How to Cite:
Hussein, M. A., & Maaroof, M. N. (2022). Isolation and characterization of Mycoplasma hominis from the semen of infertile men and detection of virulence factors genes. *International Journal of Health Sciences*, *6*(S2), 15024–15039.
https://doi.org/10.53730/ijhs.v6nS2.8981

Isolation and characterization of Mycoplasma hominis from the semen of infertile men and detection of virulence factors genes

Mohammed Ali Hussein
Department of biology, college of education for pure sciences college of medicine, Tikrit university, Iraq
Corresponding author email: mohammed79bio79@gmail.com

Mohammed Nadhir Maaroof
Department of biology, college of education for pure sciences college of medicine, Tikrit university, Iraq

***Abstract*---**The bacterial infection of seminal fluid is one of the most important factors behind infertility in many cases. The biological diversity of the *Mycoplasma* have widely been reported to be associated with the infertility grade in many cases. This study was conducted to identify and characterize the identity and the phylogenetic positioning of Fourteen isolates of *Mycoplasma* in seminal fluids. three genetic loci were included in this study *16S rRNA*, *recA*, and *Vaa* for *Mycoplasma* sequences. Each amplified locus was investigated in both bacterial infections to identify the identity of the possible bacterial infection of those patients and to assess the pattern of the genetic variation for each genetic locus in the investigated patients. Our results indicated a complete homology with the referring sequences of *Mycoplasma hominis* (GenBank acc. no. CP055150.1). Concerning the 16S rRNA fragment, no genetic variation was found in the investigated samples. It was inferred from the tree that the investigated samples were suited in the immediate vicinity to many strains of *Mycoplasma hominis* belonging to various European origins. Furthermore, the patterns of the neighbour phylogenetic distances in this tree indicated a specific biological diversity of the incorporated *Mycoplasma hominis* sequences. In addition to the ribosomal fragment, no genetic variation was also not detected in the *Vaa* fragment. The phylogenetic analyses of the *Vaa* fragment indicated obvious European origins of the investigated sample. In contrast to both 16S rRNA and *Vaa* fragments, our results showed the presence of one genetic variant (329G>A) in the *recA* fragment with silent consequences on the protein. The phylogenetic analyses of the

recA fragment indicated that this fragment has also been deposited from several European ancestors.

***Keywords*---**mycoplasma hominis, genotype, recA gene, vaa gene.

Introduction

Mycoplasma is classified in the *Mollicutes* category, which means "soft-skinned" because it lacks a tough bacterial cell wall, so it can take so many different forms that it can be difficult to identify. It has a filamentous organelle attached at the tip. Colonies show the shape of fried eggs on agar. They are notorious for being difficult to cultivate in vitro and are frequently identified as pathogens .Furthermore, they rely entirely on sterols from host cells as well as a variety of synthesized precursors (amino acids, nucleotides, and fatty acids) (1). These organisms are resistant to active cell wall antimicrobials like penicillins and cephalosporins because they lack an unusual cell wall containing peptidoglycan (2).Many differences between strains cannot be easily detected by comparative genetic sequence analysis. It can be difficult to predict the complex interactions between proteins and regulatory pathways, which means that the loss of one gene can affect the expression of other genes or the function of other proteins. , gene regulation at transcriptional and post-transcriptional levels can affect phenotype, whereby changes more subtle than gene deletions, such as polymorphisms or mutations, can lead to unexpected changes in protein function (3).The ability of *Mycoplasma hominis* to adhere to host cells is a critical step towards host colonization. The Vaa antigen is the main adhesion material for *M. hominis* and displays a clear mutational change in size as well as variability in sequence and antigen. A variable adhesion surface protein has been identified for *M. hominis*. Initially over 20 years ago as a potential adhesive for *M. hominis* (4), this protein was first recognized as a 49 kDa surface protein in the *M. hominis* strain PG21 as a multidomain transmembrane lipoprotein important for the process of The adhesion in a strain known as P50 has a molecular weight of 50 kDa and different nomenclature has been used for it (Vaa vs. P50), and the surface lipoprotein is the same as vaa (4). The adhesion-associated fat antigen Vaa often has varying molecular masses with conserved regions so p50 is a less accurate description. Variation in the sequence and size of Vaa proteins results from the allelic variants of a single copy of the *vaa gene* in *M. hominis* and that the importance of Vaa lies in the pathogenesis of Vaa. *M. hominis* infection and immune evasion (5).This lipoprotein often has varying molecular masses with conserved regions and the variation in the composition and size of Vaa proteins results from the allelic variants of a single copy of the *vaa gene* in *M. hominis*. The *recA gene* encodes for a key protein involved in DNA recombination and repair in bacteria, is important for all homologous recombination pathways and is required for the initiation of the SOS response after DNA damage (6).Studies have confirmed that *Mycoplasma* is highly specific with respect to its DNA repair mechanisms however, in *M. pulmonis*, induced mutations of the genes *recA*, *ruvA*, *ruvB* and *ruvC* produced incomplete phenotypes (6), by analysis of the genomes sequentially, the only complete repair mechanism found in *Mycoplasma* is nucleotide excision repair (NER), possibly complemented by enzymes from the BER Base Excision Repair and Recombination Mechanism .(7) identified The NER

mechanism is also in *M. genitalium*, in which the RecA protein plays an active role in the antigenic contrast mechanism but not in DNA repair (8). a histone-like protein that recognizes DNA nucleotide incorporation errors has been identified in *M. gallisepticum,* and that *Mycoplasma* has lost during its evolution the MMR (9) Mismatch Repair Mechanism. However, it is clear That *mycoplasma* possesses an appropriate set of genes and gene products, even if not yet fully identified, to promote DNA repair in such a way that its evolution is possible, causing serious disease on a large scale.

Material and method

Isolation method

50 semen samples were collected in sterile special containers, inoculated in PPLO-modified medium (MDCS) (10) and incubated anaerobically for 24-72 hours at 37°C. On the sloped surface of the solid media, colonies appeared. After being stained with Denise dye, the developing colonies were microscopically diagnosed using a dissection microscope. The carbohydrate fermentation test, the arginine group depletion test, and the urease detection test were all used to diagnose genital bacteria.A prepared DNA extraction kit (Geneaid, Taiwan) was used to extract DNA from each sample, then the samples were kept at -20°C until PCR was used for DNA extraction. The bacterial genotypes were determined based on the DNA sequence of the *16S rRNA* after amplifying the gene using the PCR method, and the *recA, and Vaa* was also detected using the PCR method, the primers used in this study were designed by the Korean company Macrogen and were based on a database site NCBI-Gene Bank for the purpose of obtaining the nucleotide sequence of the gene used in the study.Inside this study, amplified by pcr was ramped up through using additions.20µl plus 5 µl of dna templet , 1 µl (10 pmol) of each forword : 5-CAATGGCTAATGCCGGATACGC-3 and reverse : 5-GGTACCGTCAGTCTGCAAT-3primers for *16S rRNA*, and 13µl of Water containing free nuclease to the pcr tube ,pcr premix kit (bioneer , Korea) which include additional pcr response requirements(dNTPs , KCl , MgCl2 , Taq DNA polymerase ,tris HCl pH :9.0 , tracking dye and stabilizer) , The *16S rRNA gene* 334bp fragment underwent for amplification in following conditions: initial denaturation at 94 °C for 3 min. , subsequently 30 cycles of denaturation at 94 °C for 1min. , annealing at 62°C for 1min. , elongation at 72°C for 1min. after final extension at 72°C for 10 min. and holding at 4°C . Using agarose gel electrophoresis, the pcr product of the *16S rRNA gene* (334bp) was analyzed.The same materials were also added to amplify the *recA gene* and *Vaa gene* except for the primers each forward : 5-ACGCTATTGCCGAAATACAA-3 and reverse : 5-TGAAACTATATCACGAGCCCT-3 primers for *recA gene* and forward:5-CCCCGGAGATTATTAAGTCT-3 and reverse: 5-GTGCCCATTAGTAGCACTAT-3 for *Vaa gene* . Then the primers were added together to make multiple PCR for the two genes in the following way by add 20µl including 5 µl of dna templet , 1 µl (10 pmol) of each primer(*recA* and *Vaa*) , and 11µl of Free nuclease water to the pcr tube, The multiplex pcr It was completed under the following conditions: initial denaturation at 95 °C for 3 min. , followed by 30 cycles of denaturation at 95 °C for 30 sec. , annealing at 55°C for 1min. , elongation at 72°C for 1min. followed final extension at 72°C for 5 min. ,and holding at 4°C . Using agarose gel electrophoresis, the pcr product of the multiplex pcr was analyzed.

Sequencing methods

Nucleic acids sequencing of PCR amplicons

The resolved PCR amplicons were commercially sequenced from the forward direction, following to instruction manuals of the sequencing company (Macrogen Inc. Geumchen, Seoul, South Korea). Only clear chromatographs obtained from ABI (Applied Biosystem) sequence files were further analyzed, ensuring that the annotation and variations are not because of PCR or sequencing artifacts. By comparing the observed nucleic acid sequences of the investigated bacterial samples with the retrieved nucleic acid sequences, the virtual positions, and other details of the retrieved PCR fragments were identified

Interpretation of sequencing data

The sequencing results of the PCR products of the targeted samples were edited, aligned, and analyzed as long as with the respective sequences in the reference database using BioEdit Sequence Alignment Editor Software Version 7.1 (DNASTAR, Madison, WI, USA). The observed variations in each sequenced sample were numbered in PCR amplicons as well as in its corresponding position within the referring genome. The observed nucleic acids were numbered in PCR amplicons as well as in their corresponding positions within the referring genome. Each detected variant within the bacterial sequences was annotated by SnapGene Viewer ver. 4.0.4 (https://www.snapg ene.com).

Translation of nucleic acid variations into amino acid residues

The amino acid sequences of the targeted proteins were retrieved online from the protein data bank (http://www.ncbi.nlm.nih.gov). The observed nucleic acid variants in the coding portions of the analyzed genetic loci were translated into a reading frame corresponding to the referring amino acid residues in the encoded protein using the *Expasy* online program (http://web.expasy.org/translate/). Multiple amino acid sequence alignment was conducted between the referring amino acid sequences and their observed mutated counterpart using the “align” script of the BioEdit server.

Comprehensive phylogenetic tree construction

A specific comprehensive tree was constructed in this study according to the neighbour-joining protocol described by (11). The observed variants were compared with their neighbour homologous reference sequences using the NCBI-BLASTn server (12). Then, a full inclusive tree, including the observed variant, was built by the neighbour-joining method and visualized as a circular cladogram using the iTOL suit (13https://itol.embl.de/). The sequences of each incorporated species in the comprehensive tree were colored in an appropriate color to be differentiated from the other species. Whereas the sequences color of each species was taken as one unified color.

Result and Discussion

The Results of Isolating Colonies on Culture Media

The results of this study showed that *Mycoplasma hominis* bacteria were isolated from the semen of infertile men. 50 samples were collected the isolates positive for the examination were 14 isolates. The percentage of *M.hominis* bacteria was 28%. *M. hominis* appeared after 72-48 hours in the form of a fried egg as in the figure (1)

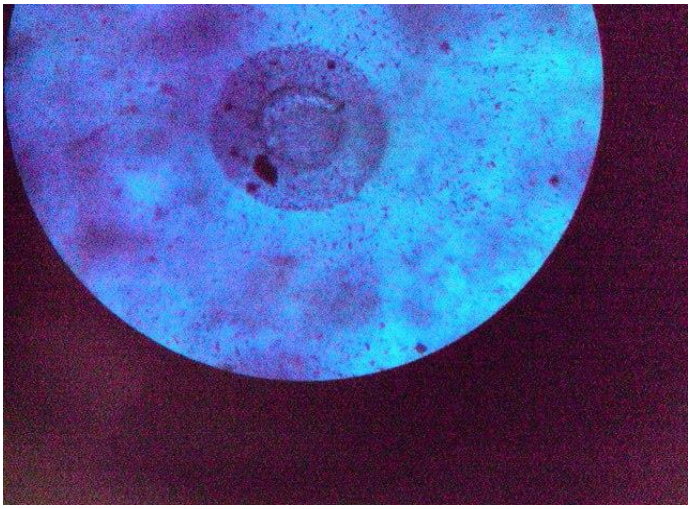

Figure (1) Under the microscope, Mycoplasma hominis. on power 40x , stained with Denise

Sequencing results

PCR Technique

The results of this technique showed the success of the process of amplifying the DNA extracted from cell of the for the *16S rRNA gene* , *vaa gene ,recA gene* and multiplex pcr for two genes together. Figure (2,3,4,5) shows the positive results of DNA samples that were amplified for *16S rRNA gene* , *vaa gene ,recA gene* and multiplex extracted from cell after being transferred onto an agarose gel.

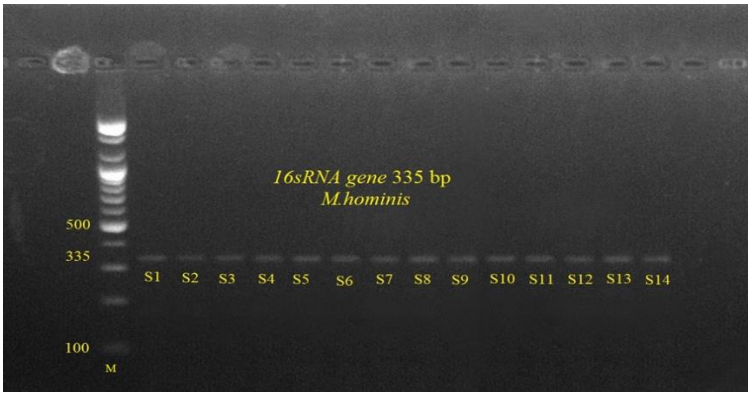

Figure (2) representing electrophoresis on agarose gel to amplify the *16S rRNA gene* extracted from the DNA

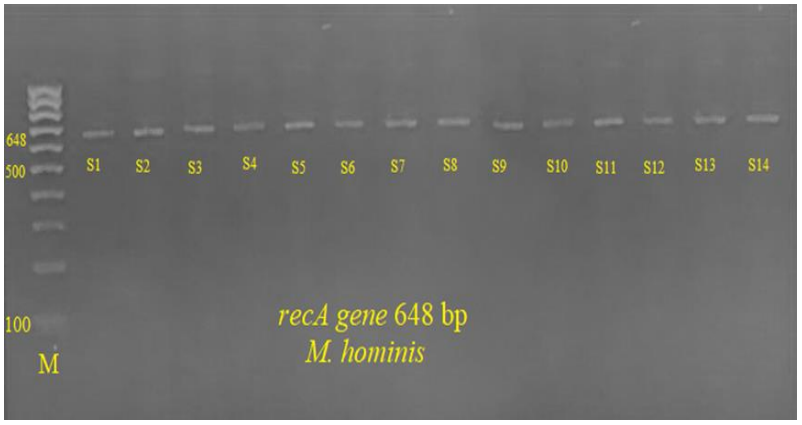

Figure (3) representing electrophoresis on agarose gel to amplify the *recA gene* extracted from the DNA

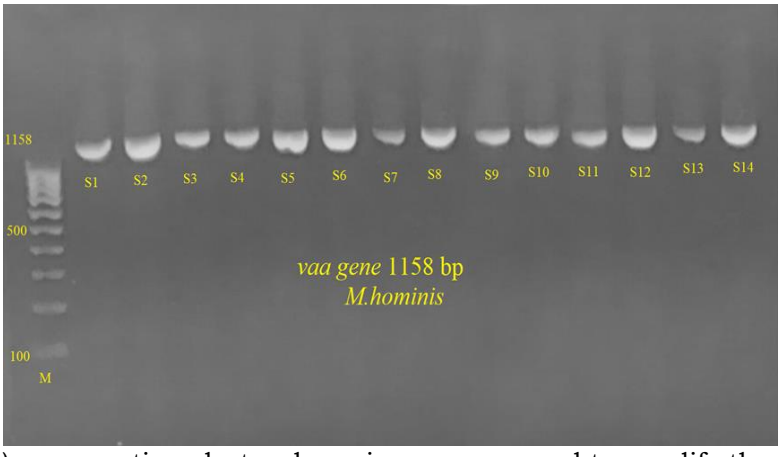

Figure (4) representing electrophoresis on agarose gel to amplify the *vaa gene* extracted from the DNA

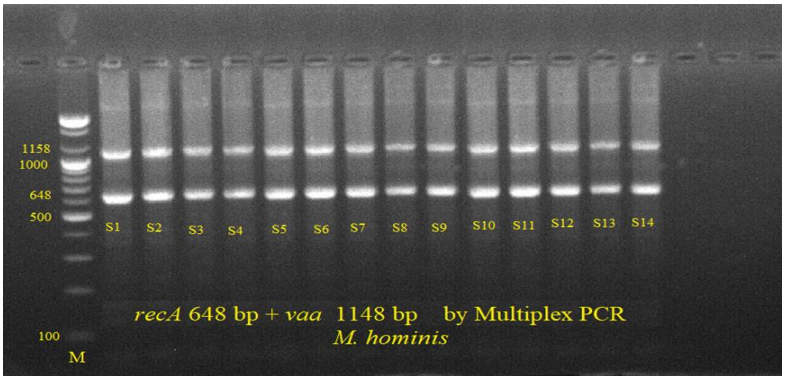

Figure (4) representing electrophoresis on agarose gel to amplify,the *recA* & *vaa* by multiplex PCR gene extracted from the DNA

16S rRNA of *Mycoplasma hominis* sequences

Within this locus, four samples were included in the present study. These samples were screened to amplify 16SrRNA sequences of *Mycoplasma*. It was well known that the variation of the 16SrRNA can be used for *Mycoplasma hominis* genotyping due to the high ribosomal ability to adapt to variable genetic diversity as it was seen in different bacterial types. The sequencing reactions indicated the exact identity after performing NCBI blastn for these PCR amplicons. Concerning the 335 bp amplicons, the NCBI BLASTn engine showed 100% sequences similarities between the sequenced samples and the reference sequences of *Mycoplasma hominis*. By comparing the observed nucleic acid sequences of these investigated samples with the retrieved nucleic acid sequences (GenBank acc. CP055150.1), the accurate positions and other details of the retrieved PCR fragments were identified. The total length of the targeted locus was localized in the NCBI server, and the positions of the start and end of the targeted locus were also confirmed within the most homologous bacterial target.

After positioning the 335 bp amplicons' sequences within the genomic sequences of the *Mycoplasma hominis*, the details of its sequences were highlighted, and the total length of the amplified amplicons was also determined .In contrast to the *ureB/ureA* amplicons, the alignment results of the 335 bp samples revealed the absence of any nucleic acid variation in the analyzed samples in comparison with the most similar referring reference nucleic acid sequences (GenBank acc. no. CP055150.1).A comprehensive phylogenetic tree was generated in the present study according to nucleic acid variations observed in the amplified 335 bp of the 16S rRNA amplicons. This phylogenetic tree was contained S1 to S4 samples alongside other relative nucleic acid sequences of *Mycoplasma* sequences. Within this tree, our investigated samples were incorporated alongside other relative sequences to constitute only one type of incorporated sequence within the cladogram, which is Mycoplasma. In contrast to the *Ureaplasma parvum* tree, no other bacterial was found to be related to *Mycoplasma* sequences. This data indicated the highly specific ability of 16S rRNA gene-based amplicons to detect these bacterial particles without including any noticeable homology with other sequences of other species within the same genus. The total number of the aligned nucleic acid sequences in this comprehensive tree was 108. In the

constructed cladogram, the investigated samples were clustered into three phylogenetic clades within the *Mycoplasma* sequences. The most interesting fact observed in our investigated bacterial isolates is correlated with the ability of the utilized 16S rRNA gene-based amplicons to categorize the *Mycoplasma* sequences into three related clades, each of which differs distinctly from the other clades within the same tree. This observation indicated that the utilized 16S rRNA gene-based amplicons provided a noticeable accuracy in the detection of these bacterial species (Fig. 2). Within the largest clade, our samples (S1 – S4) were incorporated in the clade of *Mycoplasma* hominis. This clade consists of 42 sequences of different *Mycoplasma hominis* strains. The distribution of our samples showed one position in evenly distributed distances due to the absence of any detectable nucleic acid variations. Despite the ability of 16S rRNA in the accurate description of *Mycoplasma hominis* isolates, this sort of positioning indicated no possible role for the analyzed 16S rRNA sequences in detecting any possible discrimination among them in their cladogram. Accordingly, there is no slight tilt toward any observed deviation from that indicated position of these isolates in the original clade. However, our samples were positioned in the immediate vicinity to the GenBank acc. no. of CP038014.1 and CP035542.1 that belonged to two strains of *Mycoplasma hominis* respectively deposited from France and the GenBank acc. no. of CP055150.1 that belonged to one strain of *Mycoplasma hominis* deposited from Germany. Thus, the European sources of our investigated samples are highly possible and could not be excluded from the explanation. This indicated the presence of high genetic homology among the bacterial sequences of these strains within the referred clade. However, the aggregation of these investigated bacterial samples with each other may refer to the presence of only one close pattern of the phylogenetic distribution of these sequences, which may add another layer of confirmation about the European sources of these investigated isolates.The current observation of this tree has confirmed sequencing reactions because it explained the actual neighbour-joining-based positioning in such observed variations. Interestingly, the European origins of our investigated samples could not be ignored. This was due to their positioning in the vicinity to variable European strains that belonged to the same bacterial species.

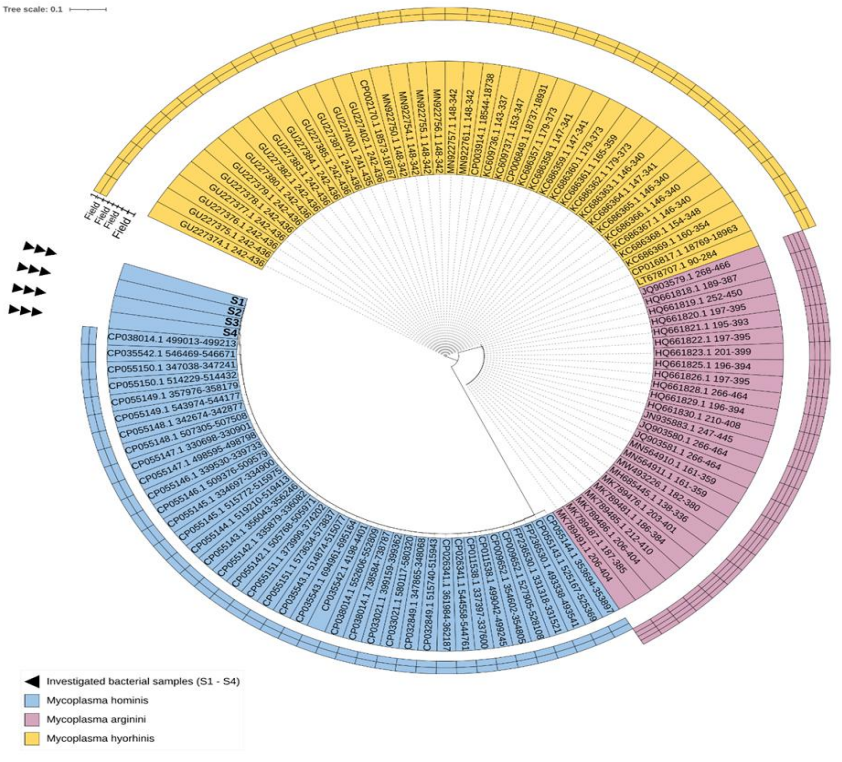

Fig. 2. The comprehensive cladogram phylogenetic tree of genetic variants of the 16S rRNA fragment of four human-infecting *Mycoplasma hominis* samples. The black-colored triangle refers to the analyzed bacterial variants. All the mentioned numbers referred to the GenBank accession number of each referring species. The number "0.1" at the top portion of the tree refers to the degree of scale range among the comprehensive tree categorized organisms. The letter "S#" refers to the code of the investigated samples.

Noteworthy, the utilization of the 16S rRNA sequences in this study has given further indication for the presence of the precise identification of the actual genotype of this bacterial organism. Interestingly, distinct phylogenetic differences were found between the clade of *Mycoplasma hominis* and the other two incorporated clades within the same tree. Despite the far genetic position of the *Mycoplasma argini* from the clade of *Mycoplasma* hominis, this tree indicated that *Mycoplasma argini* occupies the closest phylogenetic distance toward *Mycoplasma hominis* sequences. Furthermore, the clade of *Mycoplasma hyorthinis* is the most distant from the clade of *Mycoplasma hominis* sequences due to the observed low genetic homology between them. However, this 16S rRNA gene-based comprehensive tree has provided comprehensive evidence about the high competency of such genetic fragments to efficiently identify *Mycoplasma hominis* using the 16S rRNA genetic fragment. This, in turn, gives a further indication of the ability of the currently utilized 16S rRNA gene-specific primers to describe the investigated *Mycoplasma hominis* and their accurate phylogenetic positions. Irrespective of the inability of the ribosomal sequences to discriminate among

Mycoplasma hominis isolates, the high ability of such genetic fragments to efficiently identify bacterial sequences was proved by using these 16S rRNA gene-based fragments. This, in turn, gives a further indication of the ability of the currently utilized 16S rRNA gene-specific primers to describe the currently investigated *Mycoplasma hominis* and their accurate phylogenetic positions.

recA gene of Mycoplasma hominis sequences

Within this locus, only one sample (assigned *recA*) was included in the present study. This sample was screened to partially amplify the DNA *recA* gene that codes for recombinase RecA in the *Mycoplasma hominis* sequences. Thus, the variation of the DNA *recA* gene can be used for genotyping of this bacteria due to its possible ability to adapt to variable genetic diversity as it was seen in different bacterial types. The sequencing reactions indicated the exact identity after performing NCBI blastn for these PCR amplicons. Concerning the 648 bp amplicons, the NCBI BLASTn engine showed up to 99.5% sequences similarities between the sequenced samples and *Mycoplasma hominis* reference target sequences. By comparing the observed nucleic acid sequences of these investigated samples with the retrieved nucleic acid sequences (GenBank acc. CP03302.1), the accurate positions and other details of the retrieved PCR fragments were identified. In addition to the partial amplification of the *recA* locus, the currently utilized amplicons were also partially covered the *recA* locus, which is concerned with the production of recA recombinase. The total length of the targeted loci was localized in the NCBI server, and the positions of the start and end of the targeted locus were also confirmed within the most homologous bacterial target. After positioning the 648 bp amplicons' sequences within the genomic sequences of the *Mycoplasma hominis*, the details of its sequences were highlighted, and the total length of the amplified amplicons was also determined. Interestingly, the alignment results of the 648 bp samples revealed the presence of one nucleic acid variation in the analyzed sample (assigned recA sample) in comparison with the most similar referring reference nucleic acid sequences (GenBank acc. no. CP03302.1).our results indicated the presence of one nucleic acid variant observed in the investigated recA sample, namely 329G>A. This variation was observed in the currently observed nucleic acid sequences in the analyzed sample was not found in the corresponding reference sequences. To confirm this variation, the sequencing chromatograms of the investigated samples, as well as their detailed annotations, were verified and documented, and the chromatograms of their sequences were shown according to their positions in the PCR amplicons . The presence of this variant was confirmed in its original chromatogram and the absence of any possible technical error was also confirmed.The investigated nucleic acid sequences were converted to their corresponding positions in the recA recombinase. All nucleic acid sequences of the investigated samples were translated to their corresponding amino acid sequences using the Expasy translate suite. Amino acid alignment of these amino acid sequences with their references showed that the investigated 648 bp consisted of 215 amino acid sequences in the entire amino acid sequences in the recA recombinase . The observed nucleic acid variations were further analyzed to identify whether such substitution induces a possible alteration in their corresponding positions in the entire protein. Due to the observed nucleic acid variation, the nucleic acid sequences of the recA sample were translated to their

corresponding amino acid sequences using the Expasy translate suite. Amino acid alignment of these amino acid sequences with their references showed that the observed variant exhibited a silent (synonymous) effect on the recA recombinase. Then, this synonymous variant was exemplified in the recA recombinase sequences according to its corresponding positions in the entire protein, namely p.193T=. As was expected in the other variants, the bacterial sequences usually alter their sequences to adapt to the host environment in which they are living.

To give a phylogenetic understanding of the actual distances of the investigated recA sample, a comprehensive phylogenetic tree was generated in the present study according to nucleic acid sequences observed in the amplified 648 bp of the *recA* gene amplicons. This phylogenetic tree was generated to incorporate this investigated sample alongside other relative nucleic acid sequences of human-infecting *Mycoplasma hominis* sequences. In addition to *Mycoplasma hominis*, other bacterial species were also incorporated within the same tree as outgroup sequences to assess the patterns of biological intra-species and variations within the incorporated sequences of the tree. These outgroup sequences are *Mycoplasma hyorhinis* (outgroup-1), *Mycoplasma hyopneumoniae* (outgroup-2), *Mycoplasma bovis* (outgroup-3), and *Mycoplasma mycoides* (outgroup-4). Within this tree, our investigated recA sample was aligned alongside other relative sequences to constitute the currently incorporated sequences within the cladogram. The total number of the aligned nucleic acid sequences in this comprehensive tree was sixty. In the constructed cladogram, the incorporated samples were clustered into five phylogenetic clades within the human-infecting *Mycoplasma* sequences. The most interesting fact observed in our investigated bacterial isolates is correlated with the positioning of our investigated sample within the *Mycoplasma hominis* clade. This clade consisted of the majority of incorporated samples of *Mycoplasma* sequences as twenty sequences of variable strains of *Mycoplasma* sequences were incorporated within extremely close phylogenetic distances in this clade (Fig. 3). Within this major clade, it seems that the recA sample did not exhibit a noticeable tilt from the other reference samples. This style of positioning was due to the presence of only one genetic variation in the recA sample, 329G>A. However, this variant showed no evolutionary role with respect to the original positioning of this bacterial sample occupied within the major clade of this cladogram. Furthermore, the aggregation of all incorporated bacterial samples within the clade of *Mycoplasma hominis* may refer to the presence of extremely close patterns of the phylogenetic distribution of the *recA*-based sequences within this clade. The current observation of this tree has confirmed sequencing reactions because it explained the actual neighbour-joining-based positioning in such investigated sequences. The recA sample was suited beside the GenBank acc. no. of CP026341.1 that belonged to *Mycoplasma hominis* (isolate 2539) deposited from Germany. Other closely related strains were also positioned in the vicinity of this strain, including two strains isolated from France (GenBank acc. no. CP038014.1) and Russia (GenBank acc. no. CP033021.1) respectively. In addition to the European origins of these samples, only one American origin was also observed in the same clade, such as the GenBank acc. no. CP011538.1. However, the European ancestors of our investigated samples could not be excluded from this explanation as the other investigated strains were also deposited from other portions of Europe. In the

vicinity of the *Mycoplasma hominis* clade, in which the recA samples were positioned, it was found that the *recA* sequences have relatively close phylogenetic positions to the clade of *Mycoplasma hyorhinis* (outgroup-1). Thus, both types of *Mycoplasma* sequences share a considerable ratio of homology in these sequences. This clade consists of seven strains and has a close connection with the clade of *Mycoplasma hyopneumoniae* (outgroup-2), in which only twelve strains were incorporated. However, the *Mycoplasma bovis* clade (outgroup-3) was found to be the closest clade to the clade of *Mycoplasma hyopneumoniae*. In the vicinity of the outgroup-3, the clade of *Mycoplasma mycoides* (outgroup-4) was suited. Within this clade, thirteen strains were incorporated in this clade. However, all four outgroups exhibited relatively similar phylogenetic positions with respect to the *Mycoplasma hominis* clade. Noteworthy, the utilization of the DNA *recA* gene sequences in this study has given further indication for the presence of the precise identification of the actual identity of this bacterial organism.

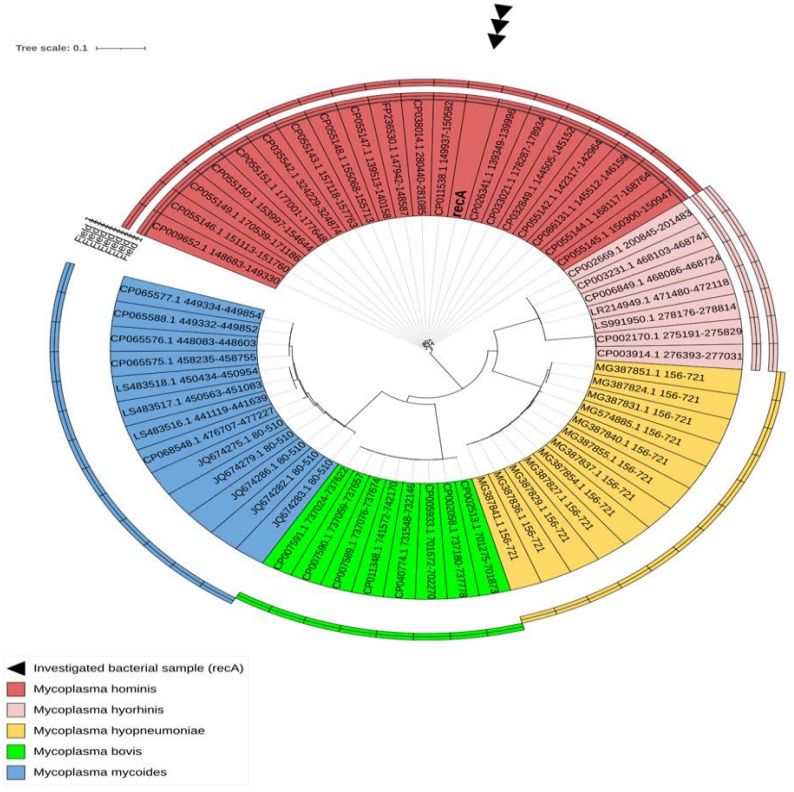

Fig. 3. The comprehensive cladogram phylogenetic tree of genetic variants of the DNA *recA* gene fragment of one sample of human-infecting *Mycoplasma hominis*. The black-colored triangle refers to the analyzed bacterial sequences. All the mentioned numbers referred to the GenBank accession number of each referring species. The number “0.1” at the top portion of the tree refers to the degree of scale range among the comprehensive tree categorized organisms. The letter “recA” refers to the code of the investigated sample.

Vaa gene of *Mycoplasma hominis* sequences

Within this locus, one sample (assigned Vaa) was included in the present study. These samples were screened to partially amplify the DNA *Vaa* gene that codes for variable adherence-associated lipoprotein adhesion in the *Mycoplasma hominis* sequences. Thus, the variation of the DNA *Vaa gene* can be used for genotyping of this bacteria due to its possible ability to adapt to variable genetic diversity as it was seen in different bacterial types. The sequencing reactions indicated the exact identity after performing NCBI blastn for these PCR amplicons. Concerning the 1158 bp amplicons, the NCBI BLASTn engine showed up to 100% sequences similarities between the sequenced samples and *Mycoplasma hominis* reference target sequences. By comparing the observed nucleic acid sequences of these investigated samples with the retrieved nucleic acid sequences (GenBank acc. CP032849.1), the accurate positions and other details of the retrieved PCR fragments were identified. The currently utilized amplicons were entirely covered the *Vaa* locus, which is concerned with the production of variable adherence-associated lipoprotein adhesion. The total length of the targeted loci was localized in the NCBI server, and the positions of the start and end of the targeted locus were also confirmed within the most homologous bacterial target. After positioning the 1158 bp amplicons' sequences within the genomic sequences of the *Mycoplasma hominis*, the details of its sequences were highlighted, and the total length of the amplified amplicons was also determined. Interestingly, the alignment results of the 1158 bp samples revealed the absence of any nucleic acid variation in the investigated Vaa sample in comparison with the most similar referring reference nucleic acid sequences (GenBank acc. no. CP032849.1). Our results indicated the presence of no detectable nucleic acid variants in the investigated Vaa sample. To confirm these variations, the sequencing chromatograms of the investigated samples, as well as their detailed annotations, were verified and documented, and the chromatograms of their sequences were directly visualized. The absence of any variants was confirmed in its original chromatogram and the absence of any possible technical error was also confirmed in the visualized chromatogram.The investigated nucleic acid sequences were however converted to their corresponding positions in the variable adherence-associated lipoprotein adhesion to see the actual position for the amplified fragment within the PCR amplicons and the entire protein. All nucleic acid sequences of the investigated samples were translated to their corresponding amino acid sequences using the Expasy translate tool. Amino acid alignment of these amino acid sequences with their references showed that the investigated 1158 bp consisted of 344 amino acid sequences in the variable adherence-associated lipoprotein adhesion. This coverage constituted the entire sequences of the investigated protein.

To give a phylogenetic understanding of the actual distances of the investigated Vaa sample, a comprehensive phylogenetic tree was generated in the present study according to nucleic acid sequences observed in the amplified 1158 bp of the *Vaa* gene amplicons. This phylogenetic tree was generated to incorporate this investigated sample alongside other relative nucleic acid sequences of human-infecting *Mycoplasma hominis* sequences. In addition to *Mycoplasma hominis* sequences, four other bacterial sequences of the same investigated genus were also incorporated within the same tree as outgroup sequences to assess the

patterns of biological intra-species variations within the incorporated sequences of the tree. These outgroup sequences are *Mycoplasma hyorhinis* (outgroup-1), *Mycoplasma hyopneumoniae* (outgroup-2), *Mycoplasma bovis* (outgroup-3), and *Mycoplasma mycoides* (outgroup-4). Within this tree, all our investigated sample was aligned alongside other relative *Mycoplasma* sequences to constitute the currently incorporated sequences within the generated cladogram. The total number of the aligned nucleic acid sequences in this comprehensive tree was 53. In the constructed cladogram, the incorporated samples were clustered into five phylogenetic clades within the *Mycoplasma* sequences. The most interesting fact observed in our investigated bacterial isolates is correlated with the positioning of our investigated sample within the *Mycoplasma hominis* clade. This clade consisted of the majority of incorporated samples of *Mycoplasma hominis* sequences as 21 sequences of variable strains of *Mycoplasma hominis* sequences were incorporated within extremely close phylogenetic distances within this clade (Fig.4). Within this major clade, the investigated Vaa sample was not tilted from the other investigated samples due to the absence of any detectable variation. The Vaa sample was suited beside one French strain (GenBank acc. no. CP035543.1), one English strain (GenBank acc. no. CP086131.1), two German strains (GenBank acc. no. CP026341.1 and CP055142.1), and two Russian strains (GenBank acc. no. CP033021.1 and CP032849.1). This pattern of positioning was clearly indicated obvious European ancestors of the investigated Vaa sample. Furthermore, the aggregation of all incorporated bacterial samples within the clade of *Mycoplasma hominis* may refer to the presence of extremely close distances of the phylogenetic distribution of these sequences within this clade. The current observation of this tree has confirmed sequencing reactions because it explained the actual neighbour-joining-based positioning in such investigated sequences. However, the positioning of this major clade with regard to the other *Mycoplasma* clades was also determined.

In the vicinity of the major *Mycoplasma hominis* clade, in which the Vaa sample was positioned, it was found that the *Vaa* sequences have relatively close phylogenetic positions to the clade of *Mycoplasma hyorhinis* (outgroup-1) with 6 incorporated strains. Thus, both types of *Mycoplasma* sequences share a relatively high ratio of homology in these sequences based on the *Vaa* gene sequences. In the vicinity of the *Mycoplasma hyorhinis* clade, *Mycoplasma hyopneumoniae* (outgroup-2) was located with 6 closely related strains. Apart from both outgroup-1 and outgroup-2 clades, the *Mycoplasma bovis* (outgroup-3) was located with 7 closely related strains. The second-largest clade within this tree is the clade of *Mycoplasma mycoides* (outgroup-4). This clade consists of 13 strains that have close distances to each other. Within the same clade. As in the case of the *recA*-gene based fragments, the utilization of the *Vaa* gene sequences in this study has given further proof for the presence of the precise identification of the actual identity of these bacterial organisms. This *Vaa* gene-based comprehensive tree has provided an inclusive tool about the high ability of *recA* and *Vaa* fragments to efficiently discriminate *Mycoplasma hominis* from the other investigated bacterial isolates of the same genus. This, in turn, gives a further indication of the ability of both *recA* and *Vaa* genes in the discrimination of the currently investigated human-infecting *Mycoplasma hominis* and its phylogenetic positions.

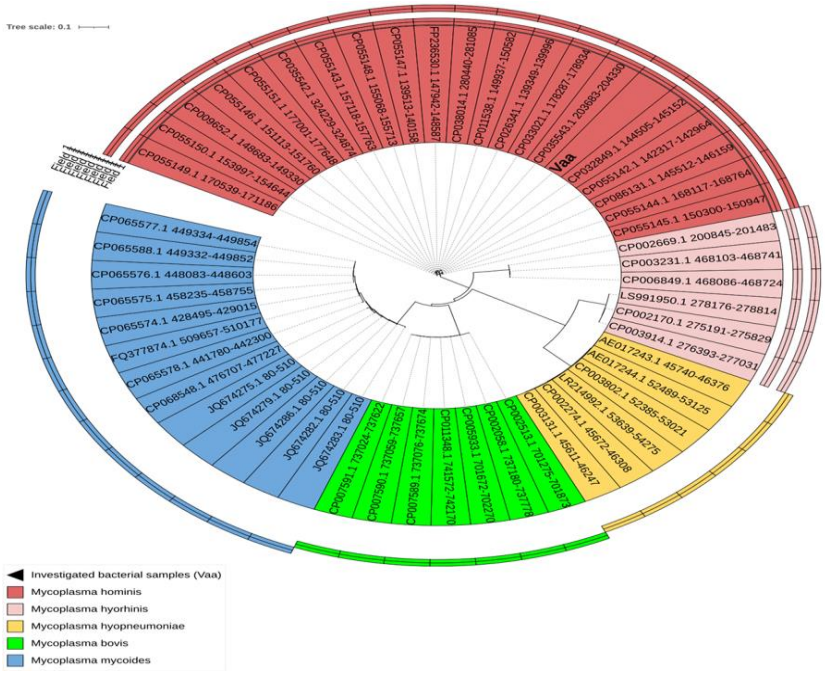

Fig. 4. The comprehensive cladogram phylogenetic tree of genetic variants of the DNA *Vaa* gene fragment of one sample of human-infecting *Mycoplasma hominis*. The black-colored triangle refers to the analyzed bacterial sequences. All the mentioned numbers referred to the GenBank accession number of each referring species. The number "0.1" at the top portion of the tree refers to the degree of scale range among the comprehensive tree categorized organisms. The letter "Vaa" refers to the code of the investigated sample.

Conclusions

In this study, the respective ability of 16S rRNA to detect the currently analyzed sequences of and *Vaa* in the detection of *Mycoplasma* are confirmed in the investigated patients. Concerning the *Vaa* fragment showed the lowest genetic diversity with no genetic variations. Accordingly, all the *Mycoplasma* genetic loci have proven efficiency in the detection of biological diversity and can be used to assess the biological diversity of bacterial infections in a broader spectrum. These promising amplicons can also be explored to discover further details within these identified serotypes in the other bacterial infections in the seminal fluids

References

1. Boesen, T.; Fedosova, . N. U;.and Kjeldgaad,.(2001). Molecular design of Mycoplasma hominis Vaa adhesion. Protein Sci. 10:2577-2569.[
2. Meygret A.; Le Roy C.; Renaudin H.; Bebear C.; and Pereyre S. (2018) Tetracycline and fluoroquinolone resistance in clinical *Ureaplasma* spp. and *Mycoplasma hominis* isolates in France between 2010 and 2015. *J. Antimicrob. Chemother.* 73, 2696–2703.
3. Matthew J. Allen-Daniels, B.S., Myrna G. Serrano, Ph.D.2 Lindsey P. Pflugner, B.S Jennifer M. Fettweis, Ph.D.1, Melissa A. Prestosa, B.S. et al

(2015)Identification of a gene in *Mycoplasma hominis* associated with preterm birth and microbial burden in intra-amniotic infection ,Department of Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, VA, USA *Am J Obstet Gynecol*.June ; 212(6): 779.e1–779.e13. doi:10.1016/j.ajog..01.032.

4. Ahmadi M. H.; Mirsalehian A.; Sadighi Gilani M. A.; Bahador A. & Talebi M. (2017)Asymptomatic Infection With *Mycoplasma hominis* Negatively Affects Semen Parameters and Leads to Male Infertility as Confirmed by Improved Semen Parameters After Antibiotic Treatment. Urology 100, 97–102.
5. Thomas Boesen, Jeppe Emmersen, Agata Baczynska1, Svend Birkelund and Gunna Christiansen (2004).The vaa locus of Mycoplasma hominis contains a divergent genetic islet encoding a putative membrane protein http://www.biomedcentral.com/1471-2180/4/37.
6. French Ct, Lao P, Loraine Ae, Matthews Bt, Yu H And Dybvig K. (2008). Large-scale transposon mutagenesis of *Mycoplasma pulmonis*. Mol Microbiol 69: 67-76.
7. Carvalho FM, Fonseca MM, Batistuzzo DE Medeiros S, Scortecci KC, Blaha CA And Agnez-Lima LF.(2005). DNA repair in reduced genome: the Mycoplasma model. Gene 360: 111-119.
8. Burgos R, Wood Ge, Young L, Glass Ji And Totten PA. (2012). RecA mediates MgpB and MgpC phase and antigenic variation in *Mycoplasma genitalium*, but plays a minor role in DNA repair. Mol Microbiol 85: 669-683.
9. Kamashev D, Oberto J, Serebryakova M, Gorbachev A, Zhukova Y, Levitskii S, Mazur Ak And Govorun V. (2011). *Mycoplasma gallisepticum* produces a histonelike protein that recognizes base mismatches in DNA. Biochemistry 50: 8692-8702.
10. Al-Ghizawi G. J. (2001) Typical and atypical pneumonia characteristics and bacterial profile of cases. Ph.D. thesis, College of Education, Basrah University.
11. Hashim HO, Mohammed MK, Mousa MJ, Abdulameer HH, Alhassnawi AT, Hassan SA, Al-Shuhaib MB. (2020) Infection with different strains of SARS-CoV-2 in patients with COVID-19. Archives of Biological Sciences. 25;72(4):575-85.
12. Zhang Z, Schwartz S, Wagner L, Miller W. (2000). A greedy algorithm for aligning DNA sequences. J Comput Biol. 7(1-2):203-14.
13. Letunic I, Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 2;47(W1): W256-W259.
14. Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2021). The COVID-19 pandemic. *International Journal of Health Sciences*, *5*(2), vi-ix. https://doi.org/10.53730/ijhs.v5n2.2937
15. Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2022). Post-pandemic health and its sustainability: Educational situation. *International Journal of Health Sciences*, *6*(1), i-v. https://doi.org/10.53730/ijhs.v6n1.5949