



# Differences in *Entamoeba histolytica* Cysteine Proteinase 5 Gene Isolated From Bandar Abbas and Tabriz, Iran

Sima Rostami<sup>1,2\*</sup>, Mostafa Rezaeian<sup>3</sup>, Rasool Jamali<sup>4</sup>, Sassan Rezaie<sup>3</sup>, Zahra Babaei<sup>5</sup>, Hossein Hooshyar<sup>6</sup>

<sup>1</sup>Department of Research and Technology, Alborz University of Medical Sciences, Karaj, Iran

<sup>2</sup>Department of Parasitology, School of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran

<sup>3</sup>Division of Molecular Biology, Department of Medical Mycology & Parasitology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

<sup>4</sup>Department of Parasitology, School of Medicine, Tabriz University of Medical Sciences, Iran

<sup>5</sup>Department of Parasitology, School of Medicine, Kerman University of Medical Sciences, Kerman, Iran

<sup>6</sup>Department of Parasitology, School of Medicine, Kashan University of Medical Sciences, Kashan, Iran

## \*Corresponding Author:

Sima Rostami, Ph.D;

Email:

srostamy1382@gmail.com

## Abstract

**Background:** Amebiasis with up to 100000 human deaths each year is the third cause of human deadly parasitic disease. With regard to the fact that cysteine protease 5 is known to be one of the most important pathogenicity factors of the *Entamoeba histolytica* and also, *CP5* gene has been observed only in *E. histolytica*, hence we discriminated *E. histolytica* from *E. dispar* on *CP5* gene by polymerase chain reaction (PCR) and characterized *CP5* gene variation in *E. histolytica* isolated from patients in both cold regions and tropical regions of Iran at molecular level.

**Materials and Methods:** In the present study, a total of 2332 stool samples (1550 from Tabriz and 782 from Bandar Abbas) were studied microscopically. DNA extraction and PCR method were performed on the positive specimens, infected with *E. histolytica/E. dispar*. Finally we characterized *CP5* gene in *E. histolytica* isolates from 10 positive samples in the cold regions (Tabriz) and 10 positive samples in the tropical regions (Bandar Abbas) by sequencing and studied the polymorphism of the gene.

**Results:** Of 1550 subjects studied from Tabriz and 782 from Bandar Abbas, 83/1550 (8.3%) and 65/782 (5.35%) persons were infected with *E. histolytica/E. dispar*, respectively. The molecular results on 20 *E. histolytica* PCR positive isolates from both regions revealed that nucleotides substitution and polymorphism on *CP5* gene was more in samples from Bandar Abbas than those from Tabriz.

**Conclusion:** Prevalence of amebiasis was high in the tropical region (Bandar Abbas) compared with the cold region (Tabriz). In this study, *CP5* gene variation in the pathogenicity and virulence of this parasite in the tropical region was higher than that in the cold region.

Published Online May 28, 2016

**Keywords:** Amebiasis, *Entamoeba dispar*, *Entamoeba histolytica*, Nucleic acid Sequencing, Cysteine proteinase



Received April 2, 2017; Revised April 28, 2017; Accepted May 20, 2016

## Introduction

*Entamoeba histolytica* is a known eukaryotic parasite of human large intestine. It is the cause of various infectious diseases ranging from dysentery to abscess of liver or other organs.<sup>1,2</sup> The World Health Organization (WHO) has reported that *E. histolytica* causes approximately 50 million cases of diseases and 100000 deaths annually worldwide.<sup>3</sup> Amebiasis is the third most common cause of death due to parasitic infection after malaria and schistosomiasis, as estimated by the WHO.<sup>1</sup> The amebiasis infection seems to be more common in tropical and subtropical areas, or at least symptomatic cases are revealable. Furthermore, only 10% of *E. histolytica* strains can be invasive; consequently individuals who show clinical manifestation, develop severe disease.<sup>2</sup>

One of the most important factors which affects the pathogenicity and virulence of *E. histolytica* is cysteine protease (CP); therefore, study of the gene encoding cysteine proteinase in *E. histolytica* is very important.<sup>4,5</sup> So far several CPs with molecular weights of 19-96 kDa have been characterized.<sup>3</sup> These molecules have critical functions such as digestion of ingested bacteria or red blood cell (RBC) cells, activation of other enzymes, and degradation of extracellular matrix of intestinal tissue; so their shortage can impact the amoeba virulence.<sup>6</sup> *E. histolytica* trophozoite secretes 10 to 1000 fold higher levels of CPs than non-invasive *E. dispar*.<sup>7</sup> There are at least 6 genes encoding CPs in *E. histolytica* (*EHCP1* – *EHCP6*), but more than 90% of them belong to *EHCP1*, *EHCP2*, and *EHCP5* and expression of 2 genes (*EHCP1*

and *EHCP5*) are missing in *E. dispar*.<sup>8</sup> In this study, we attempted to explore genetic encoding differences of *CP5* of *E. histolytica* in its various infectious properties in East Azerbaijan as opposed to those of tropical areas in relevance to the possibility of genetic polymorphism of this parasite in these 2 geographic areas.

### Materials and Methods

Subjects were selected from infected persons in 2 geographic areas, Tabriz for cold regions and Bandar Abbas for tropical regions, during 2005-2006. This study was performed in 3 distinct stages: collection of samples, culture of samples, molecular preparation and evaluation.

1. Collection of samples: A total of 2332 stool samples (1550 from Tabriz and 782 from Bandar Abbas) were randomly collected from patients referred to local hospital or laboratory of public health center located in East Azerbaijan and Bandar Abbas. The direct examination and formalin ether concentration method were used for detection of *E. histolytica*/*E. dispar* trophozoite or cyst, respectively.

2. Culture and preservation: Coagulated horse serum medium was used to transform cysts to trophozoites for culturing positive samples. Then Robinson culture medium was used for mass culture and the adaptation of trophozoites.<sup>9,10</sup>

3. Preparation and molecular study: After 3-4 subcultures, the upper layer of Robinson medium was removed and the deposit was kept in centrifuge tube. Then 10 mL of PBS solution with pH:7.2 was added to the tube and mixed adequately twice at 1600 rpm for 5 minutes using centrifugation. The upper layer was removed and deposit was mixed with 10 mL of PBS suspension and centrifuged again. The sediment was then suspended in 1 mL of PBS and finally divided equally in 1.5 mL Eppendorf tubes and kept at -80°C until DNA extraction.

### Method of DNA Extraction From Trophozoites

DNA extraction was performed according to the phenol-chloroform method described by Sambrook et al with a slight modification.<sup>11</sup>

### DNA Extraction From the Cyst

DNA extraction from the cyst was carried out using the QIAamp stool Mini Kit.

### Design of Primers

Since there is not *CP5* gene in *E. dispar*, specific primers from highly conserved regions of *CP5* gene was designed by investigation of GenBank. The primer could differentiate *E. histolytica* from *E. dispar* through the observation of a 950-bp band only for *E. histolytica*.

### Target DNA Amplification by Polymerase Chain Reaction

Polymerase chain reaction (PCR) amplification of the 950-bp fragment of the *CP5* gene was performed on DNA from the positive clinical samples. Amplification of the *CP5* gene

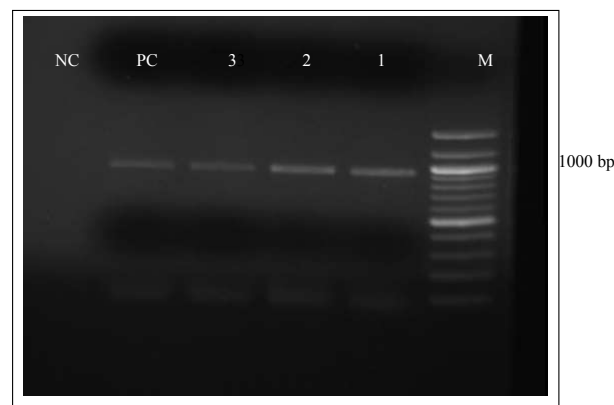
was performed as a single PCR with specifically designed primers as follows: 5' GTT CACTGTCTCGTTATTAG 3' as forward and 5' CATCAGCAACCCCAACTG 3' as reverse. The primers were tested by standard *E. histolytica* (HM1 strain) and *E. dispar* (AS161R) DNA. The PCR reaction mixture consisted of 3 µL of template DNA, 10 µL of 10X PCR buffer (Roche), 0.2 mM of each deoxynucleoside triphosphate (dNTPs), 3 µL of MgCl<sub>2</sub> 50 mM, 1 U of Taq DNA polymerase (CinnaGen, Tehran, Iran), and 20 picomol of each primer, and 30 µL of distilled water.

DNA was amplified using Primus, MWG-BIOTECH, Thermal Cycler under the following conditions: 5 minutes at 95°C as an initial hot start step, followed by 35 cycles of 1 minute at 95°C, 2 minutes at 46°C, 3 seconds at 72°C, and 10 minutes at 72°C as the final extension step. Distilled water was used as negative control. The PCR products were electrophoresized on ethidium bromide-staining 1% (W/V) agarose gel. In the case of detection of a banding pattern with approximate size of 950 bp, it was considered as *E. histolytica* and sequencing of PCR fragments were then performed. This process was done for 10 samples from each mentioned area.

### Results

Microscopic examination of 2332 stool samples (1550 from Tabriz and 782 from Bandar Abbas) showed that a total of 148 samples (6.35%) were positive for *E. histolytica*/*E. dispar* (65 specimens from Tabriz and 83 specimens from Bandar Abbas). PCR amplification of the 950 bp fragment of the *CP5* gene was performed on DNA from clinical samples diagnosed with *E. histolytica*/*E. dispar*. We characterized *CP5* gene in *E. histolytica* isolates by observation of the 950 bp band after electrophoresis of PCR product (Figure 1).

In order to sequence the isolates, 20 positive samples for *E. histolytica* (10 positive samples from each region) were selected.



**Figure 1.** Electrophoretic Separation of PCR Product From DNA Amplified at the *CP5* Locus of *Entamoeba histolytica* and *Entamoeba dispar*.

Lanes 1-3: PCR products from clinical samples; lane M: molecular marker (100 bp); PC: positive control (*E. histolytica*, HM1 strain); NC: negative control (*E. dispar*, AS161R strain).

The PCR products were sequenced on both strands and multiple alignment was produced (Figure 2). Results of the sequencing analysis for CP5-PCR, matched with GenBank data. A summary of these results is shown in Figure 2.

Multiple alignments of sequences showed the results as follows:

- One SNP (T or A at position 2202). The substitution occurred in only 2 Tabriz isolates (TAB-12, 15) while this variation was seen in most of the isolates from Bandar Abbas except for 2 isolates (BAN 25, 59).
- Two SNPs (A or T at position 2359) and (T or G at position 2360). The substitution was seen only in 4 samples of Bandar Abbas (BAN 42, 48, 59, 72).

BAN-1	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
BAN-114 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
BAN-20 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
BAN-25 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
BAN-42 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
BAN-48 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
BAN-52 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
BAN-59 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
BAN-65 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
BAN-72 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-9 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-15 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-12 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-54 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-50 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-78 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-5 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-22 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-85 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
TAB-80 70	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	129
CP5-ST 2155	TCGAATCTGGGTTGCAAAATAATAAACAATTTTACTATTGGTGA	2214

2202		
BAN-1	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
BAN-114 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
BAN-20 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
BAN-25 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
BAN-42 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
BAN-48 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
BAN-52 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
BAN-59 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
BAN-65 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
BAN-72 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-9 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-15 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-12 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-54 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-50 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-78 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-5 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-22 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-85 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
TAB-80 250	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	309
CP5-ST 2335	AACCAATTTGTCATGTAATAACAACATGAGAGGATGTTTTATGATTC	2394

2359-2360		
BAN-1	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
BAN-114 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
BAN-20 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
BAN-25 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
BAN-42 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
BAN-48 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
BAN-52 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
BAN-59 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
BAN-65 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
BAN-72 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-9 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-15 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-12 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-54 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-50 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-78 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-5 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-22 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-85 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
TAB-80 670	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	729
CP5-ST 2755	GAAGTGA AAAAGCCATTACCGTCTGCTGCTGAAGGACCACTTC	2814

2781		
BAN-1	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909
BAN-114 850	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909
BAN-20 850	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909
BAN-25 850	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909
BAN-42 850	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909
BAN-48 850	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909
BAN-52 850	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909
BAN-59 850	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909
BAN-65 850	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909
BAN-72 850	ATTGGATTGTAGA AATTAAGTGGACTATTTGGGGAGATCAAGGATATGTTTAAATG	909

Figure 2. Multialignment of Nucleotide Acid Sequences.

- Two SNPs (C or G at position 2781). The substitution occurred in only 2 Tabriz isolates (TAB-12, 15), and the variation (C or T at position 2781) was seen only in 6 Bandar Abbas isolates.
- Two SNPs (A or G at position 2955). The substitution occurred in only 6 samples of Bandar Abbas labeled as BAN 20, 42, 52, 65, 72, 114 and the variation (G or C at position 2958) was seen only in 6 Bandar Abbas isolates labeled as BAN 20, 42, 52, 59, 65, 72.

The results revealed the similarity of investigated genes with standard sample in GeneBank. It was found that 98%-100%. However, the differences in nucleic acid sequences were indicated. Nucleotides substitution and polymorphism on CP5 gene was more in samples from Bandar Abbas

BAN-1	MFTVLLLVVAYATNFWANNKHFIVE	60
BAN-114	MFTVLLLVVAYATNFWANNKHFIVE	60
BAN-20	MFTVLLLVVAYATNFWANNKHFIVE	60
BAN-25	MFTVLLLVVAYATNFWANNKHFIVE	60
BAN-42	MFTVLLLVVAYATNFWANNKHFIVE	60
BAN-48	MFTVLLLVVAYATNFWANNKHFIVE	60
BAN-52	MFTVLLLVVAYATNFWANNKHFIVE	60
BAN-59	MFTVLLLVVAYATNFWANNKHFIVE	60
BAN-65	MFTVLLLVVAYATNFWANNKHFIVE	60
BAN-72	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-9	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-15	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-12	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-54	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-50	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-78	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-5	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-22	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-85	MFTVLLLVVAYATNFWANNKHFIVE	60
TAB-80	MFTVLLLVVAYATNFWANNKHFIVE	60
CP5-ST	MFTVLLLVVAYATNFWANNKHFIVE	60

32		
BAN-1	GFFAAMTEAYNSMLKPFVDRQHEE	120
BAN-114	GFFAAMTEAYNSMLKPFVDRQHEE	120
BAN-20	GFFAAMTEAYNSMLKPFVDRQHEE	120
BAN-25	GFFAAMTEAYNSMLKPFVDRQHEE	120
BAN-42	GFFAAMTEAYNSMLKPFVDRQHEE	120
BAN-48	GFFAAMTEAYNSMLKPFVDRQHEE	120
BAN-52	GFFAAMTEAYNSMLKPFVDRQHEE	120
BAN-59	GFFAAMTEAYNSMLKPFVDRQHEE	120
BAN-65	GFFAAMTEAYNSMLKPFVDRQHEE	120
BAN-72	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-9	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-15	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-12	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-54	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-50	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-78	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-5	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-22	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-85	GFFAAMTEAYNSMLKPFVDRQHEE	120
TAB-80	GFFAAMTEAYNSMLKPFVDRQHEE	120
CP5-ST	GFFAAMTEAYNSMLKPFVDRQHEE	120

84		
BAN-1	IMQEKDYPVAAEETCTYDKKKVAVIK	240
BAN-114	IMQEKDYPVAAEETCTYDKKKVAVIK	240
BAN-20	IMQEKDYPVAAEETCTYDKKKVAVIK	240
BAN-25	IMQEKDYPVAAEETCTYDKKKVAVIK	240
BAN-42	IMQEKDYPVAAEETCTYDKKKVAVIK	240
BAN-48	IMQEKDYPVAAEETCTYDKKKVAVIK	240
BAN-52	IMQEKDYPVAAEETCTYDKKKVAVIK	240
BAN-59	IMQEKDYPVAAEETCTYDKKKVAVIK	240
BAN-65	IMQEKDYPVAAEETCTYDKKKVAVIK	240
BAN-72	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-9	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-15	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-12	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-54	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-50	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-78	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-5	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-22	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-85	IMQEKDYPVAAEETCTYDKKKVAVIK	240
TAB-80	IMQEKDYPVAAEETCTYDKKKVAVIK	240
CP5-ST	IMQEKDYPVAAEETCTYDKKKVAVIK	240

231		
BAN-1	FQLYKSGIYNSKESSTQLNHGVA	300
BAN-114	FQLYKSGIYNSKESSTQLNHGVA	300
BAN-20	FQLYKSGIYNSKESSTQLNHGVA	300
BAN-25	FQLYKSGIYNSKESSTQLNHGVA	300
BAN-42	FQLYKSGIYNSKESSTQLNHGVA	300
BAN-48	FQLYKSGIYNSKESSTQLNHGVA	300
BAN-52	FQLYKSGIYNSKESSTQLNHGVA	300
BAN-59	FQLYKSGIYNSKESSTQLNHGVA	300
BAN-65	FQLYKSGIYNSKESSTQLNHGVA	300
BAN-72	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-9	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-15	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-12	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-54	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-50	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-78	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-5	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-22	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-85	FQLYKSGIYNSKESSTQLNHGVA	300
TAB-80	FQLYKSGIYNSKESSTQLNHGVA	300
CP5-ST	FQLYKSGIYNSKESSTQLNHGVA	300

Figure 3. Multialignment of Several Amino Acid Sequences.

compared to those from Tabriz.

The nucleotide and deduced amino acid sequences of the *CP5* gene of 2 isolates of *E. histolytica* from the south (Bandar Abbas) and the northwest (Tabriz) of Iran was submitted to the GenBank (NCBI) and would be released for public access under the numbers: DQ899178, DQ899179.

Protein analysis was also performed and the result revealed some substitutions as follows (Figure 3):

1. The amino acid substitution was shown at position 32. Tryptophan was replaced with cysteine in 6 Bandar Abbas isolates ( BAN 20, 42, 48, 52, 65 and 72) while in only 2 Tabriz isolates ( TAB12, 15).
2. At position 84, lysine was replaced with histidine in only 6 Bandar Abbas isolates (BAN 42, 48, 52, 59, 72 and 114).
3. At position 232, valine was replaced with alanine in only 6 Bandar Abbas isolates (BAN 48, 52, 59, 65, 72 and 114).
4. At position 233, cysteine was replaced with tryptophan in only 7 Bandar Abbas isolates (BAN 20, 42, 52, 59, 65, 72 and 114).

## Discussion

Since most of the published epidemiological data were based on the methods that cannot discriminate *E. histolytica* from *E. dispar*, a noninvasive intestinal amoeba, the true prevalence of infection by *E. histolytica* is unknown exactly. Although *E. dispar* morphologically resembles to *E. histolytica*, its species not clinically aggressive, differentiated diagnosis of these 2 would save considerable expenses of treating these cases and diminish cost of unnecessary prescript medication along with their side effects.<sup>12,13</sup> Furthermore, only 10% of *E. histolytica* strains can be invasive; consequently individuals who are infected with them develop symptomatic disease.<sup>2</sup>

Assays such as isoenzyme,<sup>14</sup> detection of specific antigens by monoclonal antibodies,<sup>15</sup> and PCR based assays<sup>9,16</sup> have shown to be successful in identifying and differentiating these amoebas.

One of the significant factors as a best target is infectious elements as well as differentiating factors between pathogenic and non-pathogenic *E. histolytica*.<sup>3</sup>

Cysteine proteinase is the most important pathogenicity factor of *E. histolytica*. This enzyme has been recognized in many organisms including parasites, bacteria, plants as well as vertebrate and invertebrate animals.<sup>4</sup> Que and Reed reported that extracellular release of cysteine proteinase in *E. histolytica* is 10-1000 fold more than that in noninvasive *E. dispar*.<sup>3</sup> Among several documented genes which encode cysteine proteinases in both species, only *EHCP1* and *EHCP5* are specific for *E. histolytica*.<sup>17</sup> *EHCP5* is also the most important pathogenicity factor in *E. histolytica*.<sup>18</sup>

To the best of our knowledge, there was no study on the genetic polymorphism of *CP5* gene and its sequences in different areas of Iran; therefore, the present study was developed to distinguish *E. histolytica* from *E. dispar* by

single PCR using specific primer of *CP5* gene. It was also studied the polymorphism of *CP5* gene on *E. histolytica* isolated from 2 regions, Tabriz and Bandar Abbas.

In this study, amplification of the *CP5* gene was performed as a single PCR and the 950 bp expected size fragment was amplified only in *E. histolytica*. The designed primers successfully allowed the discrimination of *E. dispar* from *E. histolytica*. In addition, the lack of *CP5* gene in *E. dispar* was revealed which is in agreement with previous findings.<sup>17</sup>

Hooshyar et al used PCR-RFLP method in order to differentiate *E. histolytica* from *E. dispar* in Iranian specimens. They used restriction enzymes for discrimination of 2 species on 101 isolates. They showed that 92.1% were *E. dispar* and only 7.9% were *E. dispar* and/or mixed infections.<sup>9</sup>

In this study, among 2332 collected samples, 148 were distinguished as *E. histolytica/dispar*. Present research in comparison with previous studies<sup>12,13</sup> clearly showed that the prevalence of *E. histolytica/dispar* in these geographic areas reduced, despite the fact that the prevalence is still higher in warmer climates. Decline of percentage of infected cases in comparison to past reported incidence is because of recent advances in diagnosis, increased education, and people awareness, and have been linked with efforts to improve sanitation and general health especially in rural communities.

Moreover, the amplification of *CP5* gene and sequence analysis of 10 samples selected from Tabriz and 10 samples from the tropical region of Bandar Abbas were performed. Sequence analysis at the *CP5* locus and comparison of these isolates with reference sequences in GeneBank showed variation in 7 SNPs and nucleotides substitution.

Protein analysis revealed some important substitutions as follows: the amino acid substitution in position 32 was shown and cysteine was replaced with tryptophan in 6 Bandar Abbas isolates, but the variation was seen in only 2 Tabriz isolates. It should also be noted that cysteine is fundamental for 3D structure of enzyme and consequently for its function.

The findings of our study indicated that amino acid variations in each section might change function of cysteine proteinase enzyme. As mentioned in the results, in position 232 there was valine instead of alanine (data not shown). This variation occurred in catalytic domain of *CP5* gene of 6 isolates from Bandar Abbas and 2 isolates from Tabriz. It is possible that this type of variation may help the parasite to escape from immune system or resist to the drug.

The most variations were seen in polypeptides encoding *CP5* of *E. histolytica* isolates from tropical region (Bandar Abbas). It seems that *E. histolytica* infection depends on the factors including protozoa habitat in human, age, sanitation level, and economic and social circumstances. Besides, various weather conditions in the tropical climate have lead to some variation in nucleotide and amino acid sequences and as a consequence, pathogenetic features have differed. This opinion is in agreement with the

results of some studies that reported high prevalence of amebiasis in the south of Iran.<sup>19</sup>

On the basis of our results, the genetic polymorphism in *CP5* gene from Bandar Abbas isolates is more than that from Tabriz isolates.

It seems this can be due to the effect of temperature on *E. histolytica* and its genes. As a result, this variation may increase its aggressive potential and virulence in this area.

#### Authors' Contributions

This article was extracted from the MSc project of SR where MR and RJ supervised the project; SR was advisor of the project; ZB and HH participated in collecting of samples and sample preparation procedure.

#### Ethical Approval

This study was approved by Ethics Committee of Tabriz University of Medical Sciences.

#### Conflict of Interest Disclosures

The authors declare that they have no conflict of interests.

#### Financial Support

This study was supported by Research Committee of Tabriz University of Medical Sciences.

#### Acknowledgments

We would like to thank Dr Niyati, Dr Noorbakhsh, Dr Rokni for their advice and assistance throughout this study. The authors gratefully acknowledge Miss Farnia, Mrs Tarighi, Mr Safari, and Mr Saedi for their help and thank other people who helped. Moreover, Research Deputy of Tabriz University of Medical Science is acknowledged for providing financial support for this project.

#### References

1. Tanyuksel M, Petri W. Laboratory diagnosis of amebiasis. *Clin Microbiol Rev.* 2003;16(4):713-729.
2. Haque R, Huston CD, Hughes M, Houpt E, Petri WA. Amebiasis. *N Engl J Med.* 2003;348:1565-1573. doi:10.1056/NEJMra022710.
3. Que X, Reed SL. Cysteine proteinases and the pathogenesis of amebiasis. *Clin Microbiol Rev.* 2000;13:196-206.
4. Bruchhaus I, Loftus B, Hall N, Tannich E. The intestinal protozoan parasite *Entamoeba histolytica* contains 20 cysteine proteins genes, of which only a small subset is expressed during in vitro cultivation. *Eukaryot Cell.* 2003; 2(30): 501-509. doi:10.1128/EC.2.3.501-509.2003.
5. Willhoeft U, Hamann L, Tannich E. A DNA sequence corresponding to the gene encoding cysteine proteinase 5 in *Entamoeba histolytica* is present and positionally conserved but highly degenerated in *Entamoeba dispar*. *Infect Immun.* 1999; 67:5925-5929.
6. Stanley S Jr, Reed S. Microbes and microbial toxins: paradigms for microbial-mucosal interactions ,VI. *Entamoeba histolytica*: parasite- host interactions. *Am J Physiol Gastrointest Liver Physiol.* 2001;280(6):G1049-G1054.
7. Reed SL, Keene WE, McKerrow JH. Thiol proteinase expression correlates with pathogenicity of *Entamoeba histolytica*. *J Clin Microbiol* 1989;27:2772-2777.
8. Bruchhaus I, Jacobs T, Leippe M, Tannich E. *Entamoeba histolytica* and *Entamoeba dispar*: differences in numbers and expression of a cysteine proteinases genes. *Mol Microbiol.* 1996;22:255-263.
9. Hooshyar H, Rezaian M, Kazemi B, Jeddi-tehrani M, Solaymani-ohammadi S. The distribution of *Entamoeba histolytica* and *Entamoeba dispar* in northern, central and southern Iran. *Parasitol Res.* 2004;94(2):96-100. doi:10.1007/s00436-004-1175-8.
10. Robinson GL. The laboratory diagnosis of parasitic amoeba. *Trans R Soc Trop Med Hyg.* 1968;62(2):285-294.
11. Sambrook J, Russell DW. *Molecular Cloning: A Laboratory Manual.* 3rd ed. New York: CSHL Press; 2001.
12. Sheiban M, Rezaeian M. Study of intestinal protozoa in 7 villages in Bandar Abbas area. *Iran J Public Health.* 1981;10:45-55
13. Mazloomi Gavvani E. Study of intestinal infections in rural area in Tabriz province [thesis]. Tehran: Tarbiat Moddares University; 1987. [Persian].
14. Sargeant PG. The differentiation of invasive and noninvasive *Entamoeba histolytica* by isoenzyme electrophoresis. *Trans R Soc Trop Med Hyg.* 1978;72(5):519-521.
15. Haque R. Detection of *Entamoeba histolytica* and *Entamoeba dispar* directly in stool. *Am J Trop Med Hyg.* 1991;50(5):595-596.
16. Tachibana H, Lhara S, Kobayashi S, Kaneda Y, Takeuchi T, Watanabe Y. Differences in genomic DNA sequences between pathogenic and non-pathogenic isolates of identified by polymerase chain reaction. *J Clin Microbiol.* 1991;29(10):2234-2239.
17. Willhoeft U, Bub H, Tannich E. DNA sequences corresponding to the ariel gene family of *Entamoeba histolytica* are not present in *E. dispar*. *Parasitol Res.* 1999;85:787-789.
18. Jacobs T, Bruchhaus I, Dandekar T, Tannich E, Leippe M. Isolation and molecular characterization of a surface-bound proteinase of *Entamoeba histolytica*. *Mol Microbiol.* 1998;27: 269-276.
19. Rezaeian M. *Epidemiology of Common Diseases in Iran.* 2nd ed. Tehran: Eshtiagh Publishing Co;1999:377-386. [Persian].