

EXPLORING MYCOPLASMAOIDES PNEUMONIAE THROUGH BIOINFORMATICS

From unknown DNA fragment to functional gene analysis

Abstract

This case study explores the contents of an unknown DNA fragment. BLAST analysis revealed that the sequence originates from ***Mycoplasmaoides pneumoniae***. Further investigation using bioinformatic tools such as **NCBI, RCSB, KBase, Benchling, InterPro, and AlphaFold** identified three key proteins: a **high-molecular-weight cytodherence protein, the chaperone protein DnaK, and the ATP-dependent 6-phosphofructokinase**. These proteins are involved in **adhesion, protein folding, and glycolysis**, respectively, highlighting the bacterium's adaptations for host dependence and survival.

Introduction

M. pneumoniae is a **wall-less, pleomorphic** bacterium from the class Mollicutes and the main cause of respiratory infections such as “**walking pneumonia**.” (Waites & Talkington, 2004). It has one of the **smallest genomes** of any self-replicating bacterium (~816 kb, 687 genes) and relies heavily on its host for essential nutrients and energy due to **reductive evolution**. As it lacks a **peptidoglycan cell wall**, it maintains a **cholesterol rich membrane** obtained from the host.

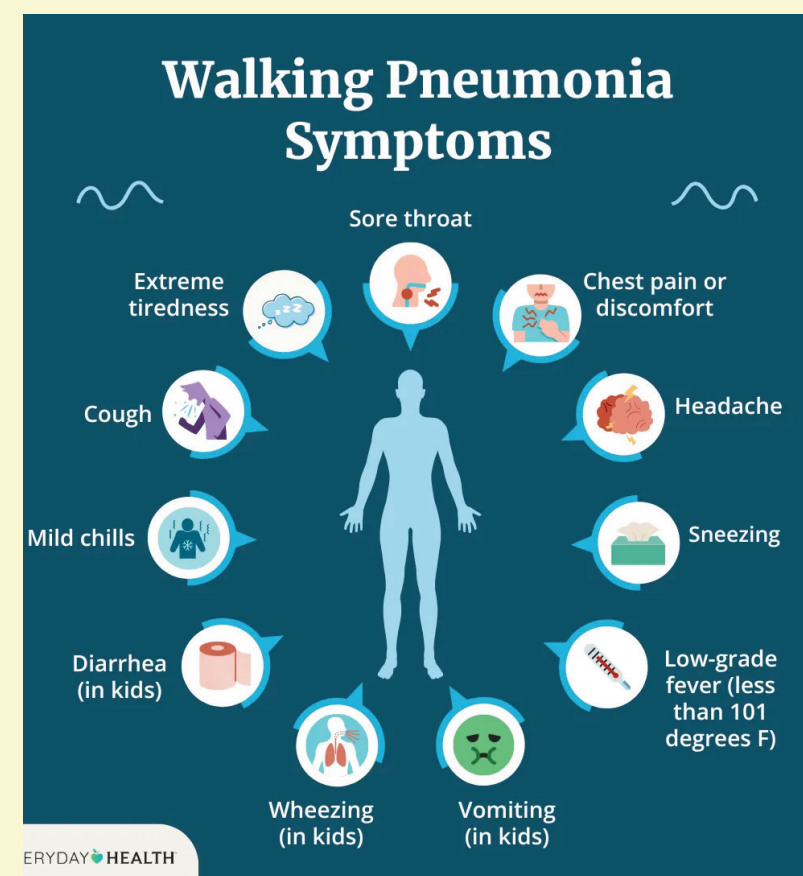


Figure 1: Walking pneumonia symptoms

The aim of this study was to **identify the origin of an unknown DNA fragment** using **bioinformatic tools** and to **analyze the biological functions** of its encoded **genes and proteins**. By examining **predicted ORFs, encoded proteins, phylogenetic relationships, and protein structure**, the DNA sequence was linked to functions contributing to the organism's **pathogenicity**.

Materials and Methods

Organism Identification and Gene Annotation

A BLASTx search of DNA Segment 21 was performed against the UniProt/ Swiss-Prot database to identify the organism.

To further analyze the sequence:

- ORFs were predicted for each of the three fragments within the segment, using the NCBI ORF Finder.
- From each fragment, the longest ORF was selected for further analysis using BLASTp, to identify homologous proteins.

Phylogenetic Analysis

- The ATP-dependent 6-phosphofructokinase (pfkA) sequence was compared across species using BLASTp.
- 25 homologous sequences were selected and aligned using Clustal Omega
- A phylogenetic tree was generated and visualized using iTOL to explore evolutionary relationships.

In Silico Cloning

- The pfkA gene was cloned in silico into the pET27b+R expression vector using Benchling.
- Vector sequence was retrieved from Addgene and annotated in Benchling.
- Primers containing BamHI and EcoRI restriction sites were designed using Primer-BLAST.
- Virtual PCR, restriction digestion, and ligation were carried out in Benchling's Assembly Wizard to create pET27b+R_pfkA.

Structural and Functional Protein Analysis

- The protein sequence of pfkA in *M. pneumoniae* was analyzed using InterProScan to gain insight into the function of the protein and its regulation.
- AlphaFold was utilized to obtain a prediction of the 3D model.

Literature research on human infection:

To better understand *M. pneumoniae* pathogenesis, a literature review was conducted on its interaction with human cells, focusing on adhesion, toxin activity, inflammation, and immune evasion

Results

BLASTp Results of longest ORFs

Fragment	ORF ID	nt length	aa length	Encoded Protein	BLAST match (organism/strain)	% ident
1	ORF22	2,883	960	protein Cytadherence	<i>Mycoplasmaoides pneumoniae</i> M129	99.90%
2	ORF52	1,728	576	protein Chaperone protein DnaK	<i>Mycoplasmaoides pneumoniae</i> M129	100%
3	ORF29	987	329	protein Probable ATP-dependent 6-phosphofructokinase	<i>Mycoplasmaoides pneumoniae</i> M129	100%

Table 1: Longest ORFs of each DNA fragment analysed with BLASTp

The longest ORFs identified correspond to key *Mycoplasmaoides pneumoniae* proteins: ORF22 encodes an adhesin critical for host cell attachment, ORF52 encodes the molecular chaperone protein DnaK, involved in protein folding under stress, and ORF29 encodes the enzyme ATP-dependent 6-phosphofructokinase, a central regulator in glycolysis. The high sequence identity confirms that the DNA fragments originate from *Mycoplasmaoides pneumoniae*.

Phylogenetic tree of phosphofructokinase homologs

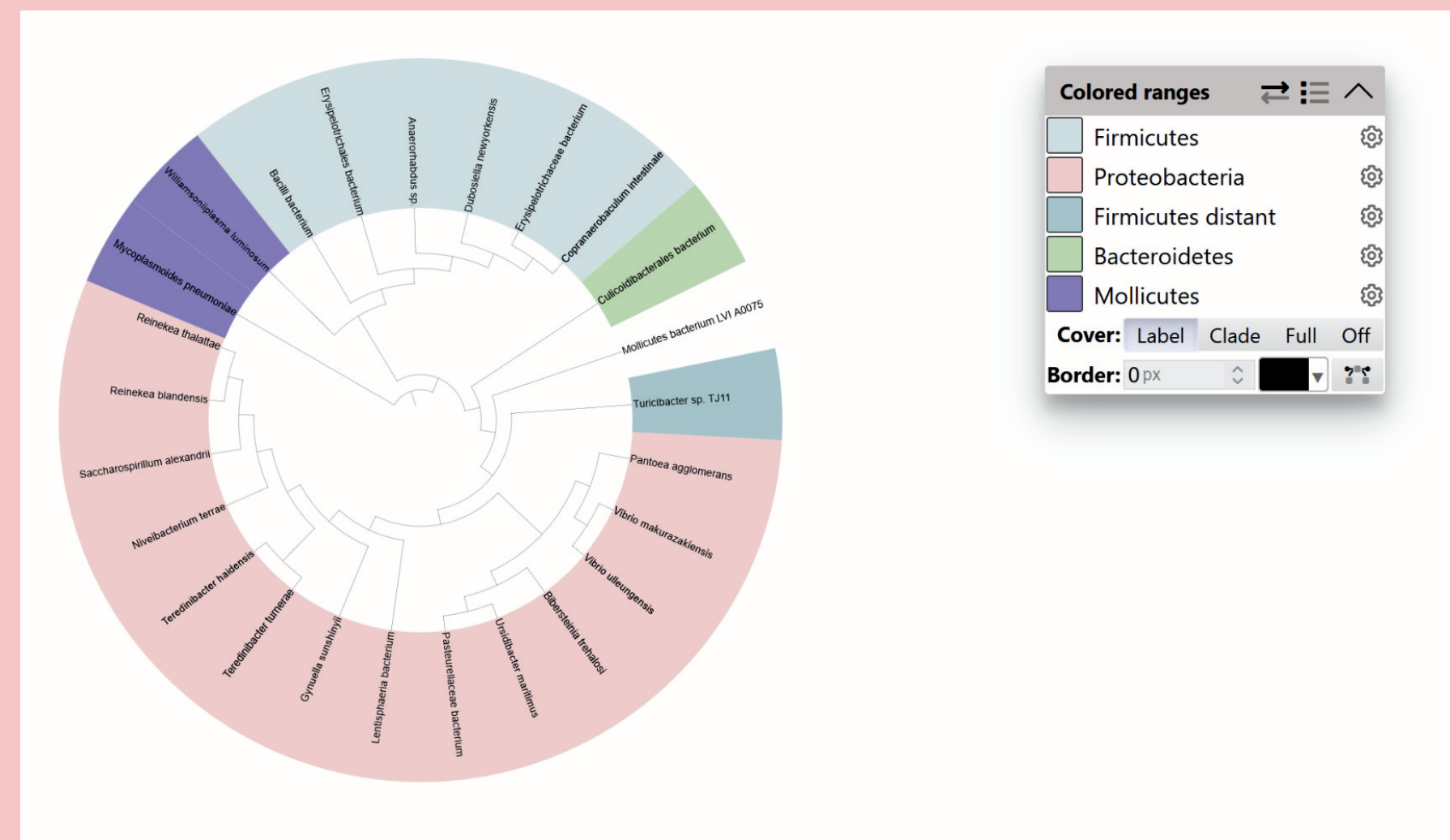


Figure 2: Phylogenetic tree of ATP-dependent 6-phosphofructokinase (PFK) homologs for 25 bacterial species:

The PfkA sequence from *Mycoplasmaoides pneumoniae* clusters with other Mollicutes, originating from Firmicutes. Unlike the highly regulated ATP-PFKs of Proteobacteria, *M. pneumoniae*'s enzyme functions constitutively due to its reduced genome. The tree illustrates the evolutionary plasticity of the PFK superfamily, shaped by phosphate-donor switching, regulatory diversification, and occasional horizontal gene transfer, while maintaining a conserved role in glycolysis across bacterial lineages.

Recombinant plasmid pET27b+R_GOI (6,378 bp) in silico

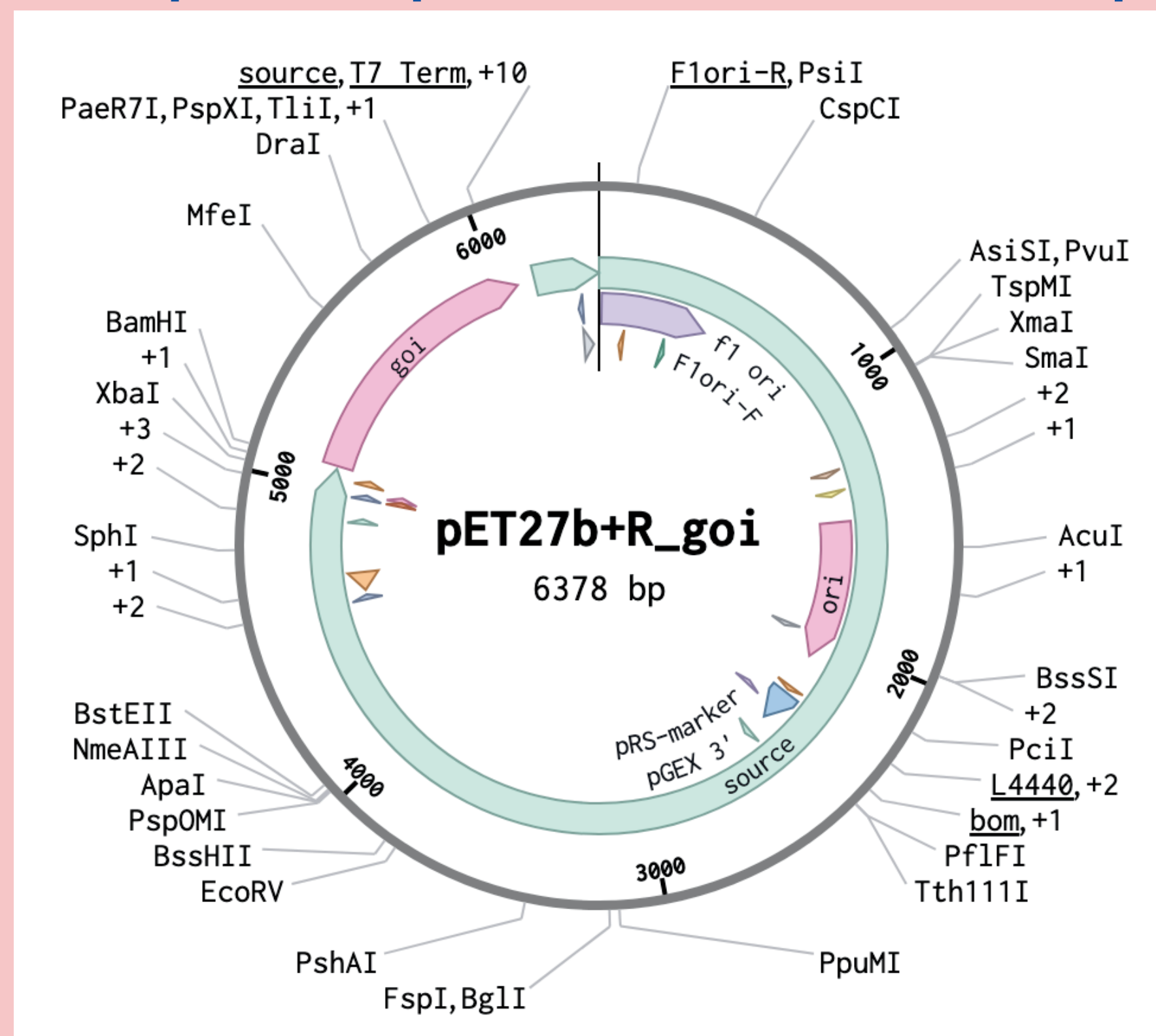


Figure 3: Benchling map of the pET27b+R_GOI construct (6,378 bp) showing the GOI insert (pink) cloned into the MCS using BamHI (5') and EcoRI (3') sites.

The recombinant plasmid pET27b+R_GOI (6,378 bp) was built by inserting the gene of interest GOI (pink) into the vector between the BamHI and EcoRI restriction sites. The construct includes a T7-promoter and a His-tag for protein expression and purification in *E. coli*. The gene is inserted in the correct orientation and reading frame, confirming a successful design.

Structural and Functional Protein Analysis

Function of pfkA

- Catalyzes F6P → F1,6BP, committing F6P to glycolysis (occurs in cytoplasm)
- Binds ATP and Mg-ATP at the catalytic site.
- Switches between open and closed conformations upon ligand binding.
- Allosterically regulated:
 - Activated by ADP/AMP
 - Inhibited by ATP/citrate

A predicted 3D model was obtained, using AlphaFold, to assess the structural features of the *M. pneumoniae* phosphofructokinase. The probable ATP-dependent 6-phosphofructokinase protein consists of two distinct domains. Domain 1 spans residues 7-142 and 265-312, while Domain 2 covers residues 145-260. The two domains belong to two distinct CATH homologous superfamilies.

AlphaFold model of phosphofructokinase domains



Figure 4: Domain organization of the probable ATP-dependent 6-phosphofructokinase protein predicted by AlphaFold. Domain 1 is green, and Domain 2 is orange.

Pathogenicity

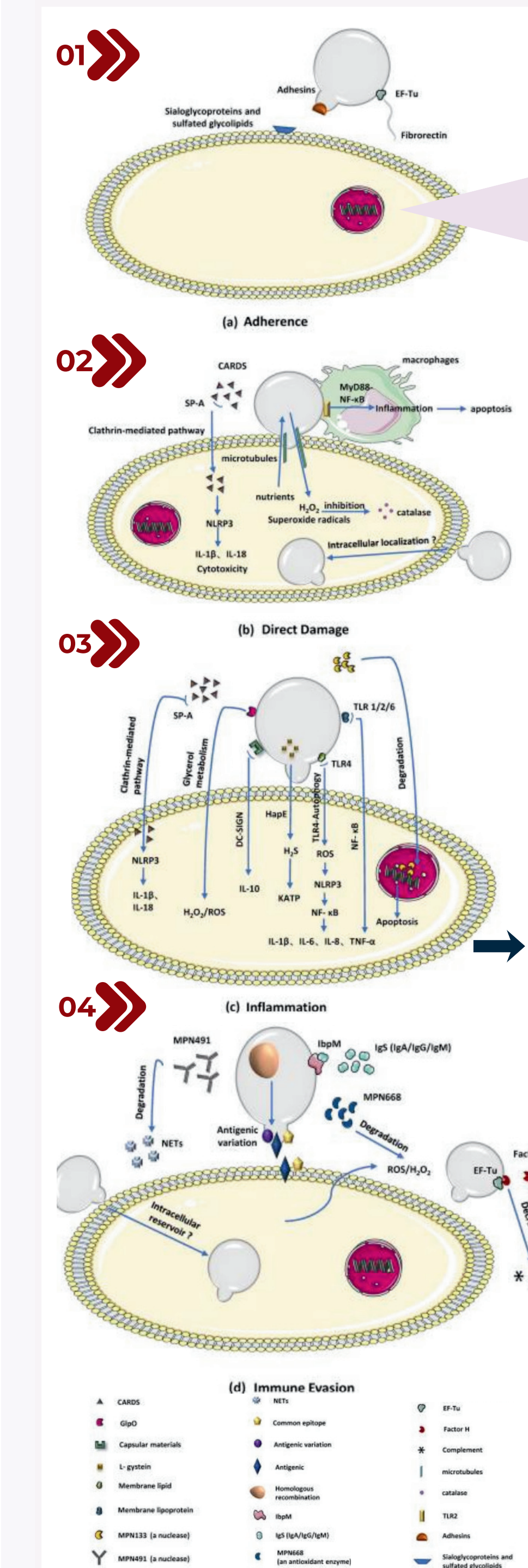


Figure 5. Pathogenic Mechanisms of *Mycoplasmaoides pneumoniae*:

- (1) Adherence via adhesins and EF-Tu, (2) Cytotoxicity by CARDS & ROS, (3) Inflammatory signaling through TLR and NLRP3 pathways
- (4) Immune evasion via nucleases, antioxidants, and antigenic variation.

01 Adherence

- *M. pneumoniae* adhesion causes cell damage
- Adhesins tightly bind to sialylated sugars on our airway cells steal nutrients and hijack the cell's metabolism
- EF-Tu proteins can also bind to fibronectin and improve colonisation

Cytadherence = the ability of a bacterium to stick to the surface of host cells

- *M. pneumoniae* cannot move on its own → terminal organelle
- complex (P1, P30, P116, HMW) → mediates host attachment
- glides slightly along surfaces

START INFECTION

- outside → adhesins P1, P30 are the main sticky proteins + accessory proteins (help P1 & P30)
- inside → gives it structure & stability

terminal button + paired plates + bowl complex

02 Direct Damage

- CARDS toxin → triggers inflammation and ciliary damage
- hydrogen peroxide (H₂O₂) and superoxide → damage membranes, proteins, and DNA → holes → cell injury & death

03 Inflammation (Over-active)

- toxins → irritate cell → activate NLRP3 → cell releases IL-1β & IL-18 → swelling, mucus, cough, tissue irritation

04 Immune Evasion

- Escapes neutrophil traps (NETs) → Produces nuclease (MPN491) that cuts apart DNA webs, freeing the bacterium
- Neutralizes oxidative attacks → Makes antioxidant enzyme (MPN668) that breaks down peroxides (H₂O₂) and protects from ROS
- Avoids antibody recognition → Performs homologous recombination, changing its surface antigens so immune cells can't detect it easily

Conclusion

The unknown DNA fragment was identified as originating from *Mycoplasmaoides pneumoniae* through BLAST analysis of the largest ORFs. *M. pneumoniae* is a small, wall-less bacterium that infects the respiratory tract using a specialized attachment organelle.

Using bioinformatic tools key proteins were identified — including a high-molecular-weight **cytodherence protein** essential for host attachment, the **molecular chaperone DnaK**, and the **ATP-dependent 6-phosphofructokinase** enzyme vital for glycolysis. Literature Research revealed pathogenicity involves adhesion-driven cell damage and virulence factors such as the CARDS toxin and ROS, which trigger inflammation and tissue injury.

Further analysis revealed evolutionary relationships of the pfkA gene, and in silico cloning of this gene into the pET27b+R vector was designed to enable functional protein expression. Future research may focus on overexpressing and characterizing this enzyme as a potential metabolic drug target, alongside exploring surface adhesins for clinical treatments of “walking pneumonia”.

