THE IRAQI JOURNAL OF
VETERINARY MEDICINE
eISSN:2410-7409
pISSN: 1609-5693

OPEN ACCESS

# *In Silico* Discovery of Novel Drug and Vaccine Targets in *Chlamydia abortus* Through Subtractive Genomics

**Zainab O Ali[1], Sura A Al-Asadi[2], Batool H Abdul Wahhab[3*], Jamila Rampurawala[4]**

*[1]Department of Laboratory for Medical, College of Health and Medical Techniques, Middle Technical University, Baghdad, Iraq, [2]Biotechnology Research Centre, Al Nahrain University, Baghdad, Iraq, [3]Department of Microbiology, Faculty of Medicine, Mustansiriyah University, Baghdad, Iraq, [4]Department of genetics, Oxford College of Science, Bangalore, India*

***Correspondence**: aldsbatool@gmail.com
Received: 06 May 2025
Revised: 10 August 2025
Accepted: 31 August 2025
Published : 28 December 2025

**DOI**: https://doi.org/10.30539/8gd4hm63



**Cite**: Ali ZO, Al-Asadi SA, Abdul Wahhab BH, Rampurawala J. *In silico* discovery of novel drug and vaccine targets in *Chlamydia abortus* through subtractive genomics. Iraqi J. Vet. Med. 2025;49(2):16-26

## ABSTRACT

*Chlamydia abortus* is a Gram-negative bacterium that causes chlamydiosis, a zoonotic disease leading to abortions and stillbirths in livestock and posing significant public health risks. With rising antimicrobial resistance and limited treatment efficacy, there is a pressing need for novel, targeted therapeutics to mitigate its economic and zoonotic impact. This study employed a subtractive genomics approach to analyze the complete proteome of *C. abortus* (strain DSM 27085 / S26/3), aiming to identify essential, non-host homologous proteins involved in unique bacterial metabolic pathways. Out of 932 proteins, five integral membrane proteins YidC, YajC, SecY, CAB503, and CAB746 were selected based on their essentiality, pathogen-specific roles, and absence of host homology. Anti-target screening confirmed no similarity to host proteins, ensuring minimal off-target effects. Antigenicity prediction (VaxiJen > 0.4) identified YidC, YajC, SecY, and CAB503 as strong vaccine candidates, while CAB746 exhibited variable antigenicity across species. Conservancy analysis showed YidC and YajC were highly conserved across *C. abortus* strains, while the others displayed strain-specific variations. Domain analysis revealed conserved motifs (e.g., IPR002208 in SecY) and transmembrane regions, supporting their structural and immunological relevance. In conclusion, YidC, YajC, SecY, CAB503, and CAB746 represent promising drug and vaccine targets. Their bacterial specificity, immunogenicity, and safety profile warrant further experimental validation to support targeted interventions against chlamydiosis.

**Keywords**: *Chlamydia abortus*, subtractive genomics, vaccine candidate, antimicrobial target, veterinary pathogen genomics

## INTRODUCTION

*Chlamydia abortus* (*C. abortus*) is a Gram-negative, obligatory intracellular bacterium belonging to the Chlamydiaceae family. It is recognized as a major cause of reproductive disorders in livestock, particularly among small ruminants such as sheep and goats. The pathogen is implicated in enzootic abortion, stillbirths, and the delivery of weak offspring, contributing to significant economic losses in animal production systems (**1-4**). In addition to ruminants, *C. abortus* can infect a broad host spectrum, including cattle, horses, guinea pigs, mice, and rabbits, highlighting its adaptability and zoonotic potential, causing keratoconjunctivitis in infected animals (**5**).

Transmission of *C. abortus* primarily occurs through direct exposure to infected placental tissues, amniotic fluid, or uterine discharges during or after parturition. Environmental contamination can sustain the pathogen,

facilitating indirect transmission via fomites, contaminated feed, or water sources (**3**). Asymptomatic carriers may shed the organism intermittently, complicating herd-level disease control. In humans, zoonotic transmission typically results from inhalation of aerosolized materials from infected animals, particularly during lambing seasons. While most human infections are subclinical or flu-like, severe cases in pregnant women can lead to miscarriage or systemic illness, underscoring the pathogen's public health importance (**6, 7**).

Given the challenges associated with traditional antimicrobial therapy, particularly the risk of resistance and lack of pathogen specificity, there is an urgent need to identify novel therapeutic strategies. Recent advances in computational biology have enabled high-throughput screening of microbial genomes to uncover species-specific targets. Subtractive genomics, in particular, has emerged as a valuable approach for identifying proteins essential to the pathogen's survival but absent in the host, thereby minimizing potential host toxicity (**8-11**).

In this study, we employed subtractive genomics to interrogate the proteome of *C. abortus* strain DSM 27085 / S26/3, selected for its virulence across multiple hosts. The analysis included removal of paralogous sequences, filtration against host homologs (**12**), and selection of non-redundant, pathogen-specific proteins. Further prioritization was performed through subcellular localization to identify membrane-associated proteins, which are more accessible to drugs and immune responses. Only conserved targets across multiple *C. abortus* strains were retained, increasing the likelihood of broad-spectrum efficacy. This study aims to identify novel drug and vaccine candidates for targeted intervention against *C. abortus*, ultimately contributing to the reduction of its zoonotic transmission and veterinary burden.

## MATERIALS AND METHODS

### Proteome Retrieval and Removal of Duplicated/Paralogous Sequences

The analysis of the whole proteome of *C. abortus* was performed to identify potential drug targets and vaccine candidates through a subtractive genomics approach (**Figure 1**).

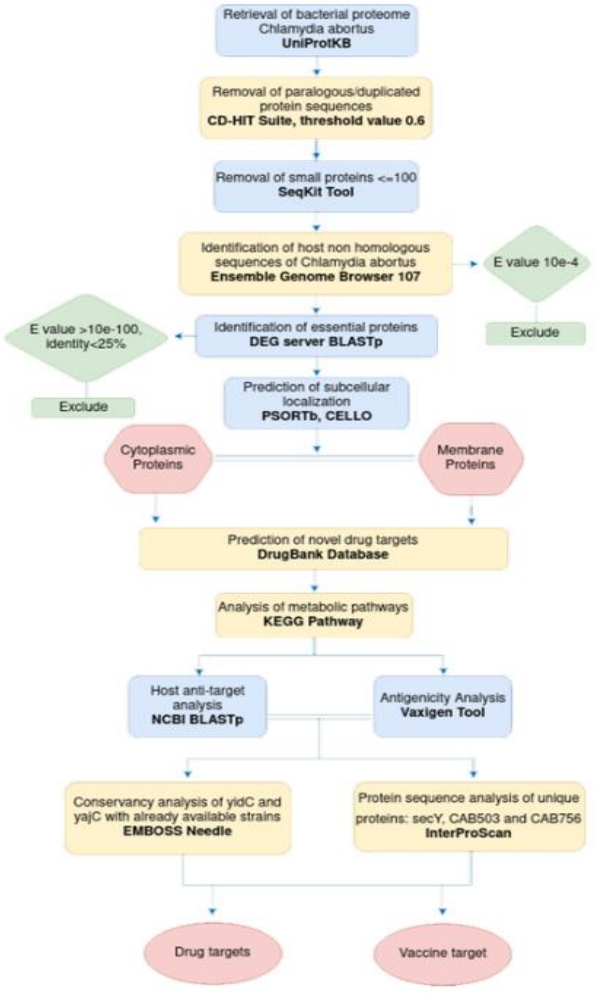


**Figure 1.** Schematic representation of complete roadmap to identity common drug and vaccine targets against *C. abortus*

The complete proteome of *C. abortus* (strain DSM 27085 / S26/3) was retrieved from the UniProtKB Proteomes database (https://www.uniprot.org/proteomes/UP000001012) (13). To remove duplicated or paralogous protein sequences, the CD-HIT suite (http://weizhong-lab.ucsd.edu/cdhit_suite/cgi-bin/index.cgi) was employed with a sequence identity threshold of 0.6. This step ensured the exclusion of functionally redundant proteins from the dataset (**14**). Uncharacterized proteins—those lacking functional annotation—were excluded from further analysis to focus only on biologically relevant targets. The resulting set of non-paralogous, characterized proteins was further refined using the SeqKit toolkit (https://bioinf.shenwei.me/seqkit/) (**15**), where proteins with a sequence length of fewer than 100 amino acids were filtered out. These short sequences, typically considered less likely to represent full-length functional proteins, were excluded and stored in a separate dataset referred to as Bin 1 for documentation purposes.

## Identification of Non-Homologous Proteins to Host Proteomes

This study excluded homologous protein sequences of *C. abortus* from host proteomes to prevent potential interactions with host proteins during drug targeting. The host proteomes included *Bos taurus* (Bovine), *Cavia porcellus* (Guinea pig), *Equus caballus* (Horse), *Mus musculus* (Mouse), *Oryctolagus cuniculus* (Rabbit), and *Ovis aries* (Sheep). The BLAST tool from Ensembl Genome Browser 107 (https://asia.ensembl.org/index.html) (**16**) was used to identify homologous proteins in each host species. Protein sequences with E-values greater than $10^{-4}$ were considered homologous and subsequently excluded. The non-homologous sequences were placed into bin 2a, 2b, 2c, 2d, 2e, and 2f, with each alphabet corresponding to a specific host species.

## Identification of Essential Non-Homologous Proteins of *C. abortus*

Essential proteins among the non-homologous *C. abortus* sequences were identified using the Database of Essential Genes (DEG) version 10 (http://origin.tubic.org/deg/public/index.php) (**17**). The BLASTp algorithm was employed using all bacterial genomes from the DEG database. The resulting output was filtered based on a minimum bit score of 100 and an E-value threshold of $10^{-10}$. Proteins with an identity score ≥ 25% and E-value ≤ $10^{-100}$ were considered essential. These essential proteins were then sorted into bins 3a, 3b, 3c, 3d, 3e, and 3f, corresponding to the non-homologous protein bins from the previous step.

## Analysis of Metabolic Pathways

The analysis of bacteria and host species metabolic pathways was performed by the Kyoto Encyclopedia of Genes and Genomes (KEGG) (https://www.genome.jp/kegg/) (**18, 19**). The bacterial species, C abortus, was analyzed and utilized the code “cab”. The metabolic pathways of host organisms, including the same host proteomes, were retrieved using codes “oas”, “mmu”, “bta”, “ocu”, “ecb” and “cpoc”. These pathways were used to compare against the metabolic pathways of *C. abortus*. The Kegg Orthology (KO) database was utilized to verify the availability of the metabolic proteins concerning these pathways. Only those metabolic proteins associated with unique metabolic pathways were considered and listed in bin 4a, bin 4b, bin 4c, bin 4d, bin 4e, and bin 4f.

## Prediction of Subcellular Localization

*C. abortus*, a Gram-negative bacterium, has proteins categorized into five subcellular compartments: cytoplasm, extracellular space, inner membrane, periplasm, and outer membrane (**20**). Subcellular localization was predicted using CELLO v2.5 (http://cello.life.nctu.edu.tw/) (**21**) and PSORTb v3.0.3 (https://www.psort.org/psortb/) (**22**). Only the location identified as common by both tools and with the highest prediction score was selected. Due to its significance in identifying drug targets and vaccine candidates, the periplasmic compartment was specifically validated for each protein. Periplasmic proteins commonly predicted by both tools with strong confidence scores were compiled into bin 5a, 5b, 5c, 5d, 5e, and 5f.

## Prediction of Novel Drug Targets

The periplasmic proteins in bin 5 were screened using the DrugBank 5.0 (version 6.0) database (https://dev.drugbank.com/guides/drugbank/citing) (**23**) to assess drug ability. Proteins without existing drug associations were classified as novel drug targets and stored in bin 6 for further investigation.

## Antitarget Analysis of Non-Homologous, Essential, and Novel Drug Targets

During drug development, certain proteins may unintentionally interact with host biomolecules, known as off-targets or "antitargets," potentially leading to adverse side effects. To minimize such risks, the *C. abortus* (DSM 27085/S26/3) proteome was subjected to antitarget analysis (**24**) using NCBI BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) against various host species. Proteins showing significant similarity to host proteins were flagged as antitargets and excluded. Selection criteria included E-value < 0.005, query coverage > 30%, and sequence identity < 25%. Remaining non-antitarget proteins were documented in bin 7a, bin 7b, bin 7c, bin 7d, bin 7e, and bin 7f.

## Antigenicity Analysis of Novel Drug Targets

Antigenicity reflects the potential of a protein to elicit an adaptive immune response, which is important for vaccine design. VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) (**25**) was used to predict the antigenicity of proteins in bin 7, using a threshold score of 0.4. Proteins exceeding this threshold were considered potential vaccine candidates with clinical relevance.

### Conservancy Analysis of Antigenic Proteins of *C. abortus* (strain DSM 27085 / S26/3)

Conservancy analysis across multiple bacterial strains was performed to assess the potential drug spectrum of *C. abortus* (DSM 27085 / S26/3) antigenic proteins. Homologous sequences for each antigenic protein were retrieved from the UniProtKB database. These sequences were aligned using EMBOSS Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/), a tool designed for pairwise sequence alignment. For antigenic proteins that were identified in only a single strain, domain analysis was performed using InterProScan (https://www.ebi.ac.uk/interpro/search/sequence/) (**26**) to evaluate structural and functional domains.

## Results and Discussion

The objective of this study was to identify novel therapeutic targets against *C. abortus* (strain DSM 27085 / S26/3). To achieve this, a subtractive genomics approach was employed to comprehensively analyze the complete proteome of *C. abortus* (strain DSM 27085 / S26/3) using a range of bioinformatics tools and databases. A schematic overview of the subtractive genomics workflow is presented in **Table 1**.

**Table 1.** A subtractive genomics analysis scheme is employed to identify common and novel therapeutic drug and vaccine targets against *C. abortus*

| Sr. No. | Subtractive Approaches | Bioinformatics Tools and Databases | Number of Proteins (respectively) |
|---|---|---|---|
| 1 | Retrieval of whole proteome of B. abortus (strain 2308) | UniProtKB database | 932 |
| 2 | Removal of non-paralogous (>60% identity) and smaller proteins (<100 amino acids) | CD-Hit suite, SeqKit toolkit | 931 |
| 3 | Proteins non homologous to Ovis aries, *Mus musculus*, *Bos taurus*, Oryctolagus cuniculus, *Equus caballus*, *Cavia porcellus* | Ensembl BLASTp (E value 10e-4) | 198, 184, 405, 215, 143, 147 |
| 4 | Identification of essential proteins against Ovis aries, *Mus musculus*, *Bos taurus*, Oryctolagus cuniculus, *Equus caballus*, *Cavia porcellus*. | DEG 15.2 server (E value <= 10e-100, Bit score > 100) | 137, 105, 91, 73, 43, 39 |
| 5 | Essential membrane proteins against Ovis aries, *Mus musculus*, *Bos taurus*, Oryctolagus cuniculus, *Equus caballus*, *Cavia porcellus*. | PSORTb, CELLO | 14, 11, 13, 6, 10, 5 |
| 6 | Essential proteins that are found to be novel in recommended database | DrugBank 5.1.9 (default parameters) | 7, 6, 11, 3, 5, 3 |
| 7 | Metabolic proteins with assigned KO | KEGG Orthology | 3, 2, 3, 1, 2, 1 |
| 8 | Novel drug target proteins non-homologous to anti targets | NCBI BLASTp (E value < 0.005, identity < 25%, Query length >30%) | 3, 2, 3, 1, 2, 1 |
| 9 | Proteins showing antigenicity (bin 8a, 8b, 8c, 8e, 8f) | VaxiJen v2.0 (Threshold value > 0.4) | 3, 2, 3, 2, 1 |
| 10 | Highly conserved proteins (bin 9a, 9b, 9c, 9e, 9f) | EMBOSS Needle | 2, 1, 1, 1, 1 |
| 11 | Protein sequence analysis (bin 10a, 10b, 10c, 10e) | InterProScan | 1, 1, 1, 1 |

### Removal of Paralogous Sequences

The complete proteome of *C. abortus* (strain DSM 27085 / S26/3) comprised 932 protein sequences. The CD-HIT suite was employed to remove duplicated sequences, using a similarity threshold of 0.6 to exclude sequences sharing more than 60% identity. Only one duplicated protein was identified and removed, resulting in 931 non-paralogous sequences. Uncharacterized proteins, those lacking defined biological function (**14**) were excluded, leaving 901 functionally characterized proteins. Additionally, proteins shorter than 100 amino acids, which are often non-essential or lack structural domains (**27**), were filtered out using the SeqKit toolkit. This resulted in 801 proteins, which were stored in bin 1. The minimal presence of paralogous proteins highlights the compact nature of the *C. abortus* proteome, which simplifies the identification of potential therapeutic targets.

### Identification of Host Homologous Protein Sequences

Homologous proteins shared between *C. abortus*, and host organisms may exhibit structural similarities that increase the risk of cross-reactivity with host proteins during drug treatment (**28**). To mitigate this, protein sequences of *C. abortus* (strain DSM 27085 / S26/3) were compared against the proteomes of six host species: *Ovis aries* (sheep), *Mus musculus* (mouse), *Bos taurus* (bovine), *Oryctolagus cuniculus* (rabbit), *Equus caballus* (horse), and *Cavia porcellus* (guinea pig). Non-homologous proteins, those not aligning with host sequences—were retained, with 198 identified for *Ovis aries*, 184 for *Mus musculus*, 405 for *Bos taurus*, 215 for *Oryctolagus cuniculus*, 143 for *Equus caballus*, and 147 for *Cavia porcellus*. These were organized into bin 2a through bin 2f, respectively. The variation in non-homologous protein counts reflects differing levels of evolutionary divergence, with *Bos taurus* showing the greatest dissimilarity to *C. abortus*, indicating a larger set of potential therapeutic targets specific to the pathogen.

### Selection of Essential Proteins of *C. abortus* (strain DSM 27085 / S26/3)

The survival of any organism is contingent upon the presence of essential proteins, which serve as critical drug targets for therapeutic agents designed to eliminate bacterial pathogens (**12**). To identify such targets, host non-homologous proteins of *C. abortus* (strain DSM 27085 / S26/3) from bin 2a (*Ovis aries*), bin 2b (*Mus musculus*), bin 2c (*Bos taurus*), bin 2d (*Oryctolagus cuniculus*), bin 2e (*Equus caballus*), and bin 2f (*Cavia porcellus*) were analyzed using the BLASTp tool from the Database of Essential Genes (DEG). Proteins were classified as essential if they exhibited an E-value of $< 10^{-100}$ and an identity score greater than 25% when aligned against the DEG bacterial essential protein dataset. This analysis yielded 137 essential proteins

in bin 2a (stored in bin 3a), 105 in bin 2b (bin 3b), 91 in bin 2c (bin 3c), 73 in bin 2d (bin 3d), 43 in bin 2e (bin 3e), and 39 in bin 2f (bin 3f). These proteins represent key functional vulnerabilities of *C. abortus* across various host species and are prime candidates for downstream drug target prioritization.

## Analysis of Subcellular Localization of Shortlisted Proteins

As a Gram-negative bacterium, *C. abortus* (strain DSM 27085 / S26/3) produces proteins that may localize to one of five subcellular compartments: cytoplasm, extracellular space, inner membrane, periplasm, or outer membrane. Determining the subcellular location of proteins is crucial for therapeutic design. Cytoplasmic proteins are typically targeted by small-molecule drugs, while membrane and periplasmic proteins are more accessible and thus preferred as vaccine or drug candidates (**29**). Membrane proteins, in particular, perform essential biological roles such as transport, signal transduction, and metabolic regulation (**30**).

The essential proteins identified in bins 3a through 3f were subjected to subcellular localization prediction.

For bin 3a *Ovis aries* (sheep), 122 proteins were predicted to be cytoplasmic, while 14 were membrane-associated. These were subsequently stored in bin 4a. Results for the remaining bins followed a similar classification, though only bin 3a data is shown here (**Table 2**).

**Table 2.** Subcellular localization of essential *C. abortus* proteins non-homologous to *Ovis aries* (sheep)

| SeqID | UniProt Acc. | Localization | Score |
|---|---|---|---|
| Pfp | Q5L5Q8 | Cytoplasmic membrane | 8.46 |
| murG | Q5L524 | Cytoplasmic membrane | 7.88 |
| sdhB | Q5L4S1 | Cytoplasmic membrane | 7.88 |
| CAB867 | Q5L4Y8 | Cytoplasmic membrane | 10 |
| CAB855 | Q5L500 | Cytoplasmic membrane | 10 |
| yajC | Q5L504 | Cytoplasmic membrane | 9.86 |
| lpxB | Q5L586 | Cytoplasmic membrane | 9.99 |
| CAB483 | Q5L5Z6 | Cytoplasmic membrane | 10 |
| fabD | Q5L607 | Cytoplasmic membrane | 8.46 |
| yidC | Q5L620 | Cytoplasmic membrane | 10 |
| Lgt | Q5L621 | Cytoplasmic membrane | 10 |
| CAB375 | Q5L6A0 | Cytoplasmic membrane | 9.82 |
| secY | Q5L702 | Cytoplasmic membrane | 10 |
| CAB750 | Q5L595 | Periplasmic | 9.76 |

For bin 3b *Mus musculus* (mouse), 94 proteins were cytoplasmic, and 11 were membrane proteins (**Table 3**), stored in bin 4b. For bin 3c *Bos taurus* (bovine), 60 proteins were cytoplasmic, 18 had unknown localization, and 13 were membrane proteins (**Table 4**), stored in bin 4c. For bin 3d *Oryctolagus cuniculus* (rabbit), 60 proteins were cytoplasmic, 1 was extracellular, 6 had unknown localization, and 6 were membrane proteins (**Table 5**), stored in bin 4d. It was observed that 30 proteins from bin 3e were located at the cytoplasmic region, 3 proteins had unknown subcellular locations, and 10 proteins showed membrane location for *Equus caballus* (horse) as in **Table 6**. These 13 membrane proteins were listed in bin 4e. For bin 3f *Cavia porcellus* (guinea pig), 33 proteins were cytoplasmic, 4 had unknown localization, and 5 were membrane proteins (**Table 7**), stored in bin 4f.

**Table 3.** Subcellular localization of essential *C. abortus* proteins non-homologous to *Mus musculus* (mouse)

| SeqID | UniProtKB | Localization | Score |
|---|---|---|---|
| Psd | Q5L4W1 | Cytoplasmic membrane | 7.88 |
| dacA | Q5L5K2 | Cytoplasmic membrane | 10 |
| CAB483 | Q5L5Z6 | Cytoplasmic membrane | 10 |
| fabD | Q5L607 | Cytoplasmic membrane | 8.46 |
| yidC | Q5L620 | Cytoplasmic membrane | 10 |
| Lgt | Q5L621 | Cytoplasmic membrane | 10 |
| npt2 | Q5L6Y7 | Cytoplasmic membrane | 10 |
| CAB951 | Q5L4Q7 | Cytoplasmic membrane | 9.27 |
| CAB746 | Q5L599 | Cytoplasmic membrane | 10 |
| ompA | P16567 | Outer Membrane | 10 |
| tolB | Q5L4R3 | Periplasmic | 9.44 |

**Table 4.** Subcellular localization of essential *C. abortus* proteins non-homologous to *Bos taurus* (bovine)

| SeqID | UniProt Acc. | Localization | Score |
|---|---|---|---|
| Pfp | Q5L5Q8 | Cytoplasmic membrane | 8.46 |
| murG | Q5L524 | Cytoplasmic membrane | 7.88 |
| lspA | Q5L6Q8 | Cytoplasmic membrane | 10 |
| CAB855 | Q5L500 | Cytoplasmic membrane | 10 |
| yajC | Q5L504 | Cytoplasmic membrane | 9.86 |
| dacA | Q5L5K2 | Cytoplasmic membrane | 10 |
| CAB641 | Q5L5K3 | Cytoplasmic membrane | 10 |
| CAB483 | Q5L5Z6 | Cytoplasmic membrane | 10 |
| yidC | Q5L620 | Cytoplasmic membrane | 10 |
| Lgt | Q5L621 | Cytoplasmic membrane | 10 |
| secY | Q5L702 | Cytoplasmic membrane | 10 |
| Lnt | Q5L726 | Cytoplasmic membrane | 10 |
| CAB049 | Q5L761 | Cytoplasmic membrane | 10 |

**Table 5.** Subcellular localization of essential *C. abortus* proteins non-homologous to *Oryctolagus cuniculus* (rabbit)

| SeqID | UniProt Acc | Localization | Score |
|---|---|---|---|
| Lgt | Q5L621 | Cytoplasmic membrane | 10 |
| secY | Q5L702 | Cytoplasmic membrane | 10 |
| mraY | Q5L520 | Cytoplasmic membrane | 10 |
| murG | Q5L524 | Cytoplasmic membrane | 7.88 |
| bioD | Q5L5F7 | Cytoplasmic membrane | 9.82 |
| lepA | Q5L659 | Cytoplasmic membrane | 7.88 |

**Table 6.** Subcellular localization of essential *C. abortus* proteins non-homologous to *Equus caballus* (horse)

| SeqID | UniProtKB | Localization | Score |
|---|---|---|---|
| lolC | Q5L5W5 | Cytoplasmic membrane | 9.46 |
| CAB503 | Q5L5X6 | Cytoplasmic membrane | 10 |
| CAB501 | Q5L5X8 | Cytoplasmic membrane | 10 |
| CAB368 | Q5L6A7 | Cytoplasmic membrane | 10 |
| CAB013 | Q5L797 | Cytoplasmic membrane | 10 |
| yidC | Q5L620 | Cytoplasmic membrane | 10 |
| Lgt | Q5L621 | Cytoplasmic membrane | 10 |
| ftsH | Q5L5B3 | Cytoplasmic membrane | 10 |
| Pfp | Q5L5Q8 | Cytoplasmic membrane | 8.46 |
| CAB468 | Q5L610 | Outer Membrane | 10 |

**Table 7.** Subcellular localization of essential *C. abortus* proteins non-homologous to *Cavia porcellus* (guinea pig)

| SeqID | UniProtKB | Localization | Score |
|---|---|---|---|
| CAB483 | Q5L5Z6 | Cytoplasmic membrane | 10 |
| yidC | Q5L620 | Cytoplasmic membrane | 10 |
| Lgt | Q5L621 | Cytoplasmic membrane | 10 |
| Psd | Q5L4W1 | Cytoplasmic membrane | 7.88 |
| murG | Q5L524 | Cytoplasmic membrane | 7.88 |

## Novel Drug Target Analysis

Membrane proteins from bins 4a, 4b, 4c, 4d, 4e, and 4f were submitted to the DrugBank database to identify novel drug targets. For bin 4a (*Ovis aries*), 7 targets (CAB855, yajC, lpxB, CAB483, yidC, lgt, secY) showed no interaction with existing therapeutic compounds and were deemed novel, stored in bin 5a. For bin 4b (*Mus musculus*), 6 novel targets (dacA, CAB483, yidC, lgt, CAB951, CAB746) were identified and stored in bin 5b. For bin 4c (*Bos taurus*), 11 novel targets (lspA, CAB855, yajC, rnpA, dacA, CAB483, yidC, lgt, secY, lnt, CAB049) were found and stored in bin 5c. For bin 4d (*Oryctolagus cuniculus*), 3 novel targets (murG, bioD, lepA) were identified and stored in bin 5d. For bin 4e (*Equus caballus*), 5 novel targets (CAB503, CAB501, CAB013, yidC, lgt) were found and stored in bin 5e. For bin 4f (*Cavia porcellus*), 3 novel targets (CAB483, yidC, lgt) were identified and stored in bin 5f. These novel targets represent unexploited opportunities for developing specific therapeutics against *C. abortus*.

## Analysis of Metabolic Pathways

The metabolic pathways of *C. abortus* (strain DSM 27085 / S26/3) and host species (*Ovis aries*, *Mus musculus*, *Bos taurus*, *Oryctolagus cuniculus*, *Equus caballus*, *Cavia porcellus*) were retrieved from the KEGG Pathway database and compared. Of 21 proteins analyzed, 6 were involved in unique bacterial metabolic pathways, termed pathogen-specific, and listed in **Table 8**. All 6 proteins had KEGG Orthology (KO) assignments, confirming their role in bacterial-specific metabolism. These proteins are ideal for targeted interventions as they avoid interference with host metabolism. They were stored in bins 6a, 6b, 6c, 6d, 6e, and 6f.

**Table 8.** Proteins identified that are involved in pathogen specific metabolic pathways

| Protein | UniProt ID | KO Assignment | Pathways |
|---|---|---|---|
| yajC | Q5L504 | K03210 | Bacterial secretion system |
| yidC | Q5L620 | K03217 | Bacterial secretion system |
| secY | Q5L702 | K03076 | Bacterial secretion system |
| CAB503 | Q5L5X6 | K03179 | Biosynthesis of secondary metabolites |
| CAB746 | Q5L599 | K17103 | Biosynthesis of secondary metabolites |
| bioD | Q5L5F7 | K01935 | Biotin metabolism |

## Antitarget Analysis of Novel Drug Targets

Therapeutic compounds interacting with antitargets or off-targets can cause side effects, necessitating their evaluation to avoid carcinogenicity and cross-reactivity (**31**). To ensure safety, NCBI BLAST was used to assess shortlisted proteins (yidC, yajC, secY, CAB503, CAB746) from bins 6a–6f against host proteomes (*Ovis aries*, *Mus musculus*, *Bos taurus*, *Oryctolagus cuniculus*, *Equus caballus*, *Cavia porcellus*). None showed significant hits with identity ≤25%, indicating they were non-antitargets. All proteins were retained and stored in bins 7a, 7b, 7c, 7d, 7e, and 7f. This confirms their suitability for safe therapeutic development.

## Antigenicity Prediction Analysis

Drug targets with antigenicity above the threshold are effective for designing recombinant vaccines to prevent infectious diseases (32). To evaluate vaccine potential, the VaxiJen server analyzed novel drug targets (yidC, yajC, secY, CAB503, CAB746, bioD) from bins 7a–7f. All proteins except bioD (bin 7d, score 0.2773, below threshold 0.4) showed antigenicity scores >0.4 and were deemed antigenic. Antigenic proteins for *Ovis aries* (bin 8a), *Mus musculus* (bin 8b), *Bos taurus* (bin 8c), *Equus caballus* (bin 8e), and *Cavia porcellus* (bin 8f) are listed in **Table 9**.

**Table 9.** Drug targets in sheep, mice, cow, rabbit, horse, and guinea pig with their antigenicity score

| Animal | Protein | UniProt ID | Vaxijen score | Antigenicity |
|---|---|---|---|---|
| Sheep | yajC | Q5L504 | 0.7414 | Probable antigen |
| | yidC | Q5L620 | 0.4299 | Probable antigen |
| | secY | Q5L702 | 0.4142 | Probable antigen |
| Mice | yidC | Q5L620 | 0.4299 | Probable antigen |
| | CAB746 | Q5L599 | 0.4912 | Probable antigen |
| | yidC | Q5L620 | 0.4299 | Probable antigen |
| Cow | yajC | Q5L504 | 0.7414 | Probable antigen |
| | yidC | Q5L620 | 0.4299 | Probable antigen |
| | secY | Q5L702 | 0.4142 | Probable antigen |
| Rabbit | bioD | Q5L5F7 | 0.2773 | Probable non-antigen |
| Horse | yidC | Q5L620 | 0.4299 | Probable antigen |
| | CAB503 | Q5L5X6 | 0.609 | Probable antigen |
| Guinea pig | yidC | Q5L620 | 0.4299 | Probable antigen |

The antigenic protein yidC, a membrane protein insertase, is evolutionarily conserved and acts as a chaperone for folding and assembling membrane proteins (**33**). The yajC protein, known as a Sec translocon accessory complex subunit, forms part of a supercomplex required for protein insertion, membrane protein assembly, and secretion (**34**). The secY protein, the central subunit of the protein translocation pathway, facilitates the insertion of newly synthesized proteins into the inner membrane (**13**). CAB503 exhibits transferase activity, transferring aryl or alkyl groups from a donor to an acceptor compound (**13**). CAB746 also shows transferase activity, catalyzing the transfer of phosphate groups between compounds (**13**). Literature confirms that yidC, yajC, and secY, as membrane proteins, are critical for protein folding and membrane insertion, underscoring their therapeutic potential.

## Conservancy Analysis

The conservancy of shortlisted proteins (yidC, yajC, secY, CAB503, and CAB746) across all available strains was evaluated to assess their potential for broad-spectrum therapeutic applications. Protein sequences for yajC and yidC were retrieved from the UniProtKB database for two entries: *C. abortus* (species level) and *C. abortus* strain DSM 27085 / S26/3 (specific strain). Corresponding UniProt IDs for yajC were Q5L702 (species) and ABABDM16_3 (strain DSM 27085 / S26/3), and for yidC were Q5L620 (species) and ABABDM16_3 (strain DSM 27085 / S26/3). No strain-specific sequences were available for secY, CAB503, and CAB746. Pairwise alignment was performed using the EMBOSS Needle tool to assess sequence conservancy between species-level and strain-specific proteins. Results

indicated high conservancy of yajC and yidC proteins between *C. abortus* species and strain DSM 27085 / S26/3. Specifically, for yajC, two positions (amino acids 2 and 196) showed similarity indicated by dots (.), and position 103 showed high similarity indicated by a colon (:). For yidC, a colon was observed at position 136, with dots at positions 383, 395, 480, and 538. Additionally, a two-amino-acid deletion or insertion was noted at positions 793 and 794 in the strain DSM 27085 / S26/3 compared to the species reference, which may represent a mutation event (**Figure 2**). The dot (.) and colon (:) symbols indicate similarity and high similarity, respectively, according to EMBOSS Needle conventions. This analysis demonstrates strong sequence conservation of these proteins, supporting their suitability as broad-spectrum therapeutic targets.

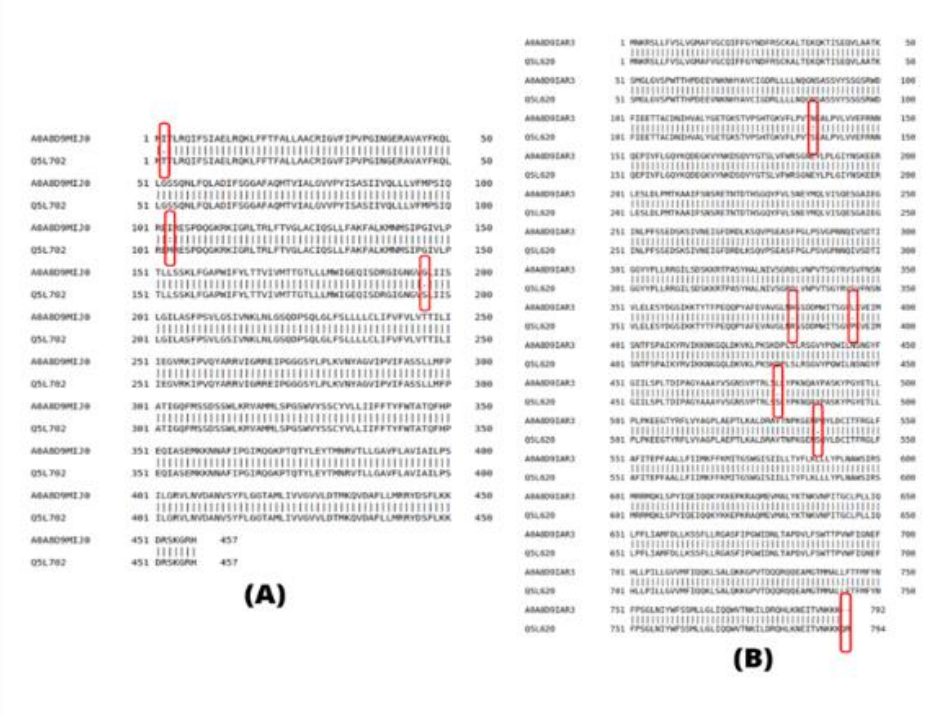


**Figure 2**. Pairwise alignment of yajC (A) and yidC (B) proteins between *C. abortus* (species level) and *C. abortus* strain DSM 27085 / S26/3 using EMBOSS Needle. **(A)** For yajC, red rectangles highlight regions of similarity, indicated by dots (".") at amino acid positions 2 and 196, and high similarity, indicated by a colon (":") at position 103. The remainder of the protein sequence is fully conserved between the species and the strain. **(B)** For yidC, the protein is nearly identical between species and strain. Red rectangles indicate high similarity (colon, ":") at position 136 and similarity (dots, ".") at positions 383, 395, 480, and 538. A two-amino-acid insertion/deletion (indel) is observed at positions 793–794, represented by a gap in the C. abortus sequence

## Domain Analysis of Protein Domains

Domain analysis through InterProScan was performed for the remaining proteins (secY, CAB503, and CAB746) because only one bacterial strain, *C. abortus* (strain DSM 27085 / S26/3), was available for these proteins. For this analysis, protein sequences were first retrieved from the UniProtKB database. These sequences were then submitted to InterProScan, which compares amino acid sequences against a nonredundant protein sequence database to identify protein families, domains, homologous superfamilies, and assign gene ontology terms (**35**). The proteins secY, CAB503, and CAB746 have lengths of 457, 299, and 268 amino acids, respectively. For secY, InterProScan assigned several protein families: IPR002208 (positions 1–456) and IPR026593 (positions 4–448) as shown in **Figure 3**. Other databases assigned the following: PIRSF004557 (1–456) by PIRSF, PF00344 (73–432) by Pfam, TIGR00967 (14–439) by TIGRFAMs, PTHR10906 (3–450) by PANTHER, and MF_01465 (4–448) by HAMAP. Homologous superfamilies identified by InterPro and CATH-Gene3D include IPR023201 and G3DSA:1.10, spanning amino acids 1–444. SUPERFAMILY assigned SSF103491 at positions 2–438. Two conserved sites were found by InterPro (IPR030659) at regions 73–92 and 181–198. PROSITE patterns (**36**) PS00756 and PS00755 were identified at positions 181–198 and 73–92, respectively. Gene ontology (GO) terms predicted for secY include the biological process of protein transport (GO:0015031) and cellular component membrane localization (GO:0016020). This domain analysis confirms secY's role as a membrane protein involved in protein transport.

InterProScan analysis for protein CAB503 depicted that protein families such as IPR006371, IPR039653, and IPR000537 were identified at regions 5–297, 39–294, and 47–256. Pfam assigned the protein family PF01040 to

CAB503 (47–256 amino acids). TIGR01475 was assigned to CAB503 (5–297 amino acids) by TIGRFAMs, and the PTHR11048 protein family was assigned to CAB503 (42–360 amino acids) by PANTHER. The protein family cd13959 was assigned to protein sequence CAB503 (39–294) by CDD. The homologous superfamilies IPR044878 and G3DSA:1.10.357.140 were identified at amino acid positions 11–166 by the InterPro homologous superfamily and CATH-Gene3D. The gene ontology terms such as molecular function, transferase activity (GO:0016765), and cellular component, that is, an integral component of membrane (GO:0016021), were predicted for protein CAB503. Transmembrane regions (TMhelix) were also identified, consistent with CAB503's membrane localization, as shown in **Figure 4**.

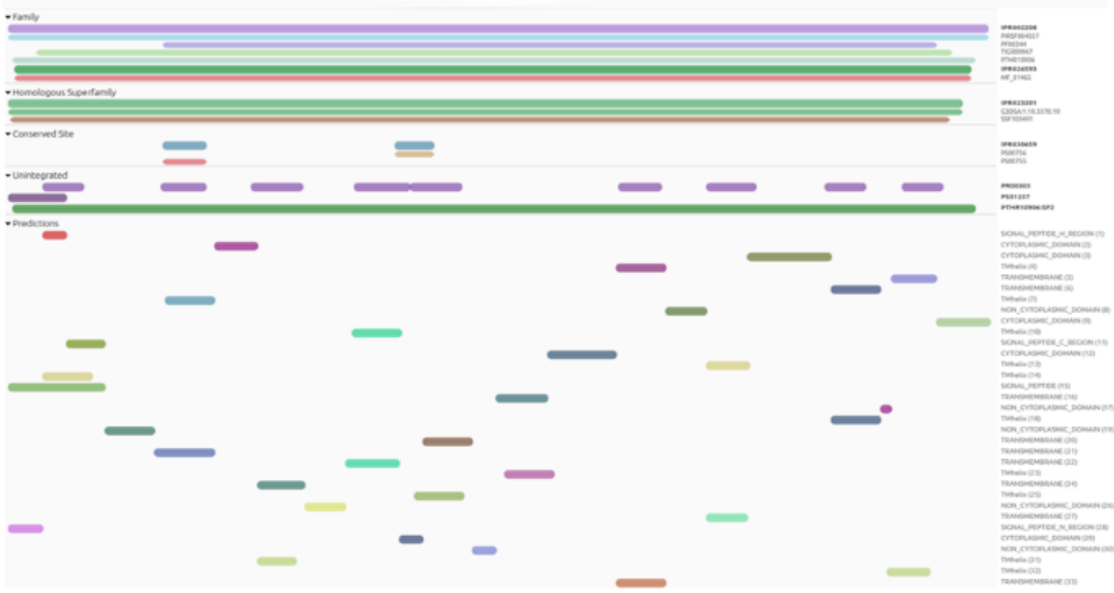


**Figure 3.** InterProScan analysis of the secY protein from *C. abortus* (strain DSM 27085 / S26/3), showing conserved regions and predicted domains along its 457-amino-acid sequence. Colored bars indicate protein family assignments (e.g., IPR002208 from amino acid 1 to 456, IPR026593 from 4 to 448), homologous superfamilies (e.g., IPR023201 from 1 to 444), conserved functional sites (e.g., IPR030659 at positions 73–92 and 181–198), and additional domain predictions (e.g., TIGR00967 from 14 to 439). Transmembrane helices (TMhelix) are also marked, confirming secY's role as a membrane protein

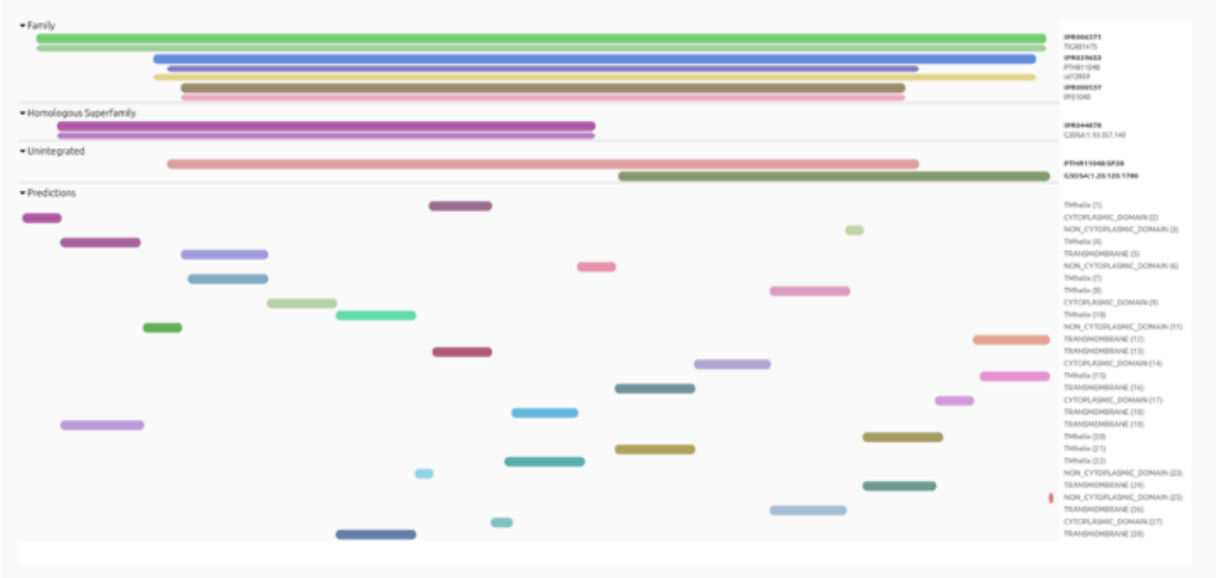


**Figure 4.** InterProScan analysis of conserved regions and domain predictions for the CAB503 protein of *C. abortus* (strain DSM 27085 / S26/3). Colored bars represent annotations across the 299-amino-acid sequence, including protein families (e.g., IPR006371, 5–297; IPR000537, 47–256), homologous superfamilies (e.g., IPR044878, 11–166), and unintegrated predictions (e.g., TIGR01475, 5–297). Transmembrane regions (TMhelix) are indicated, supporting CAB503's role as an integral membrane protein

InterProScan analysis of the protein CAB746, shown in **Figure 5**, identified protein family IPR000462 spanning amino acids 15–182. Pfam assigned the protein family PF01066 to CAB746 within the same region (15–182 amino acids). PROSITE pattern (**36**) PS00379 was detected at amino acids 66–88. Homologous superfamilies IPR043130 and G3DSA:1.20.120.1760 were found at positions 10–178 by InterPro homologous superfamily and CATH-Gene3D, respectively. Gene ontology (GO) terms for CAB746 included biological process, molecular function, and cellular component. The biological process was phospholipid biosynthetic process (GO:0008654), molecular function was phosphotransferase activity (GO:0016780), and the cellular component was membrane (GO:0016020).


**Figure 5.** InterProScan analysis of conserved regions and domain predictions for the CAB746 protein of *C. abortus* (strain DSM 27085 / S26/3). Colored bars represent annotations across the 268-amino-acid sequence, including protein families (e.g., IPR000462, 15–182), homologous superfamilies (e.g., IPR043130, 10–178), and conserved sites (e.g., PS00379, 66–88). Transmembrane regions (TMhelix) are indicated, supporting CAB746's role in membrane-associated processes

The subtractive genomics analysis identified two broadly conserved novel drug targets, yidC and yajC, as promising candidates for drug development and vaccines against *C. abortus* (strain DSM 27085 / S26/3), which causes the zoonotic disease chlamydiosis affecting hosts such as *Bos taurus*, *Cavia porcellus*, *Equus caballus*, *Mus musculus*, Oryctolagus cuniculus, and *Ovis aries*. In contrast, three novel antigenic proteins—secY, CAB503, and CAB746—were found to be unique to this bacterial strain. All five essential and novel drug targets (yidC, yajC, secY, CAB503, and CAB746) are involved in bacterial-specific metabolic pathways. Antitarget analysis confirmed no significant similarity between these drug targets and host proteins, reducing the risk of off-target effects and side effects (37). This suggests these targets could effectively reduce chlamydiosis while preserving host safety.

In conclusion, subtractive genomics effectively identified novel therapeutic targets against *C. abortus* (strain DSM 27085 / S26/3). The five shortlisted proteins are essential, non-homologous to host proteins, involved in unique bacterial metabolic pathways, and antigenic (VaxiJen score > 0.4) for most hosts. Conservancy analysis showed yidC and yajC are highly conserved across *C. abortus* strains, making them suitable for broad-spectrum therapeutics, whereas secY, CAB503, and CAB746 are strain-specific. Domain analysis revealed conserved domains and transmembrane regions, supporting their roles as membrane-associated proteins. These findings provide a strong foundation for future experimental validation through in vitro and in vivo studies, aiming to develop novel therapeutics and vaccines that can reduce the impact of chlamydiosis and improve animal health outcomes.

## Acknowledgements

N/A.

## Conflict of Interest

The authors declare no conflict of interest.

## Artificial Intelligent Declaration

The authors declare that they are responsible for the accuracy and integrity of all content of the manuscript, including part generated by AI, and it is not used as a co-author.

## References

1. Cohen S, Ho C. Review of rat (*Rattus norvegicus*), mouse (*Mus musculus*), guinea pig (*Cavia porcellus*), and rabbit (*Oryctolagus cuniculus*) indicators for welfare assessment. Animals. 2023;13(13):2167. https://doi.org/10.3390/ani13132167
2. Graham AL, Nussey DH, Lloyd-Smith JO, Longbottom D, Maley M, Pemberton JM, et al. Exposure to viral and bacterial pathogens among Soay sheep (*Ovis aries*) of the St Kilda archipelago. Epidemiol Infect. 2016;144(9):1879-1888. https://doi.org/10.1017/S0950268816000017
3. Selim A. *Chlamydophila abortus* infection in small ruminants: A review. Asian J Anim Vet Adv. 2016;11(10):587-593. https://doi.org/10.3923/ajava.2016.587.593
4. Essig A, Longbottom D. *Chlamydia abortus*: new aspects of infectious abortion in sheep and potential risk for pregnant women. Curr Clin Microbiol Rep. 2015;2:22-34. https://doi.org/10.1007/s40588-015-0014-2

5. Hughes L, Visser S, Heddema E, de Smet N, Linssen T, Wijdh RJ, et al. Zoonotic transmission of *Chlamydia felis* from domestic cats; a case series of chronic follicular conjunctivitis in humans. New Microbes New Infect. 2024;59:101412. https://doi.org/10.1016/j.nmni.2024.101412
6. Ortega N, Caro MR, Gallego MC, Murcia-Belmonte A, Álvarez D, Del Río L, et al. Isolation of *Chlamydia abortus* from a laboratory worker diagnosed with atypical pneumonia. Ir Vet J. 2015;69:8. https://doi.org/10.1186/s13620-016-0067-4
7. Al-Nuaimy WA, Al-Jandeel TJ. Immunological and molecular study of *Chlamydia trachomatis* as causative agent of abortion in Al-Muthanna province. Iraqi J Vet Med. 2018;42(1):99-104. https://doi.org/10.30539/iraqijvm.v42i1.38
8. Xia X. Bioinformatics and drug discovery. Curr Top Med Chem. 2017;17(15):1709-1726. https://doi.org/10.2174/1568026617666161116143440
9. Al-Kinani LH, Sharp MA, Wyatt KM, Coiacetto F, Sharp CR, Rossi G, et al. Haemoglobin epsilon as a biomarker for the molecular detection of canine lymphoma. Iraqi J Vet Med. 2023;47(1):21-27. https://doi.org/10.30539/ijvm.v47i1.1494
10. Chukwudozie OS, Duru VC, Ndiribe CC, Aborode AT, Oyebanji VO, Emikpe BO. The relevance of bioinformatics applications in the discovery of vaccine candidates and potential drugs for COVID-19 treatment. Bioinform Biol Insights. 2021;15:1-8. https://doi.org/10.1177/11779322211002168
11. Sudha R, Katiyar A, Katiyar P, Singh H, Prasad P. Identification of potential drug targets and vaccine candidates in *Clostridium botulinum* using subtractive genomics approach. Bioinformation. 2019;15(1):18. https://doi.org/10.6026/97320630015018
12. Goodall EC, Robinson A, Johnston IG, Jabbari S, Turner KA, Cunningham AF, et al. The essential genome of *Escherichia coli* K-12. MBio. 2018;9(1):10-128. https://doi.org/10.1128/mbio.02096-17
13. Thomson NR, Yeats C, Bell K, Holden MT, Bentley SD, Livingstone M, et al. The *Chlamydophila abortus* genome sequence reveals an array of variable proteins that contribute to interspecies variation. Genome Res. 2005;15(5):629-640. https://doi.org/10.1101/gr.3684805
14. Ijaq J, Chandrasekharan M, Poddar R, Bethi N, Sundararajan VS. Annotation and curation of uncharacterized proteins-challenges. Front Genet. 2015;6:119. https://doi.org/10.3389/fgene.2015.00119
15. Shen W, Le S, Li Y, Hu F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PloS One. 2016;11(10):e0163962. https://doi.org/10.1371/journal.pone.0163962
16. Martin FJ, Amode MR, Aneja A, Austine-Orimoloye O, Azov AG, Barnes I, et al. Ensembl 2023. Nucleic Acids Res. 2023;51(D1):D933-D941. https://doi.org/10.1093/nar/gkac958
17. Luo H, Lin Y, Liu T, Lai FL, Zhang CT, Gao F, et al. DEG 15, an update of the database of essential genes that includes built-in analysis tools. Nucleic Acids Res. 2021;49(D1):D677-D686. https://doi.org/10.1093/nar/gkaa917
18. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 2023;51(D1):D587-D592. https://doi.org/10.1093/nar/gkac963
19. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016;44(D1):D457-D462. https://doi.org/10.1093/nar/gkv1070
20. Miller SI, Salama NR. The gram-negative bacterial periplasm: Size matters. PLoS Biol. 2018;16(1):e2004935. https://doi.org/10.1371/journal.pbio.2004935
21. Yu CS, Chen YC, Lu CH, Hwang JK. Prediction of protein subcellular localization. Proteins:Struct Funct Bioinf. 2006;64(3):643-651. https://doi.org/10.1002/prot.21018
22. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics. 2010;26(13):1608-1615. https://doi.org/10.1093/bioinformatics/btq249
23. Knox, C., Wilson, M., Klinger, CM., Franklin, M., Oler E, Wilson A, et al. DrugBank 6.0: the DrugBank knowledgebase for 2024. Nucleic Acids Res. 2024;52(D1):D1265-D1275. https://doi.org/10.1093/nar/gkad976
24. Rao MS, Gupta R, Liguori MJ, Hu M, Huang X, Mantena SR, et al. Novel computational approach to predict off-target interactions for small molecules. Front Big Data. 2019;2:25. https://doi.org/10.3389/fdata.2019.00025
25. Singh J, Malik D, Raina A. Immuno-informatics approach for B-cell and T-cell epitope based peptide vaccine design against novel COVID-19 virus. Vaccine. 2021;39(7):1087-1095. https://doi.org/10.1016/j.vaccine.2021.01.011
26. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, et al. InterProScan: protein domains identifier. Nucleic Acids Res. 2005; 33(suppl_2): W116-W120. https://doi.org/10.1093/nar/gki442
27. Khan MT, Mahmud A, Iqbal A, Hoque SF, Hasan M. Hasan M. Subtractive genomics approach towards the identification of novel therapeutic targets against human *Bartonella bacilliformis*. Inform Med Unlocked. 2020;20:100385. https://doi.org/10.1016/j.imu.2020.100385
28. Gabler F, Nam SZ, Till S, Mirdita M, Steinegger M, Söding J, et al. Protein sequence analysis using the MPI bioinformatics toolkit. Curr Protoc Bioinformatics. 2020;72(1):e108. https://doi.org/10.1002/cpbi.108
29. Chitwood PJ, Hegde RS. The role of EMC during membrane protein biogenesis. Trends Cell Biol. 2019;29(5):371-384. https://doi.org/10.1016/j.tcb.2019.01.007
30. Nie D, Hu Y, Chen Z, Li M, Hou Z, Luo X, et al. Outer membrane protein A (OmpA) as a potential therapeutic target for *Acinetobacter baumannii* infection. J Biomed Sci. 2020;27:1-8. https://doi.org/10.1186/s12929-020-0617-7
31. Weiss DR, Karpiak J, Huang XP, Sassano MF, Lyu J, Roth BL, et al. Selectivity challenges in docking screens for GPCR targets and antitargets. J Med Chem. 2018;61(15):6830-6845. https://doi.org/10.1021/acs.jmedchem.8b00718
32. Smiline Girija AS. Delineating the immuno-dominant antigenic vaccine peptides against gacS-sensor kinase in *Acinetobacter baumannii*: an in silico investigational approach. Front Microbiol. 2020;11:2078. https://doi.org/10.3389/fmicb.2020.02078
33. Burland V, Plunkett III G, Daniels DL, Blattner FR. DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: organizational symmetry around the origin of replication. Genomics. 1993;16(3):551-561. https://doi.org/10.1006/geno.1993.1230
34. Komar J, Alvira S, Schulze RJ, Martin R, Lycklama a Nijeholt JA, Lee SC, Dafforn TR, Deckers-Hebestreit G, Berger I, Schaffitzel C, Collinson I. Membrane protein insertion and assembly by the bacterial holo-translocon SecYEG–SecDF–YajC–YidC. Biochem. J. 2016;473(19):3341-3354. https://doi.org/10.1042/BCJ20160545
35. Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, et al. The InterPro protein families and domains database: 20 years on. Nucleic Acids Res. 2021;49(D1):D344-D354. https://doi.org/10.1093/nar/gkaa977
36. Sigrist CJ, De Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, et al. New and continuing developments at PROSITE. Nucleic Acids Res. 2012;41(D1):D344-D347. https://doi.org/10.1093/nar/gks1067
37. Doller D, Wes P. Quality of Research Tools. Handb Exp Pharmacol. 2020; 257:119-145. https://doi.org/10.1007/978-3-030-33656-1

# اكتشاف حاسوبي لأهداف جديدة للأدوية واللقاحات في بكتيريا الكلاميديا المُجهضة من خلال الجينوميات الطرحية

**زينب عون علي [١]، سرى علي الأسدي[٢]، بتول حازم عبد الوهاب [٣*]، جميلة رامبوراولا [٤]**

**[١]فرع المختبرات الطبية، كلية الصحة والتقنيات الطبية، الجامعة التقنية الوسطى، بغداد، العراق، [٢]مركز بحوث التقانات الحيوية، جامعة النهرين، بغداد، العراق، [٣]فرع الأحياء المجهرية، كلية الطب، الجامعة المستنصرية، بغداد، العراق،[٤] فرع علم الوراثة،كلية أكسفورد للعلوم،بنغالور،الهند**

**الخلاصة**

الكلاميديا المُجهضة بكتيريا سلبية الغرام تُسبب داء الكلاميديا، وهو مرض حيواني المنشأ، يؤدي إلى الإجهاض وولادة جنين ميت في الماشية، ويشكل مخاطر صحية عامة جسيمة. مع تزايد مقاومة مضادات الميكروبات ومحدودية فعالية العلاج، تبرز حاجة ملحة إلى علاجات جديدة ومُستهدفة للتخفيف من آثارها الاقتصادية والحيوانية. اعتمدت هذه الدراسة على نهج الجينوميات الطرحية لتحليل كامل بروتيوم الكلاميديا المُجهضة (سلالة S26/3 / DSM 27085)، بهدف تحديد البروتينات الأساسية المتماثلة غير المُضيفة، والتي تُشارك في مسارات أيضية بكتيرية فريدة. من بين ٩٣٢ بروتينًا، تم اختيار خمسة بروتينات غشائية متكاملة، وهي YidC وYajC وSecY وCAB503 وCAB746، بناءً على أهميتها، وأدوارها الخاصة بمُسببات الأمراض، وعدم تشابهها مع المُضيف. أكد فحص مُضاد الهدف عدم وجود تشابه مع بروتينات المُضيف، مما يضمن الحد الأدنى من التأثيرات غير المُستهدفة. حدد التنبؤ بالمستضدات (VaxiJen > 0.4) سلالات YidC وYajC وSecY وCAB503 كمرشحات قوية للقاحات، بينما أظهر CAB746 مستضدات متفاوتة بين الأنواع. أظهر تحليل الحفظ أن سلالات YidC وYajC حافظتا على مستويات عالية من الحفظ في سلالات C. abortus، بينما أظهرت السلالات الأخرى اختلافات خاصة بكل سلالة. كشف تحليل النطاق عن أنماط محفوظة (مثل IPR002208 في SecY) ومناطق عبر غشائية، مما يدعم أهميتها البنيوية والمناعية. ختامًا، تُمثل سلالات YidC وYajC وSecY وCAB503 وCAB746 أهدافًا واعدة للأدوية واللقاحات. وتستدعي خصوصيتها البكتيرية ومناعتها وخصائص سلامتها إجراء المزيد من التحقق التجريبي لدعم التدخلات الموجهة ضد الكلاميديا.

**الكلمات المفاحية:** الكلاميديا المُجهضة، الجينوميات الطرحية، لقاح مُرشح، هدف مضاد للميكروبات، جينوميات مُسببات الأمراض البيطرية