAGGTCTAATAATAACTAATCTGCAAAGGGTTAGTGAGGCCAATGGAAACACAA  60      ATCTCATCTGAGAACTACAGTCTATATAACTGTGCGGGTATCTTATGTTACCTTGAGTTT  748
                            ************************************************************          ************************************************************

B. divergens R87            TGGGCAACACCACAACCGAATTAACAACATGGAAATATATATTTATAATGTTATGCCCCT  120     AGAATATTTGTACGCTTATCCGTATAATTAACATCGTCGCGTTTAGTGGCTTAAGGCACC  802
B. divergens Spanish sample TGGGCAACACCACAACCGAATTAACAACATGGAAATATATATTTATAATGTTATGCCCCT  120     AGAATATTTGTACGCTTATCCGTATAATTAACATCGTCGCGTTTAGTGGCTTAAGGCACC  840
B. divergens Turkiye        TGGGCAACACCACAACCGAATTAACAACATGGAAATATATATTTATAATGTTATGCCCCT  120     AGAATATTTGTACGCTTATCCGTATAATTAACATCGTCGCGTTTAGTGGCTTAAGGCACC  808
                            ************************************************************          ************************************************************

B. divergens R87            ACGACACTTCTCATAATGACACAGCAGAACGTAGAGATCACGACACCCACACCACTAACA  180     TATATCTATGTGTACTTGCGCAAATCAAGTTGATATTTCGAACGTGATAATTGATGAGAT  862
B. divergens Spanish sample ACGACACTTCTCATAATGACACAGCAGAACGTAGAGATCACGACACCCACACCACTAACA  180     TATATCTATGTGTACTTGCGCAAATCAAGTTGATATTTCGAACGTGATAATTGATGAGAT  900
B. divergens Turkiye        ACGACACTTCTCATAATGACACAGCAGAACGTAGAGATCACGACACCCACACCACTAACA  180     TATATCTATGTGTACTTGCGCAAATCAAGTTGATATTTCGAACGTGATAATTGATGAGAT  868
                            ************************************************************          ************************************************************

B. divergens R87            TCAAAAAACCAAACAAACCAAAACAGTTTGACAAAACATACCTTTTGTTTAGCGAATGAA  240     AAGCCACATATTCAATGAAAATAATGTAAAATGCAACACACATGTGACTCAATGTAAACT  922
B. divergens Spanish sample TCAAAAAACCAAACAAACCAAAACAGTTTGACAAAACATACCTTTTGTTTAGCGAATGAA  240     AAGCCACATATTCAATGAAAATAATGTAAAATGCAACACACATGTGACTCAATGTAAACT  960
B. divergens Turkiye        TCAAAAAACCAAACAAACCAAAACAGTTTGACAAAACATACCTTTTGTTTAGAGAATGAA  240     AAGCCACATATTCAATGAAAATAATGTAAAATGCAACACACATGTGACTCAATGTAAACT  928
                            ************************************************************          ************************************************************

B. divergens R87            TTTTTAGGCACGTCAGTAGTGTAAACCCTAAGTTGAAAATAGAGTGTTCCTACATATAGA  300     ATGCATAAATTAATGGCACATTGTATTGCACATAGGCAAAAATAGTATTAATAACTGACA  982
B. divergens Spanish sample TTTTTAGGCACGTCAGTAGTGTAAACCCTAAATTGAAAATAGAGTGTTCCTACATATAGA  300     ATGCATAAATTAATGGCACATTGTATTGCACATAGGCAACAATAGTATTAATAACTGACA  1020
B. divergens Turkiye        TTTTTAGGCACGTCAGTAGTGTAAACCCTAAATTGAAAATAGAGTGTTCCTACATATAGA  300     ATGCATAAATTAATGGCACATTGTATTGCACATAGGCAAAAATAGTATTAATAACTGACA  988
                            ***************************** ************************************          ************************************ ********************

B. divergens R87            GCACTAATATATATGGGAGGATAATACTGTGGGGTTTCGTGGAAGCCACGCTAGACAATA  360     TAATTAATGGTTTATTGTCTAATGTGTTGGAACCAAATTATGAGGCATTGTCCATTCCTA  1042
B. divergens Spanish sample GCACTAATATATATGGGAGGATAATACTGTGGGGTTTCGTGGAAGCCACGCTAGACAATA  360     TAATTAATGGTTTATTGTCTAATGTGTTGGAACCAAATTATGAGGCATTGTCCATTCCTA  1080
B. divergens Turkiye        GCACTAATATATATGGGAGGATAATACTGTGGGGTTTCGTGGAAGCCACGCTAGACAATA  360     TAATTAATGGTTTATTGTCTAATGTGTTGGAACCAAATTATGAGGCATTGTCCATTCCTA  1048
                            ************************************************************          ************************************************************

B. divergens R87            AGTTAGCGAAACGGTTGAAGATATTATGGTCTGAAGCATACTCATGCACATCGTCTCATG  420     GCCGCACCACAATTGTTAGATTGAGAGTGACTACTAATATTGTCATCTGTTTTAAGCACA  1102
B. divergens Spanish sample AGTTAGCGAAACGGTTGAAGATATTATGGTCTGAAGCATACTCATGCACATCGTCTCATG  420     GCCGCACCACAATTGTTAGATTGAGAGTGACTACTAATATTGTCATCTGTTTTAAGCACA  1140
B. divergens Turkiye        AGTTAGCGAAACGGTTGAAGATATTATGGTCTGAAGCATACTCATGCACATCGTCTCATG  420     GCCGCACCACAATTGTTAGATTGAGAGTGACTACTAATATTGTCATCTGTTTTAAGCACA  1108
                            ************************************************************          ************************************************************

B. divergens R87            TTACAGCAATTTTAATTAATTCGGTAGCTTGTGAATATGTTGTTATGCTGGCATCATATA  480     CGTATAGCGATTAACTATACGAGTTGTATTTCTTCCGTTATATTCTCTAAACAAAAGGTA  1162
B. divergens Spanish sample TTACAGCAATTTTAATTAATTCGGTAGCTTGTGAATATGTTGTTATGCTGGCATCATATA  480     CGTATAGCGATTAACTATACGAGTTGTATTTCTTCCGTTATATTCTCTAAACAAAAGGTA  1200
B. divergens Turkiye        TTACAGCAATTTTAATTAATTCGGTAGCTTGTGAATATGTTGTTATGCTGGCATCATATA  480     CGTATAGCGATTAACTATACGAGTTGTATTTCTTCCGTTATATTCTCTAAACAAAAGGTA  1168
                            ************************************************************          ************************************************************

B. divergens R87            AACTGCTATTATTTATCTATATCTATCTTTG-----------------------------  513     TGTTTTGTCAAACTGTTTTGGTTTGTTTGGTTTTTTGATGTTAGTGGTGTGGGTGTCGTG  1222
B. divergens Spanish sample AACTGCTATTATTTATCTATATCTATCTTTGTATATATATATATATATATATATATAT  540     TGTTTTGTCAAACTGTTTTGGTTTGTTTGGTTTTTTGATGTTAGTGGTGTGGGTGTCGTG  1260
B. divergens Turkiye        AACTGCTATTATTTATCTATATCTATCTTTG-----------------------------  516     TGTTTTGTCAAACTGTTTTGGTTTGTTTGGTTTTTTGATGTTAGTGGTGTGGGTGTCGTG  1228
                            ******************************                                          ************************************************************

B. divergens R87            ---------TATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT  562     ATCTCTACGTTCTGCTGTGTCATTATGAGAAGTGTCGTAGGGGCATAACATTATAAATAT  1282
B. divergens Spanish sample ATATATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT  600     ATCTCTACGTTCTGCTGTGTCATTATGAGAAGTGTCGTAGGGGCATAACATTATAAATAT  1320
B. divergens Turkiye        ---TATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT  568     ATCTCTACGTTCTGCTGTGTCATTATGAGAAGTGTCGTAGGGGCATAACATTATAAATAT  1288
                                                 ********************************************          ************************************************************

B. divergens R87            ACATTGATGTTATCTACAAAGCAGCTGAGGGTATTGTATCACTGTGGCTTATGGCATGAT  622     ATATTTCCATGTTGTTAATTCGGTTGTGGTGTTGCCCATTGTGTTTCCATTGGCCTCACT  1342
B. divergens Spanish sample ACATTGATGTTATCTACAAAGCAGCTGAGGGTATTGTATCACTGTGGCTTATGGCATGAT  660     ATATTTCCATGTTGTTAATTCGGTTGTGGTGTTGCCCATTGTGTTTCCATTGGCCTCACT  1380
B. divergens Turkiye        ACATTGATGTTATCTACAAAGCAGCTGAGGGTATTGTATCACTGTGGCTTATGGCATGAT  628     ATATTTCCATGTTGTTAATTCGGTTGTGGTGTTGCCCATTGTGTTTCCATTGGCCTCACT  1348
                            ************************************************************          ************************************************************

B. divergens R87            CATTCAATTCTATCACAGTGTGGGCGTTTTGCTCTAGGTCGAGTTACATGTTCTAATTTT  682     AACCCTTTGCAGATTAGTTATTATTAGACCTCATCAAA 1380
B. divergens Spanish sample CATTCAATTCTATCACAGTGTGGGCGTTTTGCTCTAGGTCGAGTTACATGTTCTAATTTT  720     AACCCTTTGCAGATTAGTTATTATTAGACCTCATCAAA 1418
B. divergens Turkiye        CATTCAATTCTATCACAGTGTGGGCGTTTTGCTCTAGGTCGAGTTACATGTTCTAATTTT  688     AACCCTTTGCAGATTAGTTATTATTAGACCTCATCAAA 1386
                            ************************************************************          ************************************
```

Figure 4: Multiple sequence alignment of the IG region within the EF-1α locus between *B. divergens* isolates. Nucleotide variations are highlighted with black boxes, while differences in TA

173 repeat numbers are also marked. In addition to *B. divergens*, the intergenic region was examined
174 in other zoonotic *Babesia* species, including *Babesia* MO1 and *B. duncani*. While the EF-1α locus
175 in *B. duncani* lacked TA repeats, the *Babesia* MO1 genome contained a short stretch of 3 TA
176 repeats. These findings indicate that the presence and expansion of TA repeats may be specific to
177 *B. divergens* (Supplementary Figure 3).

178

179 **Validation of TA repeat variations in the IG region**

180 To confirm the sequence differences in the TA repeat region among *B. divergens* isolates, a
181 combination of PCR amplification, Sanger sequencing, and NGS approaches was employed. For
182 the R87 isolate, the IG region was amplified using the Bdiv-Ef-For1/IGRev1 primer pair, yielding
183 an amplicon of approximately 2.8 kb, which was subsequently validated by Sanger sequencing.
184 The repeat structure identified in the sequencing data confirmed the presence of 5 TA repeats,
185 consistent with initial findings. For the Turkiye isolate, the IG region was amplified in two
186 overlapping fragments using the Bdiv-Ef-For1/Sez-For1 and Sez-For1-Rev/GlutamylR1 primer
187 sets. These amplicons were subjected to NGS analysis, which verified the 8 TA repeats and
188 confirmed the overall sequence integrity of the IG region. For the Spanish isolate, initial attempts
189 to amplify the full IG region using Bdiv-Ef-For1/IGRev1 primers were unsuccessful, likely due to
190 sequence divergence at the primer binding sites. An alternative primer pair, IGLMFR1/IGRev1,
191 was used to amplify ~1 kb fragment, which was validated through Sanger sequencing. The results
192 confirmed the presence of 24 TA repeats, aligning with the Nanopore sequencing data.

```
B. divergens R87 Nanopore          AACTGCTATTATTTATCTATATCTATCTTTG-----------------------------------TATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   562
B. divergens R87 Sanger            AACTGCTATTATTTATCTATATCTATCTTTG-----------------------------------TATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   562
B. divergens Spanish sample Nanopore AACTGCTATTATTTATCTATATCTATCTTTGTATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   600
B. divergens Spanish sample Sanger AACTGCTATTATTTATCTATATCTATCTTTGTATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   600
B. divergens Turkiye Nanapore      AACTGCTATTATTTATCTATATCTATCTTTG-------------------------------TATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   568
B. divergens Turkiye NGS           AACTGCTATTATTTATCTATATCTATCTTTG-------------------------------TATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   568
                                   ******************************                                  ***************************************************
```

Figure 5: Multiple sequence alignment of the TA repeat region across *B. divergens* isolates using Nanopore, Sanger, and NGS sequencing methods. The black-highlighted regions indicate TA repeats.

Discussion

Since the discovery of *Babesia*, the availability of transfection and gene editing methods arguably emerged as the most significant developments, after *Babesia in vitro* cultures. Transfection technologies (together with related gene editing methods), which involves the introduction of foreign genetic material into cells, has revolutionized the study of various diseases, including those caused by *Babesia* parasites [15]. This technique has enabled researchers to manipulate cellular processes, elucidate pathogenic mechanisms, and explore potential therapeutic interventions. Thus, transfection represents a crucial tool in advancing our understanding of *Babesia* infections and developing strategies for their prevention and treatment [13,17]. EF-1α-IG region contains bi-directional promoters capable of efficiently driving expression of *EF-1α* genes. Their utilization has significantly advanced transfection systems in *Babesia* [11,15]. In this study, we explored the structural and sequence variability of the EF-1α-IG regions across three *B. divergens* isolates from humans and cattle to better understand its potential role in gene regulation and parasite adaptation. The findings indicated that while the EF-1α coding sequences remain highly conserved, there is variability in the IG region, specifically in TA repeat numbers and the length of the terminal IR

212 sequences, suggesting that these elements could contribute to transcriptional regulation and
213 genomic stability.

214 The high level of conservation of the *EF-1α* gene observed in different strains of *B. divergens* and
215 its similarity to the EF-1α protein of the two other pathogenic bovine babesiosis species, *B. bovis*
216 and *B. bigemina*, highlights the potential use of EF-1α in subunit vaccines. These findings
217 revealed that EF-1α protein has potential to be a suitable antigen for vaccine development, as it
218 was previously suggested for other important protozoa such as *T. gondii* [9] and *C. parvum* [8],
219 which are also members of the apicomplexan phylum. Additionally, the broad applicability of the
220 vaccine in suppressing different parasitic infections was further verified by the existing cross-
221 protective immunity against *Eimeria tenella* and *Eimeria maxima* in broiler chickens vaccinated
222 with recombinant EF-1α protein [10]. Thus, EF-1α protein appears as a good candidate for the
223 development of vaccines against babesiosis, taking advantage of its potential to induce cross-
224 immunity against different species.

225 Comparisons of the IG regions among the three *B. divergens* isolates (R87, Spanish, and Turkiye
226 isolates) provided novel insights into the structural organization of this region. This study
227 challenged a previously established assumption that the IG regulatory regions of the EF-1α loci
228 are fully conserved among *Babesia* spp. and parasite isolates, both in terms of sequence and
229 structural organization. Despite having identical *EF-1α* gene ORFs organized in a typical head-to-
230 head fashion, *B. divergens* isolates exhibited striking differences in their IG regions, despite
231 sharing identical *18S rRNA* sequences. One of the most unexpected findings of this study was the
232 variability in TA repeat numbers within the IG region. The Spanish isolate exhibited the highest
233 number of TA repeats (n=24), while the Turkiye isolate contained 8, and the R87 strain had only
234 5. Given that repeat sequences in intergenic regions can affect promoter activity, transcription

235 factor binding, and chromatin accessibility [31], these differences may influence *EF-1α* gene
236 expression. The conserved 237 bp terminal IR region and the 153 bp intron within the untranslated
237 region suggest that the IG region may have multiple regulatory roles. Inverted repeats (IRs) are
238 often associated with promoter function, transcriptional regulation, while intronic sequences
239 within untranslated regions can influence alternative splicing, mRNA stability, or post-
240 transcriptional control. The conservation of this structure across isolates suggests its functional
241 importance in *EF-1α* gene regulation.

242 As mentioned, the differences within the IG regions among *B. divergens* isolates analyzed hereby
243 predominantly occur in the initial segment of the putative promoter regions. Although
244 investigations into EF-1α-IG region variations across different strains remain limited, examination
245 of these regions from *B. bovis* strains in Argentina, Mexico, and Australia suggests a notable
246 degree of conservation among them [11]. The distinctions between *B. bovis* and *B. divergens*,
247 despite both being etiological agents of babesiosis in cattle, are evident in their geographical
248 distribution, different tick vector preferences, and the clinical manifestations of babesiosis [23,32].
249 Moreover, the broader host spectrum of *B. divergens*, including its zoonotic potential, highlights
250 its adaptability and significance beyond the confines of bovine hosts. It is thus logical to speculate
251 that the discrepancies observed in the EF-1α-IG region could be attributed to the differing host
252 preferences of *B. divergens*. This is particularly noteworthy in our study where R87, Spanish, and
253 Turkiye isolates, whose elongation loci we compared, were obtained from humans (R87, and
254 Spanish isolates) and cattle (Turkiye) respectively. Surprisingly, the EF-1α-IG region from
255 parasites that infect humans presented differences that could also be related to virulence or
256 adaptation to laboratory conditions. Such findings raise intriguing questions regarding factors
257 involved in parasite evolution, and the potential influence of host-specific factors on the genetic

258 makeup and regulatory mechanisms of *Babesia* species, warranting further investigation into the
259 intricate interplay between pathogen genetics and host environments. As they conflict with a
260 previously established paradigm, the substantial differences found in the IG regions of the *EF-1α*
261 genes between three different strains of *B. divergens*, as confirmed by *18S rRNA gene* analysis are
262 found for the first time among piroplasmid parasites. Moreover, as previously noted, these
263 differences may reflect important adaptive events affecting the development of the parasites
264 among distinct invertebrate and/or vertebrate hosts.

265 Surprisingly, upon expanding our comparative analysis of the EF-1α IG regions found in other
266 zoonotic *Babesia* species, *B. duncani* and *Babesia* MO1, striking differences were also noted in
267 their EF-1α IG regions. *Babesia duncani*, on the other hand, contained no TA repeats, while only
268 3 TA motifs were present in the MO1 isolate. This suggests that TA repeat expansion within the
269 EF-1α-IG region may represent a species-specific regulatory feature rather than a common trait of
270 all zoonotic *Babesia*. Further comparative studies across more isolates will be essential to
271 determine whether these differences contribute to host tropism, gene expression variability, or
272 virulence mechanisms.

273 **Conclusion**

274 This study provides new insights into the structural variability of the EF-1α-IG region in *B.*
275 *divergens* isolates from different hosts, revealing differences in TA repeat numbers that may
276 contribute to gene regulation and parasite adaptation. While our findings highlight the potential
277 regulatory role of this region, functional validation through promoter activity assays, gene
278 expression analysis, and transfection studies is still needed to determine its precise impact on EF-
279 1α expression. Given the zoonotic nature of *B. divergens*, future research should explore whether
280 similar variations in the IG region exist in *B. divergens* isolates from cattle and humans.

281 Additionally, the gerbil model could be used to investigate potential intergenic region variations
282 across different vertebrate hosts to better understand the role of this region in host adaptation.
283 Further investigation of the functional implication of these variations may provide valuable
284 insights into transcriptional control across diverse parasite strains, provide valuable information
285 on the mechanisms driving *B. divergens* evolution and improve transfection strategies for
286 functional genomic studies. Overall, this study lays the groundwork for future research into gene
287 regulation and host-pathogen interactions in zoonotic *Babesia* species.

288 **Methods**

289 ***Babesia divergens* - Rouen 87 Strain (R87)**

290 *Babesia divergens* R87 is a reference strain originally isolated from a human case of babesiosis in
291 France [33]. This strain has been widely used in molecular and genomic studies due to its well-
292 characterized genome sequence [27]. The EF-1α locus data from this strain were utilized in this
293 study, obtained from the *B. divergens* genome assembly described by [27].

294 ***Babesia divergens*-Spanish sample**

295 *Babesia divergens* Spanish strain was isolate from a blood sample of elderly patient from Asturias,
296 Spain who suffers a fulminant fatal babesiosis [28]. The blood sample was cultured in vitro in
297 human A+ RBCs and RPMI 1640 (Gibco-Life Technology, 11875093) supplemented with 10%
298 human serum (Gibco, 11021037), 7.5% (w/v) sodium bicarbonate solution (Lonza Group Ltd,
299 Basel, Switzerland, 144-55-8), and 100 μmol/L hypoxanthine (Sigma-Aldrich Corporation, St
300 Louis, MO, H9377) at a pH of 7.3. Cells were cultured at 37°C in a humidified atmosphere of 5%
301 CO2. The culture medium was replaced every 24 h, and parasitemia was monitored by examining
302 Giemsa-stained blood smears by a light microscope.

303 ***Babesia divergens*-Turkiye**

304 *Babesia divergens*-Turkiye isolate was obtained from a 3-year-old cow exhibiting severe
305 babesiosis symptoms, including high fever, anemia, and hemoglobinuria, in Bartin province of
306 Turkiye in 2023. Microscopic examination revealed a parasitemia level of 1.5%. Following the
307 collection of blood stabilate from the animal, treatment was administered using imidocarb
308 dipropionate [34]. To determine the species of *Babesia*, Reverse Line Blot (RLB) analysis was
309 conducted. Total DNA extraction from blood samples was performed using the PureLink™
310 Genomic DNA Mini Kit (Invitrogen Corporation, Carlsbad, United States) according to the
311 manufacturer's instructions. The hypervariable V4 region of the piroplasm *18S rRNA* gene was
312 then amplified using primers RLB-F2 and RLB-R2-biotin [35] specifically for the RLB assay.
313 RLB was carried out on the PCR product, following the previously described method [36]. After
314 amplification, 20 µL of all PCR products obtained from each DNA sample were diluted to a final
315 volume of 150 µL with 2X SSPE/0.1% SDS buffer. For RLB hybridization, the samples were
316 heated at 95–100°C for 10 minutes in a Thermal Cycler and rapidly cooled on ice. Subsequently,
317 the PCR products were hybridized with probes specific to the genera and species of *Babesia* and
318 *Theileria*, which were linked to an RLB membrane [36].

319 ***Babesia MO1 and Babesia duncani***

320 In this study, we used the EF-1α locus data of these species obtained from the *Babesia* MO1 [37]
321 and *B. duncani* [38,39] assembly. Specifically, Bioproject PRJNA1032622 for *Babesia MO1* and
322 Bioproject PRJNA821606 and locus OL804102 for *B. duncani.* To assess the presence of TA
323 repeats in *Babesia* MO1 and *B. duncani*, we retrieved the EF-1α locus from *Babesia* MO1 and *B.*
324 *duncani* genome assemblies (PRJNA1032622 and PRJNA821606, respectively). The intergenic
325 regions were identified by locating the flanking genes (*ribonucleoside diphosphate reductase* and

326 *glutamyl tRNA synthetase*) and were subsequently aligned to detect potential tandem repeat
327 regions. The presence or absence of TA repeats was manually inspected by multiple sequence
328 alignment.

329 **Nanopore sequencing**

330 The total DNA from *Babesia* samples was used to prepare Oxford Nanopore libraries with the
331 SQK-LSK114 Ligation Sequencing Kit for 24 h in the MinION Platform with an R10.4 flow cell
332 following the manufacturer's protocol (Oxford Nanopore Technologies, Oxford, United
333 Kingdom). The base calling was performed using the software Dorado v0.8.3 with the duplex
334 option. The sequencing and basecalling was carried out at the Unidad Universitaria de
335 Secuenciación Masiva y Bioinformática, Instituto de Biotecnología, Universidad Nacional
336 Autónoma de México (Cuernavaca, Mexico).

337 **Designing of primers for EF-1α locus analysis and PCR**

338 Primers were designed utilizing sequences from *Ribonucleoside diphosphate reductase*
339 (Bdiv_023390c) and *Glutamyl tRNA* (Bdiv_023400) genes obtained from the Piroplasma DB
340 (https://piroplasmadb.org) and nanopore sequencing data to amplify the EF-1α locus. All primers
341 were designed using Primer Quest™ Tool (https://www.idtdna.com/pages/tools/primerquest)
342 (Supplementary Table 1, and Supplementary Figure 4). PCR amplification was carried out using
343 Phusion® High-Fidelity PCR Master Mix with GC Buffer (#M0532S; NEB). The PCR was
344 performed in a total reaction volume of 20 µL containing 10 µL of 2X Phusion Master Mix, 1 µL
345 of each forward and reverse primer, 1 µL of template DNA, 7 µL of nuclease-free water.

346 ***De novo* assembling**

347 Sequencing library was prepared using Nextera XT DNA Library Preparation Kit and sequencing
348 was performed by Illumina Miseq platform as paired end (PE) 2x150 bases reads. Raw NGS reads
349 (FASTQ) were quality checked by FASTQC [40] and trimmed by Trimmomatic v0.32 [41].
350 Demultiplexing and low-quality read filtering were performed via CLC Genomics Workbench
351 (Qiagen, Hilden, Germany). The *de novo* assembly constructed by CLC Genomic's de novo
352 assembly module with the parameters of minimum contig size as 200 bp, mismatch cost is 2, insert
353 cost is 3, deletion cost is 3, length fraction is 0.5, similarity fraction is 0.8, and paired read input
354 as minimum distance of 180.

355 **Sanger sequencing**

356 To validate the assembled EF-1α-IG region, Sanger sequencing was performed on PCR amplified
357 IG regions from the *B. divergens* isolates. PCR products were purified using the mi-Gel Extraction
358 Kit (Metabion international AG, Steinkirchen Germany) and sequenced using an ABI PRISM
359 3730XL DNA Analyzer (Applied Biosystems, San Francisco, CA, USA). The resulting
360 chromatograms were analyzed using manually inspected, corrected, and edited using Chromas Pro
361 program (McCarthy, Queensland, Australia) and Laser Gene 12.1 program (DNAStar, Madison,
362 WI, USA), and sequences were aligned against the de novo assembled IG region to confirm
363 accuracy and resolve potential misassembles.

364

365 **Phylogenetic analysis**

366 Phylogenetic relationships were inferred using Molecular Evolutionary Genetics Analysis (MEGA
367 11) [42]. The maximum likelihood (ML) method was employed to construct phylogenetic trees

368 based on EF-1α protein sequences and *18S rRNA* gene sequences from various piroplasm species.
369 Bootstrap support values were calculated with 1,000 replicates to assess tree robustness.

370 **Data availability**

371 All data is provided within the manuscript or supplementary information files.

372 **References**

373 1. Slobin, L. I. (1980). The Role of Eucaryotic Elongation Factor Tu in Protein Synthesis: The
374 Measurement of the Elongation Factor Tu Content of Rabbit Reticulocytes and Other Mammalian
375 Cells by a Sensitive Radioimmunoassay. European Journal of Biochemistry 110, 555–563. doi:
376 10.1111/j.1432-1033.1980.tb04898.x
377 2. Merrick, W. C. (1992). Mechanism and regulation of eukaryotic protein synthesis. Microbiol
378 Rev 56, 291–315. doi: 10.1128/mr.56.2.291-315.1992
379 3. Sasikumar, A. N., Perez, W. B., and Kinzy, T. G. (2012). The many roles of the eukaryotic
380 elongation factor 1 complex. WIREs RNA 3, 543–555. doi: 10.1002/wrna.1118
381 4. Riis, B., Rattan, S. I., Clark, B. F., and Merrick, W. C. (1990). Eukaryotic protein elongation
382 factors. Trends in biochemical sciences 15, 420–424.
383 5. Stapulionis, R., Kolli, S., and Deutscher, M. P. (1997). Efficient mammalian protein synthesis
384 requires an intact F-actin system. Journal of Biological Chemistry 272, 24980–24986.
385 6. Chen, E., Proestou, G., Bourbeau, D., and Wang, E. (2000). Rapid up-regulation of peptide
386 elongation factor EF-1α protein levels is an immediate early event during oxidative stress-induced
387 apoptosis. Experimental cell research 259, 140–148.
388 7. Lamberti, A., Longo, O., Marra, M., Tagliaferri, P., Bismuto, E., Fiengo, A., et al. (2007). C-
389 Raf antagonizes apoptosis induced by IFN-α in human lung cancer cells by phosphorylation and

390 increase of the intracellular content of elongation factor 1A. Cell Death & Differentiation 14, 952–
391 962.
392 8. Matsubayashi, M., Teramoto-Kimata, I., Uni, S., Lillehoj, H. S., Matsuda, H., Furuya, M., et al.
393 (2013). Elongation factor-1α is a novel protein associated with host cell invasion and a potential
394 protective antigen of Cryptosporidium parvum. Journal of Biological Chemistry 288, 34111–
395 34120.
396 9. Wang, S., Zhang, Z., Wang, Y., Gadahi, J. A., Xu, L., Yan, R., et al. (2017). Toxoplasma gondii
397 elongation factor 1-alpha (TgEF-1α) is a novel vaccine candidate antigen against toxoplasmosis.
398 Frontiers in microbiology 8, 168.
399 10. Lin, R.-Q., Lillehoj, H. S., Lee, S. K., Oh, S., Panebra, A., and Lillehoj, E. P. (2017).
400 Vaccination with Eimeria tenella elongation factor-1α recombinant protein induces protective
401 immunity against E. tenella and E. maxima infections. Veterinary Parasitology 243, 79–84.
402 11. Suarez, C. E., Norimine, J., Lacy, P., and McElwain, T. F. (2006). Characterization and gene
403 expression of Babesia bovis elongation factor-1α. International journal for parasitology 36, 965–
404 973.
405 12. Silva, M. G., Knowles, D. P., Mazuz, M. L., Cooke, B. M., and Suarez, C. E. (2018). Stable
406 transformation of Babesia bigemina and Babesia bovis using a single transfection plasmid.
407 Scientific reports 8, 6096.
408 13. Hakimi, H., Asada, M., and Kawazu, S. (2021). Recent advances in molecular genetic tools
409 for babesia. Veterinary Sciences 8, 222.
410 14. Vinkenoog, R., Sperança, M. A., van Breemen, O., Ramesar, J., Williamson, D. H., Ross-
411 MacDonald, P. B., et al. (1998). Malaria parasites contain two identical copies of an elongation
412 factor 1 alpha gene. Molecular and biochemical parasitology 94, 1–12.

413 15. Suarez, C. E., and McElwain, T. F. (2010). Transfection systems for Babesia bovis: a review
414 of methods for the transient and stable expression of exogenous genes. Veterinary parasitology
415 167, 205–215.
416 16. Fernandez-Becerra, C., Azevedo, M. de, Yamamoto, M. M., and Portillo, H. del (2003).
417 Plasmodium falciparum: new vector with bi-directional promoter activity to stably express trans
418 genes. Available at: https://www.cabidigitallibrary.org/doi/full/10.5555/20033124083 (Accessed
419 March 13, 2025).
420 17. Suarez, C. E., Alzan, H. F., Silva, M. G., Rathinasamy, V., Poole, W. A., and Cooke, B. M.
421 (2019). Unravelling the cellular and molecular pathogenesis of bovine babesiosis: is the sky the
422 limit? International journal for parasitology 49, 183–197.
423 18. Johnson, W. C., Hussein, H. E., Capelli-Peixoto, J., Laughery, J. M., Taus, N. S., Suarez, C.
424 E., et al. (2022). A transfected Babesia bovis parasite line expressing eGFP is able to complete the
425 full life cycle of the parasite in mammalian and tick hosts. Pathogens 11, 623.
426 19. Laughery, J. M., Knowles, D. P., Schneider, D. A., Bastos, R. G., McElwain, T. F., and Suarez,
427 C. E. (2014). Targeted surface expression of an exogenous antigen in stably transfected Babesia
428 bovis. PLoS One 9, e97890.
429 20. Oldiges, D. P., Laughery, J. M., Tagliari, N. J., Leite Filho, R. V., Davis, W. C., da Silva Vaz
430 Jr, I., et al. (2016). Transfected Babesia bovis expressing a tick GST as a live vector vaccine. PLoS
431 neglected tropical diseases 10, e0005152.
432 21. Suarez, C. E., Bishop, R. P., Alzan, H. F., Poole, W. A., and Cooke, B. M. (2017). Advances
433 in the application of genetic manipulation methods to apicomplexan parasites. International journal
434 for parasitology 47, 701–710.

435 22. Mazuz, M. L., Laughery, J. M., Lebovitz, B., Yasur-Landau, D., Rot, A., Bastos, R. G., et al.
436 (2021). Experimental infection of calves with transfected attenuated Babesia bovis expressing the
437 Rhipicephalus microplus Bm86 Antigen and eGFP Marker: Preliminary studies towards a dual
438 anti-tick/Babesia vaccine. Pathogens 10, 135.
439 23. Ozubek, S., Bastos, R. G., Alzan, H. F., Inci, A., Aktas, M., and Suarez, C. E. (2020). Bovine
440 babesiosis in Turkey: Impact, current gaps, and opportunities for intervention. Pathogens 9, 1041.
441 24. Hildebrandt, A., Zintl, A., Montero, E., Hunfeld, K.-P., and Gray, J. (2021). Human babesiosis
442 in Europe. Pathogens 10, 1165.
443 25. Cubillos, E. F., Snebergerova, P., Borsodi, S., Reichensdorferova, D., Levytska, V., Asada,
444 M., et al. (2023). Establishment of a stable transfection and gene targeting system in Babesia
445 divergens. Frontiers in Cellular and Infection Microbiology 13. Available at:
446 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10753763/ (Accessed April 30, 2024).
447 26. Cuesta, I., González, L. M., Estrada, K., Grande, R., Zaballos, Á., Lobo, C. A., et al. (2014).
448 High-Quality Draft Genome Sequence of Babesia divergens , the Etiological Agent of Cattle and
449 Human Babesiosis. Genome Announc 2, e01194-14. doi: 10.1128/genomeA.01194-14
450 27. Gonzalez, L. M., Estrada, K., Grande, R., Jiménez-Jacinto, V., Vega-Alvarado, L., Sevilla, E.,
451 et al. (2019). Comparative and functional genomics of the protozoan parasite Babesia divergens
452 highlighting the invasion and egress processes. PLoS neglected tropical diseases 13, e0007680.
453 28. Asensi, V., González, L. M., Fernández-Suárez, J., Sevilla, E., Navascués, R. Á., Suárez, M.
454 L., et al. (2018). A fatal case of Babesia divergens infection in Northwestern Spain. Ticks and tick-
455 borne diseases 9, 730–734.
456 29. Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix.
457 Molecular biology and evolution 25, 1307–1320.

458 30. Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the
459 control region of mitochondrial DNA in humans and chimpanzees. Molecular biology and
460 evolution 10, 512–526.
461 31. Quilez, J., Guilmatre, A., Garg, P., Highnam, G., Gymrek, M., Erlich, Y., et al. (2016).
462 Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA
463 methylation in humans. Nucleic acids research 44, 3750–3762.
464 32. Schnittger, L., Ganzinelli, S., Bhoora, R., Omondi, D., Nijhof, A. M., and Florin-Christensen,
465 M. (2022). The Piroplasmida Babesia, Cytauxzoon, and Theileria in farm and companion animals:
466 Species compilation, molecular phylogeny, and evolutionary insights. Parasitology research 121,
467 1207–1245.
468 33. Gorenflot, A., Brasseur, P., Precigout, E., L'Hostis, M., Marchand, A., and Schrevel, J. (1991).
469 Cytological and immunological responses toBabesia divergens in different hosts: Ox, gerbil, man.
470 Parasitol Res 77, 3–12. doi: 10.1007/BF00934377
471 35. Georges, K., Loria, G. R., Riili, S., Greco, A., Caracappa, S., Jongejan, F., et al. (2001).
472 Detection of haemoparasites in cattle by reverse line blot hybridisation with a note on the
473 distribution of ticks in Sicily. Veterinary parasitology 99, 273–286.
474 34. Mosqueda, J., Olvera-Ramírez, A., Aguilar-Tipacamú, G., and Cantó, G. (2012). Current
475 Advances in Detection and Treatment of Babesiosis. Curr Med Chem 19, 1504–1518. doi:
476 10.2174/092986712799828355
477 36. Aktas, M., and Ozubek, S. (2015). Molecular and parasitological survey of bovine piroplasms
478 in the Black Sea region, including the first report of babesiosis associated with Babesia divergens
479 in Turkey. Journal of medical entomology 52, 1344–1350.

480 37. Singh, P., Vydyam, P., Fang, T., Estrada, K., Gonzalez, L. M., Grande, R., et al. (2024).
481 Insights into the evolution, virulence and speciation of Babesia MO1 and Babesia divergens
482 through multiomics analyses. Emerging Microbes & Infections 13, 2386136. doi:
483 10.1080/22221751.2024.2386136
484 38. Wang, S., Li, D., Chen, F., Jiang, W., Luo, W., Zhu, G., et al. (2022). Establishment of a
485 transient and stable transfection system for babesia duncani using a homologous recombination
486 strategy. Frontiers in Cellular and Infection Microbiology 12, 844498.
487 39. Singh, P., Lonardi, S., Liang, Q., Vydyam, P., Khabirova, E., Fang, T., et al. (2023). Babesia
488 duncani multi-omics identifies virulence factors and drug targets. Nature microbiology 8, 845–
489 859.
490 40. Bioinformatics, B. (2018). Babraham Bioinformatics-FastQC a quality control tool for high
491 throughput sequence data.
492 41. Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina
493 sequence data. Bioinformatics 30, 2114–2120.
494 42. Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: molecular evolutionary genetics
495 analysis version 11. Molecular biology and evolution 38, 3022–3027.

496 **Acknowledgements**

497 We thank “Centro de Transfusiones de la Comunidad de Madrid”, which provided the human A+
498 blood from healthy volunteer donors. This work was funded by grants from Health Institute Carlos
499 III (ISCIII; PI20CIII-00037 to E.M. and L.G.M) and Comunidad Autónoma de Madrid (CAM;
500 TEC-2024/BIO-66/SALAINDEC-CM to EM and L.G.M.).

501 **Author Contrubitions**

502 Conceptualization: SO, ASF, EM and CES; Methodology and experiments: OS, LMG, HA, AGI,
503 RGB, MA and RG; Sequencing and bioinformatic analysis: SO, RG, LMG and ASF, Formal
504 analysis and data interpretation SO, EM, ASF and LMG; Writing original draft SO; Writing,
505 reviewing and editing ASF, EM, LMG. All authors read and approved of the final manuscript.

506 **Competing interests**

507 The authors declare no competing interests.

508 **Ethics statement**

509 *Babesia divergens* Spanish isolate were cultured using human A+ blood obtained from healthy
510 volunteer donors. The blood was sourced from the Blood Transfusion Center of Madrid, Spain,
511 adhering to approved protocols and in compliance with the relevant institutional guidelines and
512 regulations. *Babesia divergens* Turkiye isolate was obtained from a clinically diagnosed babesiosis
513 case in a cow. The sample was submitted for diagnostic purposes to the Department of
514 Parasitology, Faculty of Veterinary Medicine, Fırat University. The study was conducted in
515 accordance with institutional guidelines and regulations.

516 **Additional information**

517 **Supplementary Information**

518 **Supplementary Figure 1:** Percentage of identity of amino acid sequence of EF-1α among
519 different piroplasm species.

520 **Supplementary Figure 2:** Comparison of EF-1α protein sequences among *B. divergens* isolates
521 (R87, Spanish sample, and Turkiye isolates) and various piroplasm species. Three conserved EF-

522 1α regions with known Guanosine 5'-triphosphate (GTP)-Binding Site Function (GTP BS 1, 2, and
523 3) highlighted in boxes.

524 **Supplementary Figure 3:** Comparative alignment of the EF-1α-IG region across zoonotic
525 *Babesia* species. The TA repeat motifs (highlighted in black) show notable variability among *B.*
526 *divergens* isolates (R87, Spanish, and Turkiye), while *Babesia* MO1 contains only 3 TA motifs,
527 and *B. duncani* lacks any TA repeats.

528 **Supplementary Figure 4:** Schematic representation of the *B. divergens* EF-1α locus showing the
529 primers used for IG region amplification. IR (inverted repeat) and IG (intergenic region) are
530 highlighted, with primer positions indicated by arrows.

531

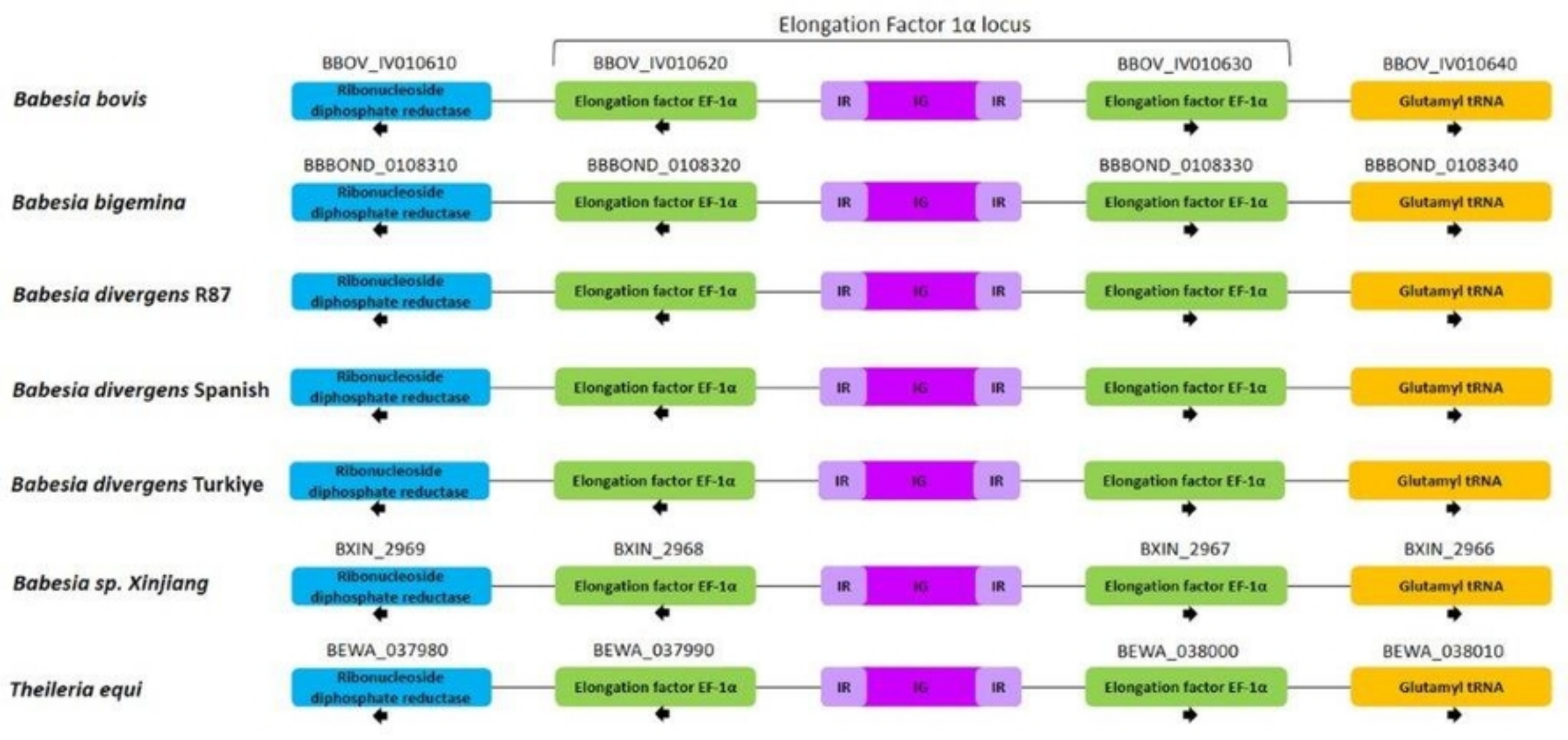

Elongation Factor 1α locus
Babesia bovis
BBOV_IV010610
Ribonucleoside diphosphate reductase
BBOV_IV010620
Elongation factor EF-1α
IR
IG
IR
BBOV_IV010630
Elongation factor EF-1α
BBOV_IV010640
Glutamyl tRNA
Babesia bigemina
BBBOND_0108310
Ribonucleoside diphosphate reductase
BBBOND_0108320
Elongation factor EF-1α
IR
IG
IR
BBBOND_0108330
Elongation factor EF-1α
BBBOND_0108340
Glutamyl tRNA
Babesia divergens R87
Ribonucleoside diphosphate reductase
Elongation factor EF-1α
IR
IG
IR
Elongation factor EF-1α
Glutamyl tRNA
Babesia divergens Spanish
Ribonucleoside diphosphate reductase
Elongation factor EF-1α
IR
IG
IR
Elongation factor EF-1α
Glutamyl tRNA
Babesia divergens Turkiye
Ribonucleoside diphosphate reductase
Elongation factor EF-1α
IR
IG
IR
Elongation factor EF-1α
Glutamyl tRNA
Babesia sp. Xinjiang
BXIN_2969
Ribonucleoside diphosphate reductase
BXIN_2968
Elongation factor EF-1α
IR
IG
IR
BXIN_2967
Elongation factor EF-1α
BXIN_2966
Glutamyl tRNA
Theileria equi
BEWA_037980
Ribonucleoside diphosphate reductase
BEWA_037990
Elongation factor EF-1α
IR
IG
IR
BEWA_038000
Elongation factor EF-1α
BEWA_038010
Glutamyl tRNA

Figure 1

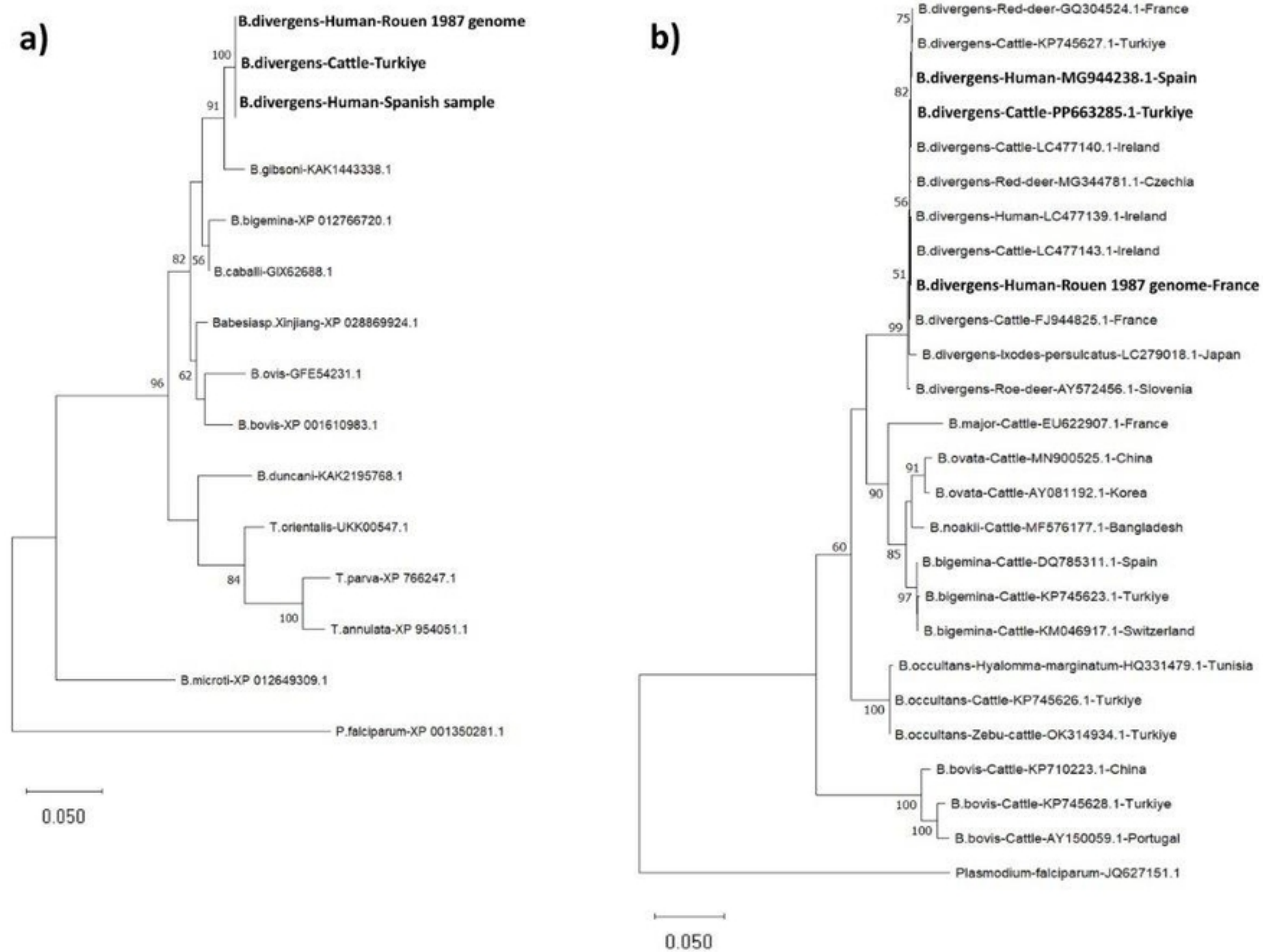

a)
B.divergens-Human-Rouen 1987 genome
B.divergens-Cattle-Turkiye
B.divergens-Human-Spanish sample
B.gibsoni-KAK1443338.1
B.bigemina-XP 012766720.1
B.caballi-GIX62688.1
Babesiasp.Xinjiang-XP 028869924.1
B.ovis-GFE54231.1
B.bovis-XP 001610983.1
B.duncani-KAK2195768.1
T.orientalis-UKK00547.1
T.parva-XP 766247.1
T.annulata-XP 954051.1
B.microti-XP 012649309.1
P.falciparum-XP 001350281.1
0.050
b)
B.divergens-Red-deer-GQ304524.1-France
B.divergens-Cattle-KP745627.1-Turkiye
B.divergens-Human-MG944238.1-Spain
B.divergens-Cattle-PP663285.1-Turkiye
B.divergens-Cattle-LC477140.1-Ireland
B.divergens-Red-deer-MG344781.1-Czechia
B.divergens-Human-LC477139.1-Ireland
B.divergens-Cattle-LC477143.1-Ireland
B.divergens-Human-Rouen 1987 genome-France
B.divergens-Cattle-FJ944825.1-France
B.divergens-Ixodes-persulcatus-LC279018.1-Japan
B.divergens-Roe-deer-AY572456.1-Slovenia
B.major-Cattle-EU622907.1-France
B.ovata-Cattle-MN900525.1-China
B.ovata-Cattle-AY081192.1-Korea
B.noakii-Cattle-MF576177.1-Bangladesh
B.bigemina-Cattle-DQ785311.1-Spain
B.bigemina-Cattle-KP745623.1-Turkiye
B.bigemina-Cattle-KM046917.1-Switzerland
B.occultans-Hyalomma-marginatum-HQ331479.1-Tunisia
B.occultans-Cattle-KP745626.1-Turkiye
B.occultans-Zebu-cattle-OK314934.1-Turkiye
B.bovis-Cattle-KP710223.1-China
B.bovis-Cattle-KP745628.1-Turkiye
B.bovis-Cattle-AY150059.1-Portugal
Plasmodium-falciparum-JQ627151.1
0.050

Figure 2

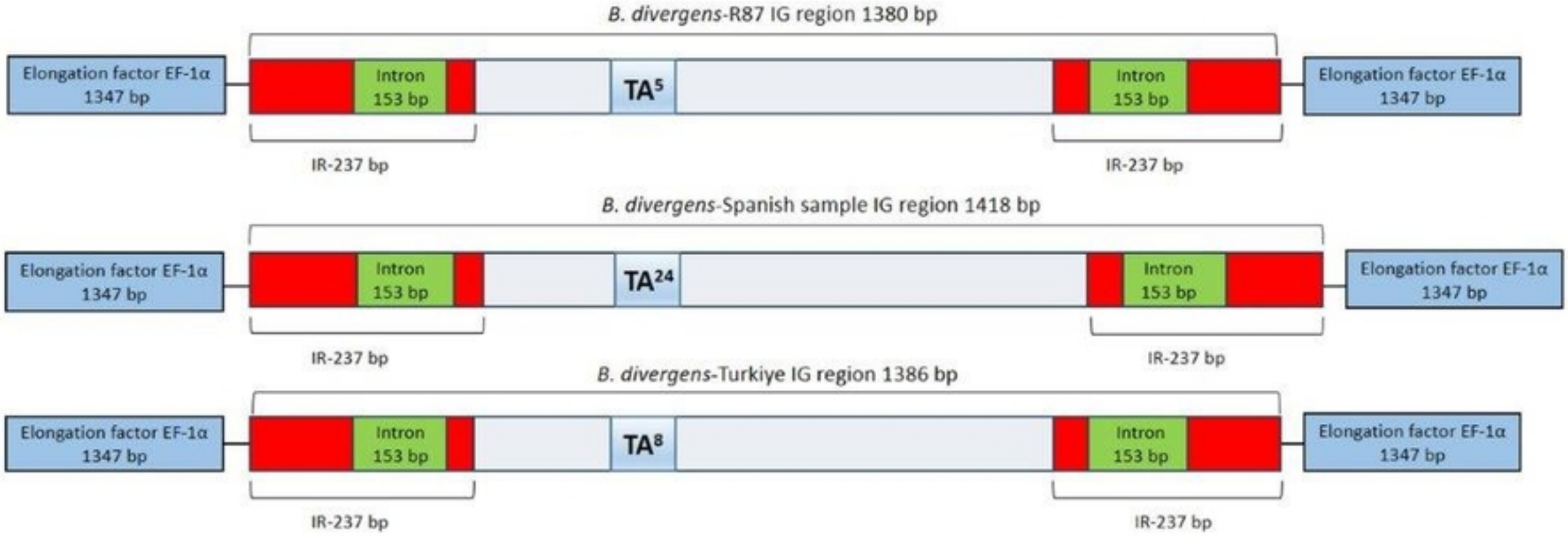

B. divergens-R87 IG region 1380 bp
Elongation factor EF-1α 1347 bp
Intron 153 bp
TA5
Intron 153 bp
Elongation factor EF-1α 1347 bp
IR-237 bp
IR-237 bp
B. divergens-Spanish sample IG region 1418 bp
Elongation factor EF-1α 1347 bp
Intron 153 bp
TA24
Intron 153 bp
Elongation factor EF-1α 1347 bp
IR-237 bp
IR-237 bp
B. divergens-Turkiye IG region 1386 bp
Elongation factor EF-1α 1347 bp
Intron 153 bp
TA8
Intron 153 bp
Elongation factor EF-1α 1347 bp
IR-237 bp
IR-237 bp

Figure 3

```
B. divergens R87          TTTGATGAGGTCTAATAATAACTAATCTGCAAAGGGTTAGTGAGGCCAATGGAAACACAA  60    ATCTCATCTGAGAACTACAGTCTATATAACTGTGCGGGTATCTTATGTTACCTTGAGTTT  742
B. divergens Spanish sample TTTGATGAGGTCTAATAATAACTAATCTGCAAAGGGTTAGTGAGGCCAATGGAAACACAA  60    ATCTCATCTGAGAACTACAGTCTATATAACTGTGCGGGTATCTTATGTTACCTTGAGTTT  780
B. divergens Turkiye      TTTGATGAGGTCTAATAATAACTAATCTGCAAAGGGTTAGTGAGGCCAATGGAAACACAA  60    ATCTCATCTGAGAACTACAGTCTATATAACTGTGCGGGTATCTTATGTTACCTTGAGTTT  748
                          ************************************************************    ************************************************************

B. divergens R87          TGGGCAACACCACAACCGAATTAACAACATGGAAATATATATTTATAATGTTATGCCCCT  120   AGAATATTTGTACGCTTATCCGTATAATTAACATCGTCGCGTTTAGTGGCTTAAGGCACC  802
B. divergens Spanish sample TGGGCAACACCACAACCGAATTAACAACATGGAAATATATATTTATAATGTTATGCCCCT  120   AGAATATTTGTACGCTTATCCGTATAATTAACATCGTCGCGTTTAGTGGCTTAAGGCACC  840
B. divergens Turkiye      TGGGCAACACCACAACCGAATTAACAACATGGAAATATATATTTATAATGTTATGCCCCT  120   AGAATATTTGTACGCTTATCCGTATAATTAACATCGTCGCGTTTAGTGGCTTAAGGCACC  808
                          ************************************************************    ************************************************************

B. divergens R87          ACGACACTTCTCATAATGACACAGCAGAACGTAGAGATCACGACACCCACACCACTAACA  180   TATATCTATGTGTACTTGCGCAAATCAAGTTGATATTTCGAACGTGATAATTGATGAGAT  862
B. divergens Spanish sample ACGACACTTCTCATAATGACACAGCAGAACGTAGAGATCACGACACCCACACCACTAACA  180   TATATCTATGTGTACTTGCGCAAATCAAGTTGATATTTCGAACGTGATAATTGATGAGAT  900
B. divergens Turkiye      ACGACACTTCTCATAATGACACAGCAGAACGTAGAGATCACGACACCCACACCACTAACA  180   TATATCTATGTGTACTTGCGCAAATCAAGTTGATATTTCGAACGTGATAATTGATGAGAT  868
                          ************************************************************    ************************************************************

B. divergens R87          TCAAAAAACCAAACAAACCAAAACAGTTTGACAAAACATACCTTTTGTTTAGCGAATGAA  240   AAGCCACATATTCAATGAAAATAATGTAAAATGCAACACACATGTGACTCAATGTAAACT  922
B. divergens Spanish sample TCAAAAAACCAAACAAACCAAAACAGTTTGACAAAACATACCTTTTGTTTAGCGAATGAA  240   AAGCCACATATTCAATGAAAATAATGTAAAATGCAACACACATGTGACTCAATGTAAACT  960
B. divergens Turkiye      TCAAAAAACCAAACAAACCAAAACAGTTTGACAAAACATACCTTTTGTTTAGCGAATGAA  240   AAGCCACATATTCAATGAAAATAATGTAAAATGCAACACACATGTGACTCAATGTAAACT  928
                          ************************************************** *******    ************************************************************

B. divergens R87          TTTTTAGGCACGTCAGTAGTGTAAACCCTAAGTTGAAAATAGAGTGTTCCTACATATAGA  300   ATGCATAAATTAATGGCACATTGTATTGCACATAGGCAAAAATAGTATTAATAACTGACA  982
B. divergens Spanish sample TTTTTAGGCACGTCAGTAGTGTAAACCCTAAGTTGAAAATAGAGTGTTCCTACATATAGA  300   ATGCATAAATTAATGGCACATTGTATTGCACATAGGCAAAAATAGTATTAATAACTGACA  1020
B. divergens Turkiye      TTTTTAGGCACGTCAGTAGTGTAAACCCTAAGTTGAAAATAGAGTGTTCCTACATATAGA  300   ATGCATAAATTAATGGCACATTGTATTGCACATAGGCAAAAATAGTATTAATAACTGACA  988
                          ****************************** ***************************    ************************************** ******************

B. divergens R87          GCACTAATATATATGGGAGGATAATACTGTGGGGTTTCGTGGAAGCCACGCTAGACAATA  360   TAATTAATGGTTTATTGTCTAATGTGTTGGAACCAAATTATGAGGCATTGTCCATTCCTA  1042
B. divergens Spanish sample GCACTAATATATATGGGAGGATAATACTGTGGGGTTTCGTGGAAGCCACGCTAGACAATA  360   TAATTAATGGTTTATTGTCTAATGTGTTGGAACCAAATTATGAGGCATTGTCCATTCCTA  1080
B. divergens Turkiye      GCACTAATATATATGGGAGGATAATACTGTGGGGTTTCGTGGAAGCCACGCTAGACAATA  360   TAATTAATGGTTTATTGTCTAATGTGTTGGAACCAAATTATGAGGCATTGTCCATTCCTA  1048
                          ************************************************************    ************************************************************

B. divergens R87          AGTTAGCGAAACGGTTGAAGATATTATGGTCTGAAGCATACTCATGCACATCGTCTCATG  420   GCCGCACCACAATTGTTAGATTGAGAGTGACTACTAATATTGTCATCTGTTTTAAGCACA  1102
B. divergens Spanish sample AGTTAGCGAAACGGTTGAAGATATTATGGTCTGAAGCATACTCATGCACATCGTCTCATG  420   GCCGCACCACAATTGTTAGATTGAGAGTGACTACTAATATTGTCATCTGTTTTAAGCACA  1140
B. divergens Turkiye      AGTTAGCGAAACGGTTGAAGATATTATGGTCTGAAGCATACTCATGCACATCGTCTCATG  420   GCCGCACCACAATTGTTAGATTGAGAGTGACTACTAATATTGTCATCTGTTTTAAGCACA  1108
                          ************************************************************    ************************************************************

B. divergens R87          TTACAGCAATTTTAATTAATTCGGTAGCTTGTGAATATGTTGTTATGCTGGCATCATATA  480   CGTATAGCGATTAACTATACGAGTTGTATTTCTTCCGTTATATTCTCTAAACAAAAGGTA  1162
B. divergens Spanish sample TTACAGCAATTTTAATTAATTCGGTAGCTTGTGAATATGTTGTTATGCTGGCATCATATA  480   CGTATAGCGATTAACTATACGAGTTGTATTTCTTCCGTTATATTCTCTAAACAAAAGGTA  1200
B. divergens Turkiye      TTACAGCAATTTTAATTAATTCGGTAGCTTGTGAATATGTTGTTATGCTGGCATCATATA  480   CGTATAGCGATTAACTATACGAGTTGTATTTCTTCCGTTATATTCTCTAAACAAAAGGTA  1168
                          ************************************************************    ************************************************************

B. divergens R87          AACTGCTATTATTTATCTATATCTATCTTTC-------------------------------  513   TGTTTTGTCAAACTGTTTTGGTTTGTTTGGTTTTTTGATGTTAGTGGTGTGGGTGTCGTG  1222
B. divergens Spanish sample AACTGCTATTATTTATCTATATCTATCTTTCTATATATATATATATATATATATATATAT  540   TGTTTTGTCAAACTGTTTTGGTTTGTTTGGTTTTTTGATGTTAGTGGTGTGGGTGTCGTG  1260
B. divergens Turkiye      AACTGCTATTATTTATCTATATCTATCTTTC-------------------------------  516   TGTTTTGTCAAACTGTTTTGGTTTGTTTGGTTTTTTGATGTTAGTGGTGTGGGTGTCGTG  1228
                          ******************************                                  ************************************************************

B. divergens R87          ---------TATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT  562   ATCTCTACGTTCTGCTGTGTCATTATGAGAAGTGTCGTAGGGGCATAACATTATAAATAT  1282
B. divergens Spanish sample ATATATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT  600   ATCTCTACGTTCTGCTGTGTCATTATGAGAAGTGTCGTAGGGGCATAACATTATAAATAT  1320
B. divergens Turkiye      ---TATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT  568   ATCTCTACGTTCTGCTGTGTCATTATGAGAAGTGTCGTAGGGGCATAACATTATAAATAT  1288
                                             ***********************************************    ************************************************************

B. divergens R87          ACATTGATGTTATCTACAAAGCAGCTGAGGGTATTGTATCACTGTGGCTTATGGCATGAT  622   ATATTTCCATGTTGTTAATTCGGTTGTGGTGTTGCCCATTGTGTTTCCATTGGCCTCACT  1342
B. divergens Spanish sample ACATTGATGTTATCTACAAAGCAGCTGAGGGTATTGTATCACTGTGGCTTATGGCATGAT  660   ATATTTCCATGTTGTTAATTCGGTTGTGGTGTTGCCCATTGTGTTTCCATTGGCCTCACT  1380
B. divergens Turkiye      ACATTGATGTTATCTACAAAGCAGCTGAGGGTATTGTATCACTGTGGCTTATGGCATGAT  628   ATATTTCCATGTTGTTAATTCGGTTGTGGTGTTGCCCATTGTGTTTCCATTGGCCTCACT  1348
                          ************************************************************    ************************************************************

B. divergens R87          CATTCAATTCTATCACAGTGTGGGCGTTTTGCTCTAGGTCGAGTTACATGTTCTAATTTT  682   AACCCTTTGCAGATTAGTTATTATTAGACCTCATCAAA  1380
B. divergens Spanish sample CATTCAATTCTATCACAGTGTGGGCGTTTTGCTCTAGGTCGAGTTACATGTTCTAATTTT  720   AACCCTTTGCAGATTAGTTATTATTAGACCTCATCAAA  1418
B. divergens Turkiye      CATTCAATTCTATCACAGTGTGGGCGTTTTGCTCTAGGTCGAGTTACATGTTCTAATTTT  688   AACCCTTTGCAGATTAGTTATTATTAGACCTCATCAAA  1386
                          ************************************************************    ************************************
```

Figure 4

```
B. divergens R87 Nanopore          AACTGCTATTATTTATCTATATCTATCTTTC------------------------------------TATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   562
B. divergens R87 Sanger            AACTGCTATTATTTATCTATATCTATCTTTC------------------------------------TATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   562
B. divergens Spanish sample Nanopore AACTGCTATTATTTATCTATATCTATCTTTCTATATATATATATATATATATATATATATATATATATATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   600
B. divergens Spanish sample Sanger AACTGCTATTATTTATCTATATCTATCTTTCTATATATATATATATATATATATATATATATATATATATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   600
B. divergens Turkiye Nanapore      AACTGCTATTATTTATCTATATCTATCTTTC------------------------------TATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   568
B. divergens Turkiye NGS           AACTGCTATTATTTATCTATATCTATCTTTC------------------------------TATATATATATATATATGAAATTATCAGCTTATACCGGAATTATGTATTAGAAATAT   568
                                   ******************************                                 **************************************************
```

Figure 5