PCR for detection of tick-borne *Anaplasma phagocytophilum* pathogens: a review

A. Rymaszewska

Department of Genetics, University of Szczecin, Szczecin, Poland

**ABSTRACT:** Tick-borne infections such as granulocytic anaplasmosis number among emerging infectious diseases. *Anaplasma phagocytophilum* is an obligate intracellular bacterial parasite infecting the granulocytes of vertebrates. This bacterium is the aetiologic agent of HGA (human granulocytic anaplasmosis). Molecular methods allow quick and accurate detection of pathogens in ticks, humans, or animals. Monitoring of the environment for *A. phagocytophilum* involves both classical and nested PCR, since these simple methods are most efficient. As markers, parts of the 16S rRNA, *ankA*, *groESL*, *msp2*, or *msp4* genes are used for analyses. Molecular methods have enabled analysis of the genomes of pathogens, differentiation between strains and thus, in association with ecological studies, they facilitate an understanding of their biology, pathogenicity and mode of diffusion in the natural environment.

**Keywords:** *Anaplasma phagocytophilum*; detection of *Anaplasma phagocytophilum*; 16S rRNA; *ankA*; *groESL*; *msp2*; *msp4*

**Contents**

1. Introduction
2. The use of conserved genes for the detection of *A. phagocytophilum*
   2.1. *A. phagocytophilum* 16S rRNA gene analysis
   2.2. Variability of the *groESL* operon and pathogenicity of *A. phagocytophilum* strains
3. Unique genes used for detecting *A. phagocytophilum* DNA and variability studies
   3.1. Analysis of genes related to environmental adaptation of *A. phagocytophilum*
   3.2. Analysis of the *ankA* gene in *A. phagocytophilum*
4. Conclusions
5. References

**1. Introduction**

In 1987, Kary Mullis announced a new method of *in vitro* amplification of DNA. Soon after, the method was recognized as a major breakthrough in biology. Known as the Polymerase Chain Reaction (or PCR), the technique spread to virtually all the world’s biological research laboratories within a few years. Classical PCR, as well as the PCR-based techniques that were later developed, are used in a number of fields, including taxonomy, evolutionary studies, for individual identification, and – most importantly – in medical diagnosis. This simple method of nucleic acid amplification can in a test tube reveal mutations that underlie a number of genetic diseases, or can allow the detection of infection-causing microbes in samples of patient tissue.

Tick-borne infections are regarded as emerging infectious diseases. Their incidence is increasing from year to year, which is partly due to the extension of the distribution ranges of the vectors that carry the pathogens, but also due to the changing lifestyle of modern people, who spend more time engaging in outdoor activities and travel more often to distant parts of the world. In the late 1980s and early 1990s, new diagnostic methods were invented, such as tests for infectious diseases, which also included tick-borne pathogens. Modern detection techniques applied to tick-borne pathogens,
based on molecular biology, are extremely sensitive, hence particular disease units are more likely to be recognized.

Human granulocytic anaplasmosis (HGA) is the disease caused by *Anaplasma phagocytophilum*, an obligate intracellular parasitic bacterium infecting the granulocytes of vertebrates. The parasite can enter the body of a human or a farm animal mostly via a *Ixodes* or *Dermacentor* tick. HGA is a flu-like disease with nonspecific symptoms and is thus difficult to diagnose. The Centres for Disease Control and Prevention recommend that diagnoses based on serological tests should be confirmed by molecular methods, e.g., PCR (www.cdc.gov).

In addition to the detection of *Anaplasma* in the blood of patients, environmental surveys have been carried out for years in order to better recognize the vectors and reservoirs of the bacterium. Monitoring the environment for the incidence of *A. phagocytophilum* involves both classical and nested PCR, as these simple methods are very efficient. As markers, portions of the 16S rRNA, *ankA*, *groESL*, *msp2*, or *msp4* genes are used for analyses. The template comprises DNA collected from ticks (tick homogenates), or blood or pieces of other tissues collected from wild animals. Amplification is based on a one-step PCR protocol for the genes *msp2*, *ankA*, or 16S rRNA, or a two-step PCR assay, which is recommended for the *groESL* or *msp4* genes.

The genome of *A. phagocytophilum* is relatively small, consisting of a single circular chromosome (Dunning Hotopp et al., 2006). The complete genomic sequence of *A. phagocytophilum* estimated at 1,471,282 bp was submitted to GenBank in 2006 (NC007797). The established number of genes is 1,411, 1,264 of which encode proteins, 42 which encode structural RNAs and 27 of which are pseudogenes (www.ncbi.nlm.nih.gov). Sequencing the complete genome of *A. phagocytophilum* has greatly facilitated research into the diversity of these microorganisms.

2. The use of conserved genes for the detection of *A. phagocytophilum*

2.1. *A. phagocytophilum* 16S rRNA gene analysis

Ribosomal RNA genes are highly conserved among all bacterial species; therefore, they are a useful tool in molecular diagnostics, as well as in the identification of pathogen species and in phylogenetic research (Massung et al., 2002). In most bacteria rRNA genes are arranged in a single operon repeated up to 15 times per genome. The localization of genes in the operon is usually the same, i.e., 16S (*rrs*)-23S (*rrl*)-5S (*rrf*). In the genome of *A. phagocytophilum* there are only single copies of *rrf*, *rrs* and *rrl*, similarly to other *Rickettsia* species (Massung et al., 2002; Dunning Hotopp et al., 2006).

Fragments of the 16S rRNA gene are used for the detection of bacterial DNA in vectors and reservoirs, as well as in medical and veterinary diagnostics. The efficiency of the marker varies between reports and depends on the region which is amplified. One of the first articles describing tick-transmitted pathogen detection in patient blood was the study of Chen et al. (1994). The authors proposed a two-step PCR using modified universal eubacterial primers (POmod and PC5), which amplify the entire 16S rRNA gene, and internal GE9f/GE10r primers, which generate a 919 bp product. The method was characterised by a relatively low efficiency, since only 50% of samples from patients with confirmed HGA (formerly HGE, human granulocytic ehrlichiosis) gave positive results. Similar results were obtained by Bakken et al. (2002) and Comer et al. (1999). The authors suggested that PCR testing of the serum of acute-phase patients using a 16S rRNA gene fragment is not a highly sensitive method for the early diagnosis of human anaplasmosis. The GE9f/GE10r primer pair, although not providing satisfactory results, is employed in the molecular diagnosis of anaplasmosis and has been used, e.g., for the identification of *A. phagocytophilum* DNA in the first European patients in Slovenia (Petrovec et al., 1997). Multiple research units have sought to develop new primers with higher sensitivity and diagnostic applicability, as well as allowing detection of the pathogen in organisms considered to be vectors or potential reservoirs. The testing of different primer pair variants has allowed the differentiation of five strains of *A. phagocytophilum* denoted as AP-ha, AP-variant 1, AP-variant 2, AP-variant 3 and AP-variant 4, with the most common variant being AP-ha. The latter has been detected, e.g., in human patients, but also in *I. scapularis* ticks, white-footed mice, eastern chipmunks and white-tailed deer (Massung et al., 2002).

Sequence analyses have proven that the differences between variants are minute and fall within the range of 1–4 nucleotides.
In order to detect bacterial DNA, primer pairs for a single-step reaction such as the PER1/PER2 set (Goodman et al., 1996; Massung et al., 2007), or EHR521 paired with EHR747 or EHR790 (Warner and Dawson, 1996; Petrovec et al., 1999, 2003; Stefancikova et al., 2008), are sometimes utilised to amplify a smaller fragment.

2.2. Variability of the groESL operon and pathogenicity of A. phagocytophilum strains

Another conserved fragment employed in A. phagocytophilum studies is the heat shock operon groESL which codes for the expression of heat shock proteins (HSP). In most cases it usually consists of a stress-inducible promoter preceding an ORF. The operon is composed of two genes, groES, encoding a 10–20 kDa protein, and groEL, encoding a 58–65 kDa protein, separated by a non-coding sequence of variable length depending on the bacterial species. The GroEL protein, constituting a bacterial antigen, is highly similar to the eukaryotic heat-shock protein HSP60 (Sumner et al., 1997). Proteins of this group are highly conserved, hence the sequences of bacterial and human HSPs exhibit a high level of similarity.

Portions of the groESL operon are often used in studies concentrating on the detection of tick-borne pathogens in ticks, wild animals and humans. The already established two-step PCR method is more sensitive than the single-step reaction (Sumner et al., 1997; Chae et al., 2000; Petrovec et al., 2002) which allows the detection of less severe bacteraemias. The PCR assay utilises the external primers HS1/HS6 which produce fragments of 1 300–1 450 bp, depending on the bacterial species. The use of these primers enables the detection of several tick-transmitted pathogens, such as A. phagocytophilum, R. rickettsi, Ehrlichia sp., E. canis and Bartonella henselae (Sumner et al., 1997). The product of the first reaction is not analysed electrophoretically as no product or many aspecific bands may be obtained due to bacterial DNA concentration being too low in the batch material. To avoid the drawing of incorrect conclusions, the amplicon needs to be diluted 10 times and this dilution is then used as a template in a further PCR. The use of internal primers named HS43/HSV in a nested PCR encompassing a fragment which is variable between species permits identification of the bacteria. In the case of A. phagocytophilum DNA the two-step PCR produces a 1 296 bp fragment (Figure 1) (Sumner et al., 1997; Petrovec et al., 1999; Chae et al., 2000). The HS43 HSVR primer pair is designed to amplify the end fragment of groES, the intragenic spacer and approximately 75% of the length of groEL (Petrovec et al., 1999). Sometimes, primer HS45 is utilized instead of HSVR and the obtained product is shorter (for A. phagocytophilum – 480 bp). The proposed primer pairs are characterized by high sensitivity and are applied in the detection of Anaplasm DNA (Petrovec et al., 1999, 2002, 2003; Chae et al., 2000; Lotric-Furlan et al., 2001; von Loewenich et al., 2003; Shukla et al., 2007; Rymaszewski, 2008; Smrde et al., 2010; Katargina et al., 2011).

PCR and subsequent sequencing have facilitated the observation of a high degree of variability within the groESL operon, although most of the variants were of neutral character for the fitness of cells, and thus do not undergo selection. Nevertheless it seems that a single-nucleotide substitution may influence the pathogenicity of Anaplasma. Petrovec et al. (2002) suggested that a substitution within codon 242 leading to amino acid replacement may be of key importance (Figure 2). The presence of serine instead of alanine was observed in all samples of DNA isolated from patients. A. phagocytophilum strains carrying one variant of the protein may...
be pathogenic in humans, whereas other strains remain neutral. Alternatively, both strains may cause disease but the spectra of symptoms may vary depending on bacterial genotype. Support of this notion was provided by Stuen et al. (2003) who detected two different variants of bacteria in sheep presenting different symptoms by analysing the 16S rRNA and groESL genes of Anaplasma phagocytophilum.

The presence of different strains of Anaplasma phagocytophilum in ticks or wild animals considered to be reservoirs of the bacterium has been confirmed many times by independent researchers. Multiple analyses of blood DNA samples from dogs, sheep, calves, goats and ticks of the Ixodes genus, a vector for Anaplasma, have revealed the presence of two different bacterial strains (von Loewenich et al., 2003; Petrovec et al., 2003; Shukla et al., 2007; Rymaszewska, 2008; Smrdel et al., 2010; Katargina et al., 2011).

3. Unique genes used for detecting A. phagocytophilum DNA and variability studies

3.1. Analysis of genes related to environmental adaptation of A. phagocytophilum

Among the unique genes employed in the detection and identification of Anaplasma phagocytophilum the most common are msp2 and msp4 belonging to the OMP-1/MSP-2/P44 superfamily, characteristic of Anaplasmataceae (Dunning Hotopp et al., 2006).

In the genome of A. phagocytophilum three omp-1, one msp-2, two msp-2-homologs, one msp4, and 113 p44 loci were found within this superfamily.

The msp2 gene codes for an immunodominant protein MSP2 (major surface protein-2), which is detected in Anaplasma, including A. phagocytophilum (Barbet et al., 2003; Dunning Hotopp et al., 2006). The high variability of msp2, which includes intra-species variability, with consequent protein polymorphism and the generation of antigenic variations facilitate bacterial survival in diverse hosts. The correlation of different variants of the MSP2 protein with geographical localization has also been observed (Barbet et al., 2003; Dumler et al., 2003; Casey et al., 2004; Lin et al., 2004).

According to Scorpio et al. (2004), MSP2 diversity could also be an important means of adaptation to alternate niches such as small mammals or cervid reservoir neutrophils or tick cells.

Primers enabling the amplification of a large region containing the msp2/p44 gene, the upstream gene p44ESup1, and the intergenic region between them were reported by Barbet et al. (2003). However, even though the designed primers could be successfully used with all U.S. Strains, they failed to amplify the European A. phagocytophilum strain.
Thus, the authors proposed different pairs of primers for the analysis of European anaplasms (Barbet et al., 2006). The size of the amplified fragment means that it is possible to analyse the variability of the entire gene. Primers amplifying a smaller fragment of msp2, e.g., MSP3F and MSP3R, which generate a 334 bp product (Zeidner et al., 2000a; Levin et al. 2002), are utilised for the detection of A. phagocytophilum DNA in animal samples (Figure 3). The proposed pair of primers is universal and very efficient, regardless of sample type (e.g., animal blood or tick isolates) used for pathogen DNA amplification or the geographical region from which the samples originate (Zeidner et al., 2000a,b; Levin et al., 2002; Reeves et al., 2006; Skoracki et al., 2006; Toledo et al., 2009; Rymaszewska, 2010; Rymaszewska and Adamska, 2011). Therefore, they may be applied in diagnostics, in which a single-step PCR, which shortens the time of identification, may be of great importance.

### 3.2. Analysis of the ankA gene in A. phagocytophilum

With regard to other genes, such as ankA or msp4 which serve as markers for the detection of A. phagocytophilum DNA, a two-step PCR is used. Every case involves a procedure similar to that of the groESL operon. The first product, generated with the use of external primers, is not visualized electrophoretically and its 10-fold dilution is then employed as a template for the nested PCR. The method allows the amplification of even small quantities of Anaplasma DNA.

The ankA gene (approx. size 3 693–3 696 bp in U.S. strains and 3 618–3 720 bp in European strains) encodes a protein antigen of molecular weight between 150–160 kDa and which has repeated ankyrin motifs at the amino terminus. Analysis of the AnkA protein indicates a localization in the bacterial cytoplasm and not in the membrane. During infection, the protein is transferred into eukaryotic host cells, where it has been found near the nuclei of granulocytes. Thus, A. phagocytophilum has the ability to manipulate its host’s neutrophils through AnkA (Caturegli et al., 2000; Massung et al., 2000; Park et al., 2004).

Several primer pairs have been proposed for the detection of A. phagocytophilum DNA, such as LA1/LA6, which generates a 444 bp product (Caturegli et al., 2000; Walls et al., 2000), or two-step reaction primers: ANK-F1/ANK-R1 for the primary reaction and ANK-F2/ANK-R2 (product of 667 bp; Figure 4) for the nested reaction (Massung et al., 2007). Walls et al. (2000) reported that the detection of anaplasma DNA through the amplification of ankA using the LA1/LA6 primer pair provides better results than 16S rRNA gene amplification (using the GE9f/GE10r primers). According to these authors, the method serves as a good molecular assay for the diagnosis of HGA. Walls et al. (2000) estimated the specificity of the utilised primer pair to be 100%, whereas the sensitivity was estimated to be 95%, compared to 48% for 16S rRNA gene amplification.

Some authors have proposed primer sets which allow the amplification of a repeated motif with significant variability between different populations of A. phagocytophilum (Massung et al., 2000). PCR and subsequent sequencing allowed a structural comparison of this genomic region in Anaplasma originating from various geographical locations and different hosts. The results of such studies are also of great importance for determining the ecology of A. phagocytophilum.

### 4. Conclusions

PCR is widely used in many laboratories, including those which carry out medical diagnostics. It is...
a very sensitive, specific and, what is particularly important, rapid method. In microbiological diagnostics, PCR allows the identification of pathogen species, which in turn may influence the choice of therapy and its efficiency. The use of a two-step PCR assay for the detection of tick-transmitted pathogens, such as those targeting the groESL operon or the 16S rRNA gene, permits differentiation between different bacterial strains in cases of co-infection, which is sporadic for tick-borne diseases. Moreover, knowing the nucleotide sequences of unique genes, such as msp2 or ankA in A. phagocytophilum facilitates studies in medicine and ecology. At the same time, the high sensitivity of PCR amplification of genes unique for anaplasms increases the value of those genes as markers in medical and veterinary diagnostics.

PCR and sequencing have greatly enhanced the study of the genomes of pathogens, and allowed differentiation between strains and thus, in association with ecological studies, have facilitated understanding of their biology, pathogenicity and mode of diffusion in the natural environment. The knowledge gained in this way will contribute greatly to the prevention of tick-borne diseases.

5. REFERENCES


Rymaszewska A (2008): Divergence within the marker region of the groESL operon in Anaplasma phagocyto-
tophilum. European Journal of Clinical Microbiology and Infectious Diseases 27, 1025–1036.


thilum among Ixodes ricinus ticks in Germany. Journal of Clinical Microbiology 41, 5033–5040.


Received: 2011–10–28
Accepted after corrections: 2011–11–26

Corresponding Author:
Anna Rymaszewska, University of Szczecin, Department of Genetics, ul. Felczaka 3C, 71-412 Szczecin, Poland
Tel./Fax +48 91 444 14 80, E-mail: ankas@univ.szczecin.pl