

Article

In Silico Prediction of Secreted Proteins in Shiga Toxin-Producing *Escherichia coli*: Identification of a Hydrolase as a Robust Genomic Marker

María Victoria Vélez ¹, Ana Elisa Juárez ¹, Rocío Colello ¹, Felipe Del Canto ^{2,*} and Nora Lía Padola ^{1,*}

¹ Laboratorio de Inmunoquímica y Biotecnología, Centro de Investigación Veterinaria de Tandil (CIVETAN), UNCPBA-CICPBA-CONICET, Facultad de Ciencias Veterinarias, Tandil 7000, Argentina; mvictoriavelez@vet.unicen.edu.ar (M.V.V.); anajuarez@vet.unicen.edu.ar (A.E.J.); rocioc@vet.unicen.edu.ar (R.C.)

² Núcleo Interdisciplinario de Microbiología, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Santiago 8320000, Chile

* Correspondence: felipedelcanto@uchile.cl (F.D.C.); nlpadola@vet.unicen.edu.ar (N.L.P.)

Simple Summary

Shiga toxin-producing *Escherichia coli* is a foodborne bacterium that can cause severe illness in humans. Identifying harmful strains remains challenging due to their high genetic and biological diversity. In this study, we analyzed more than thirty-five thousand *Escherichia coli* genomes from public databases using computer-based methods to identify disease-associated factors. We focused on proteins secreted outside the bacterial cell as potential markers. One hundred fifty-five candidate genes were identified, and a hydrolase-encoding gene showed a strong association with pathogenic strains. This hydrolase could complement existing screening methods and support the development of improved strategies for detection and control.

Abstract

Shiga toxin-producing *Escherichia coli* (STEC) are major foodborne pathogens that can cause severe human disease. Identifying molecular markers associated with pathogenicity is essential to improve detection and to better understand virulence mechanisms. In this study, 35,828 *E. coli* genomes available in a public database were analyzed with the purpose of identifying STEC genes encoding proteins secreted to the outer membrane or into the extracellular space. The strains belong to six different sequence types (STs): ST442, ST297, ST1131, ST2217, ST2387, and ST2520. Comparative genomics allowed identification of 155 genes that were initially evaluated as potential virulence-associated candidates. Among these, a hydrolase-encoding gene showed a wide distribution and a significant association with STEC genomes versus non-STEC genomes (81.6% in the STEC group and 15.8% in the non-STEC group). The protein product is predicted to be released into the extracellular space and, according to examples in other bacterial pathogens, may contribute to STEC virulence. Altogether, our findings highlight this hydrolase as a relevant molecular marker that could complement current genetic screening schemes and support the design of novel control strategies against STEC infections.

Keywords: proteins; STEC; novel diagnostic; possible therapeutic targets; hydrolase



Academic Editor: Fabrizio Bertelloni

Received: 22 December 2025

Revised: 20 January 2026

Accepted: 30 January 2026

Published: 4 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Escherichia coli (*E. coli*) is a widespread bacterium that inhabits the intestinal tract of humans and food-producing animals as part of their normal microbiota [1,2]. Most *E. coli* strains are harmless commensals that coexist with their hosts in a mutualistic relationship and rarely cause disease [3]. However, this species displays remarkable genetic and phenotypic diversity, giving rise to pathogenic variants capable of causing a wide range of infections. Taking into account the virulence factors they harbor and the clinical manifestations they cause, Shiga toxin-producing *E. coli* (STEC) has gained relevance. STEC constitutes a zoonotic food-borne pathogen that causes diarrhea and Hemolytic Uremic Syndrome (HUS) [4,5]. Cattle are recognized as the principal reservoir of STEC, excreting the pathogen in their feces [6]. This ecological relationship allows STEC to persist in the bovine gastrointestinal tract as commensal organisms, facilitating their release into the environment through intermittent or continuous shedding [7,8]. Consequently, human infection can occur through multiple exposure pathways, such as the consumption of undercooked meat, contaminated water -of particular concern given its role in the contamination of fresh produce- as well as through direct contact with colonized animals or their environments [9,10].

Shiga Toxin (Stx) is the defining virulence factor of STEC. Stx is categorized into two main groups, *stx1* and *stx2*, which are further divided into numerous subtypes [11,12]. A single STEC strain may produce one or multiple Stx subtypes. Notably, strains producing Stx1a, Stx2a, and Stx2d are primarily associated with severe diarrhea and HUS [13,14].

Although Shiga toxins (Stx) are recognized as the principal virulence factor responsible for the pathogenicity of STEC, additional determinants are likely to contribute to the development of disease [15,16]. To exert their toxic effects, bacteria must first adhere to the host gastrointestinal tract [17]. The Locus of Enterocyte Effacement (LEE) encodes genes involved in intestinal epithelial adherence and the formation of attaching and effacing lesions [18]. Nevertheless, LEE-negative STEC strains have also been isolated from clinical cases, suggesting the existence of alternative and poorly understood adherence mechanisms, and little is known about their virulence strategies toward epithelial cells [19].

Additional virulence factors, such as outer membrane proteins (OMPs), display a broad range of functions, acting as adhesins for bacterial infection [20], passive diffusion and efflux channels [21], siderophore receptors [22], protein translocation pores [23], and enzymes [24,25]. Secreted and surface-exposed proteins represent optimal targets for vaccine and diagnostic applications, as they constitute the primary interface between the pathogen and the host immune system [26,27]. The identification and characterization of additional virulence factors are essential to elucidate STEC pathogenic mechanisms and may contribute to identifying effective strategies to prevent transmission [28]. Various studies have reported successful approaches to reduce STEC shedding and spread, including vaccination of cattle using specific antigens derived from LEE-encoded virulence factors [27]. However, their limited coverage underscores the need for alternative antigens that can provide broader protection across STEC, which can protect against a broader range of serotypes. To this end, our objectives were to detect and characterize proteins in silico and to evaluate their potential as specific markers for STEC identification.

2. Materials and Methods

2.1. The Strains Selections

The bacterial strains were selected from the strain collection of the Laboratory of Immunochemistry and Biotechnology, FCV, UNCPBA, Argentina, for which whole-genome sequence data are available [28]. Genomes of 11 LEE-negative STEC strains were analyzed. The bacteria were isolated from cattle, and they harbored *stx* genes, previously confirmed by

PCR [7]. Moreover, the strains analyzed belong to six different sequence types (STs): ST442, ST297, ST1131, ST2217, ST2387, and ST2520 [28]. Genomes of non-pathogenic/commensal strains available in databases were used as controls: *E. coli* HS (NCBI Assembly Ref-seq accession code GCF_000017765.1), *E. coli* DH5 α (GCF_002899475.1), *E. coli* K-12 (GCF_000005845.2), *E. coli* MG1655 (GCF_000005845.2), *E. coli* K-12 (GCF_000019425.1), *E. coli* DH10B (GCF_000019425.1), and *E. coli* BL21 (GCF_000022665.1).

2.2. Comparative Analysis: Development of the Study Matrix

A total of 35,828 *E. coli* genomes were obtained from the Assembly Refseq database available in the National Center for Biotechnology Information (NCBI) of the United States of America (<https://www.ncbi.nlm.nih.gov/genbank/>, accessed on 8 January 2024) [29]. *E. coli* strains were classified as STEC or non-STEC group based on the presence or absence of any of the genes encoding Shiga toxin subunits (genes encoding A and B subunits of variants *stx1a*, *stx1c*, *stx2a*, *stx2b*, *stx2c*, *stx2d*, *stx2e*, *stx2f*, and *stx2g* were included). Non-STEC includes other pathotypes different from STEC.

The STs corresponding to each strain were identified using MLST [30], and genomes associated with the six above-mentioned STs were retrieved from the database. BUSCA (<https://busca.biocomp.unibo.it/>, accessed on 1 September 2024) was employed to predict protein subcellular localization specific to STEC [31]. A comparative genomic analysis based on the genetic content was performed using the Large-scale blast score ratio tool (LS-BSR) (available at <https://github.com/jasonsahl/LS-BSR>, accessed on 8 January 2024) [32]. Thus, the genetic content of all STEC strain genomes was compared to the genetic content of non-pathogenic *E. coli* strains *E. coli* HS, *E. coli* DH5 α , *E. coli* K-12 MG1655, *E. coli* K-12 DH10B, and *E. coli* BL21. For each gene, genomic sequences with BSR ≥ 0.8 were considered positive, and those with BSR < 0.4 were considered negative [32] (Supplementary Table S3). The LS-BSR pipeline generated a binary presence/absence matrix for 35,828 complete genomes. Detailed strain information and accession numbers are provided in the Supplemental (Supplementary Table S1).

2.3. Selection of Proteins of Study

Based on the matrix generated in Table S3, the fifteen most prevalent genes encoding uncharacterized proteins were selected. These genes were absent from commensal *E. coli* strains (BSR < 0.4) and were prioritized for evaluating their potential application as diagnostic markers.

To characterize the selected targets as proteins and to confirm that they encode conserved, structurally coherent proteins suitable as discriminatory genomic markers, several bioinformatic tools were used to analyze their sequences and predicted structures.

Functional annotation and structural prediction were performed through a consensus pipeline integrating multiple publicly available databases and web-based tools. Sequence similarity was evaluated by BLASTP 2.17.0+ (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome, accessed on 4 November 2024) against the UniProt (<https://www.uniprot.org/blast>, accessed on 4 November 2024) and NCBI nr databases to identify homologs and infer preliminary functions [33,34]. Conserved domains and protein families were characterized using InterProScan (<https://www.ebi.ac.uk/interpro>, accessed on 3 January 2025), CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>, accessed on 15 September 2025), and HHpred (<https://toolkit.tuebingen.mpg.de/tools/hhpred> accessed on 20 September 2025) to define domain composition and superfamily relationships [35,36].

TMHMM 2.0 (<https://services.healthtech.dtu.dk/services/TMHMM-2.0/>, accessed on 16 September 2025) was used to predict transmembrane helices [37], and AlphaFold

(<https://alphafold.ebi.ac.uk/>, accessed on 15 July 2025) to generate three-dimensional structural models supporting functional inference [38]. Protein–protein interaction networks were analyzed with STRING (<https://string-db.org/>, accessed on 16 July 2025) to explore functional associations based on genomic context, co-expression, and experimental evidence [39].

Expasy’s ProtParam server (<https://web.expasy.org/protparam/>, accessed on 16 July 2025) to calculate physicochemical parameters, including molecular weight, theoretical pI, charged residues, molecular formula, and instability index [40]. Finally, B-cell epitopes were predicted using the IEDB B-cell Epitope Prediction tool (<https://tools.iedb.org/bcell/>, accessed on 1 September 2025). Specifically, the BepiPred-2.0 algorithm was applied with default parameters [41]. This enables us to evaluate the potential of the candidate as a possible diagnostic target.

2.4. Statistical Analysis

The association between group membership (STEC vs. non-STEC) and the presence/absence of each of the most prevalent genes that codified proteins was assessed using Pearson’s Chi-square test (χ^2). Yates’ correction was applied where appropriate. Results were considered statistically significant at a level of a *p*-value of <0.05. Additionally, the strength of the association was quantified by calculating the Odds Ratio (OR) with its corresponding 95% confidence interval (CI). The statistical analyses were performed using Graphpad and MedCalc (<https://www.graphpad.com/quickcalcs/contingency1/>, accessed on 1 November 2024 and https://www.medcalc.org/en/calc/odds_ratio.php, accessed on 1 January 2026, respectively). To evaluate the discriminative ability between STEC and non-STEC strains, a Receiver Operating Characteristic (ROC) curve analysis was performed. This analysis was used to determine the Area Under the Curve (AUC), along with its 95% confidence interval (CI), as a measure of overall predictive performance. The optimal cut-off point was determined by maximizing both sensitivity and specificity. Analyses were performed using Roc Analysis software (<https://englab.rad.jhmi.edu/jrocfits/JROCFITi.html>, accessed on 1 December 2025).

3. Results

Among the 35,828 RefSeq *E. coli* genomes, 247 were identified as ST297 (107 genomes), ST442 (100 genomes), ST2520 (22 genomes), ST2217 (9 genomes), ST2387 (7 genomes) or ST1131 (2 genomes), representing STs found in STEC non-O157 strains isolated in Argentina. After a comparison of the genetic content against non-pathogenic *E. coli* genomes, 947 STEC unique genes were identified, from which 113 encoding putative outer membrane or extracellular proteins were selected. These genes were screened in the whole database (35,828 genomes), which included 4980 STEC genomes (14%) and 30,848 as non-STEC. Among the most prevalent protein-coding genes within the STEC group, the most prevalent genes, or their encoded proteins, were selected for further analysis. A total of 15 were chosen for an in-depth characterization (Table 1).

Table 1. Candidate protein-coding genes evaluated as potential STEC markers: localization, prevalence, and statistical validation.

N°	NCBI Accession	Length (aa)	Location	Functional Name	Prevalence	χ^2	OR (95% CI), <i>p</i> Value	Fitted ROC Area
1	WP_136816566.1	155	plasma membrane	Prepilin peptidase	4449/4980	*-	-	-
2	WP_108919521.1	696	extracellular space	Bifunctional siderophore receptor/adhesin Iha	3401/4980	$\chi^2 = 4885.538$	8.24 (7.71 to 8.80), <i>p</i> < 0.0001	0.801
3	AWM66617.1	105	outer membrane	Hypothetical protein	4464/4980	$\chi^2 = 11,494.595$	138,964.28 (8685.74 to 2,223,307.61), <i>p</i> < 0.0001	0.883

Table 1. Cont.

N°	NCBI Accession	Length (aa)	Location	Functional Name	Prevalence	χ^2	OR (95% CI), <i>p</i> Value	Fitted ROC Area
4	WP_085453571.1	494	extracellular space	Hydrolase	4077/4980	$\chi^2 = 5454.483$	23.81 (22.02–25.75), <i>p</i> < 0.0001	0.902
5	WP_308733525.1	243	extracellular space	TraT complement resistance	3916/4980	$\chi^2 = 6260.662$	2.71 (2.53–2.92), <i>p</i> < 0.0001	0.515
6	RDP83843.1	59	outer membrane	Probable site-specific recombinase/integrase	3488/4980	-	-	-
7	WP_407235464.1	190	extracellular space	Fimbrial protein	2648/4980	$\chi^2 = 411.148$	1.85 (1.74–1.97), <i>p</i> < 0.0001	0.623
8	WP_100008035.1	243	extracellular space	Putative lipoprotein	4192/4980	$\chi^2 = 626.068$	2.68 (2.48–2.90), <i>p</i> < 0.0001	0.811
9	OEI27840.1	93	outer membrane	Hypothetical protein	3092/4980	$\chi^2 = 3121.57$	10.48 (9.81 to 11.19), <i>p</i> < 0.0001	0.801
10	WP_283570181.1	59	extracellular space	Hypothetical protein	3300/4980	$\chi^2 = 14,633.726$	43.66 (40.29 to 47.32), <i>p</i> < 0.0001	0.95
11	WP_001470661.1	63	outer membrane	DUF4222_uncharacterized	3623/4980	$\chi^2 = 6392.410$	11.44 (10.68–12.25), <i>p</i> < 0.0001	0.84
12	WP_210177975.1	89	extracellular space	Shiga-like toxin, beta subunit	3451/4980	-	-	-
13	WP_097746304.1	87	extracellular space	Shiga toxin Stx2b subunit B	3449/4980	-	-	-
14	SQQ40187.1	83	outer membrane	Putative prophage protein/hypothetical	3641/4980	-	-	-
15	WP_249926694.1	472	extracellular space	A β -hydrolase_1 (putative functional module)	4068/4980	$\chi^2 = 9924.587$, <i>p</i> < 0.0001	23.72 (21.94–25.65), <i>p</i> < 0.0001	0.906

NCBI accession: *RefSeq* or *GenBank* accession number corresponding to the protein sequence in public databases. Length (aa) Location: predicted subcellular localization (e.g., extracellular space, outer membrane, plasma membrane), determined using BUSCA and complementary tools. Predicted function: functional annotation inferred from BLAST, InterProScan, UniProt, and HHPred analyses (e.g., hydrolase, lipoprotein, fimbrial protein, complement resistance factor). Prevalence: number of STEC genomes (out of 4980) in which the gene was detected. χ^2 : Chi-square statistic evaluating the association between gene presence and STEC status (all *p* < 0.0001). OR: Odds Ratio; 95% CI, 95% Confidence Interval. Fitted ROC area: area under the ROC curve estimating the discriminatory capacity of each gene to differentiate STEC genomes (AUC closer to 1 indicates better performance). *- Statistical analysis was not performed.

From the initial set of candidates, five were excluded from the statistical analysis. Specifically, one showed cytoplasmic localization (number 1), two encoded Shiga toxin (numbers 12 and 13), one encoded an integrase (number 6), and one was of phage-related origin (number 14) (Table 1) (Supplementary Table S2). The ten remaining proteins showed significant association with STEC strains (*p* < 0.0001). However, both the magnitude of association (χ^2) and the discriminative power (ROC AUC) varied among them (Table 1).

Three genes exhibited the strongest associations and the highest discriminatory capacities to detect STEC genomes, suggesting their potential as robust markers. One encoding a hypothetical protein (number 10: NCBI protein accession code WP283570181.1, $\chi^2 = 14,633.7$; OR: 43.66; AUC = 0.95), and two encoding putative uncharacterized hydrolases (number 4: NCBI protein accession WP_085453571.1; $\chi^2 = 5454.483$, OR: 23.81; AUC = 0.902 and number 15 NCBI protein accession code WP_249926694.1 $\chi^2 = 9924.6$; OR: 23.72; AUC = 0.906) (Table 1). Although protein number 10 showed the strongest statistical signal, it encodes a hypothetical protein, which makes it unsuitable as a biomarker. A blastn-based screening showed that genes encoding both hydrolases (4 and 10, Table 1) hit the same locus in a set of representative genomes, indicating that they were identified separately, but the gene is present only in single copies per genome. In addition, considering that proteins 4 and 15 share 81% identity and 88% sequence similarity, only one was prioritized as a potential marker. Protein 15 was ultimately selected for in-depth study based on its superior predictive performance, as indicated by a higher AUC value.

The hydrolase identified in this study showed a prevalence of 81.6% within the STEC group, while appearing in only 15.8% of non-STEC genomes. The distribution of genetic markers within the STEC group was evaluated, focusing on *eae* and *hes* as representative markers for LEE-positive and LEE-negative strains, respectively [19,42]. The *eae* gene and its variants (α , β , γ , δ , ε) were identified between 50.8% to 65.2% of the STEC group, while the *hes* gene was detected in 16.9%.

The hydrolase identified is composed of 472 amino acids. The prediction of the localization by BUSCA indicated that the extracellular space and its approximate molecular mass are ~52.19 kDa (Supplementary Table S2).

The sequence was analyzed using BLAST, UniProt, InterProScan, HHPred, CDD, STRING, and AlphaFold. These analyses confirmed that the protein belongs to the Hydrolase family (Supplementary Table S2). InterProScan analysis identified an α/β -hydrolase fold (IPR029058), a highly conserved structural core common to hydrolytic enzymes of diverse functions, suggesting the protein may indeed act as a hydrolase. Epitope prediction analyses indicated the presence of several regions with the potential to act as B-cell epitopes (Figure 1). Among them, the peptide spanning residues 364 to 376 showed the highest predicted score (8.401), making it the most promising B-cell epitope candidate (Table 2).

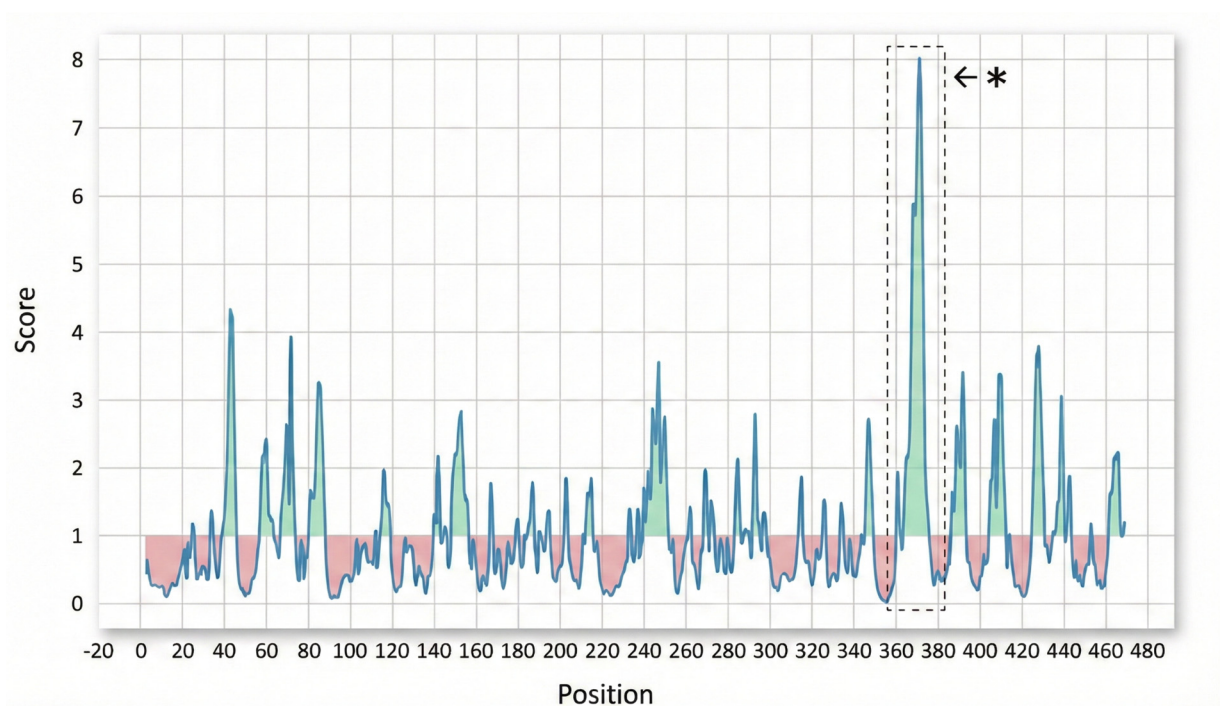


Figure 1. The graph illustrates the prediction score distribution across the amino acid sequence (X-axis: Position). Regions showing high peaks (green and pink spikes rising above the mean line) correspond to areas of the protein with the highest probability of being B-cell epitopes. * Highest-scoring epitopes.

Table 2. Candidate B-cell epitopes predicted from the protein sequence.

Start	End	Peptide	Length	Score
35	40	QERRPD	6	1.833
60	68	KYYTDTS	9	1.968
85	93	TLKWWKS	9	1.944
113	119	SGLYQVA	7	1.94
143	148	WETRDSRY	6	1.98
202	208	QLKAQPD	7	1.921

Table 2. Cont.

Start	End	Peptide	Length	Score
240	252	QLTEPDSRKDA	13	1.916
364	376	SLTGEEDSSKLDA	13	8.401
387	393	DRKGSQVS	7	1.838
402	408	QRGQSYS	7	1.86
420	426	QQSYTS	7	1.871
451	457	QEPSPSL	7	1.885

Start: N-terminal residue position of the predicted peptide segment in the full protein sequence. End: C-terminal residue position of the predicted peptide segment in the full protein sequence. Peptide: The amino acid sequence of the predicted epitope candidate. Length: The number of amino acid residues in the predicted epitope segment (End – Start + 1). Score: The aggregated prediction score for the peptide segment, representing the confidence level of its potential antigenicity. Higher scores indicate a greater likelihood of being a functional B-cell epitope.

Moreover, structural modeling using AlphaFold analysis revealed a 99% sequence identity (E-value = 0) with the α/β -hydrolase structure AF-A0A085NW94-F1-v6, confirming the conservation of the α/β -hydrolase fold and supporting the predicted hydrolase function of WP_249926694.1 (Figure 2). Furthermore, STRING analysis revealed that WP_249926694.1 is functionally associated with several extracellular or membrane-related proteins, including OmpT, AidB, and ApaH, among others. The functional network (Figure 1) shows that this hydrolase is part of a conserved functional module; its connectivity and co-occurrence suggest that it is a stable component of the STEC proteome.

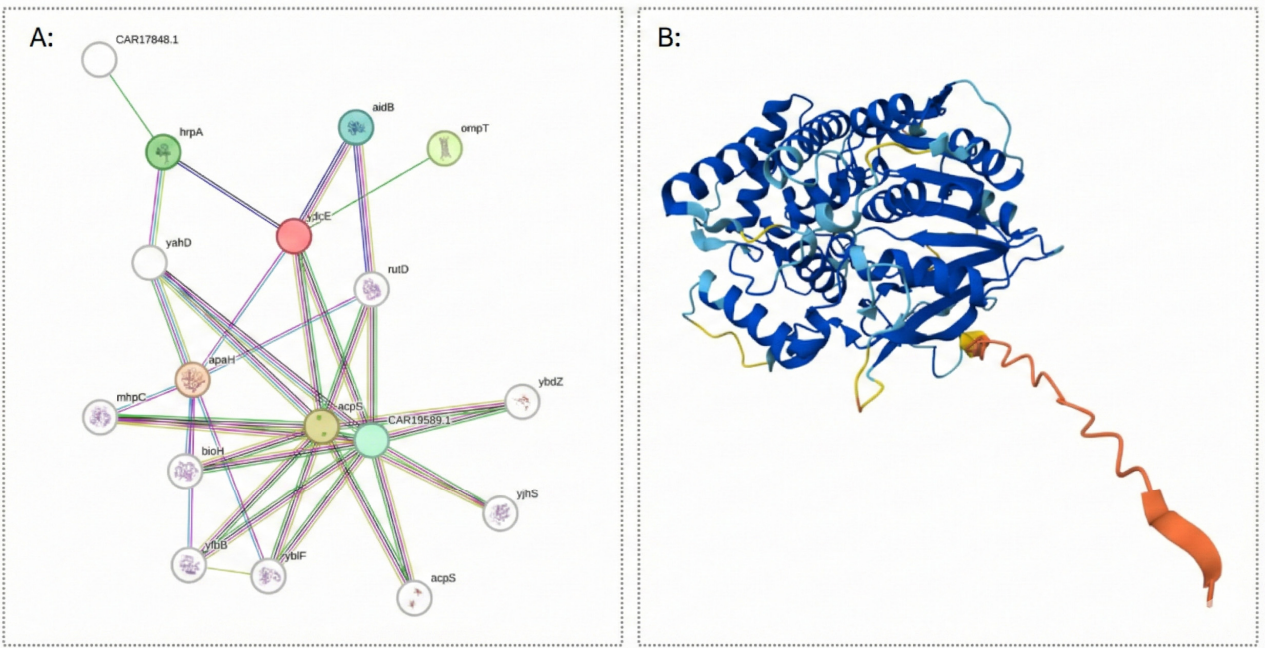


Figure 2. (A). STRING interaction network of the WP_249926694.1 protein. STRING analysis of the *E. coli* protein YlcE shows its predicted interaction partners. YlcE (red node) is connected to proteins of known or predicted function, including AidB, OmpT, ApaH, HrpA, AcpP. (B). Predicted three-dimensional structure of WP_249926694.1 based on its closest homolog (AF-A0A085NW94-F1-v6) obtained from AlphaFold. The color scale represents the confidence of the prediction, with blue indicating high-confidence regions and red representing lower-confidence regions.

4. Discussion

Reducing the zoonotic risk associated with STEC would significantly improve public health by enhancing food safety. Considering both the zoonotic and economic impact of STEC infections, along with the increasing number of cases worldwide [4], emphasizes

the importance of diagnostic tools capable of detecting a broad range of STEC strains [43]. Although Shiga toxins (Stx) are the main virulence determinants in STEC, other virulence factors contribute to their pathogenicity. However, many of these cannot yet be considered reliable markers because they are not consistently present across the extensive genomic diversity of STEC (e.g., *eae*, *saa*, *ehxA*, *hes*) [19,42,44]. As a result, defining a STEC strain as pathogenic is complex/challenging since no combination of markers predicts the potential of a STEC strain to cause human disease. In LEE-positive strains, *eae* is a well-established predictor to detect pathogenic strains [42]. In contrast, for LEE-negative strains, various pathogenicity islands and virulence-associated genes have been described; however, their prevalence is generally below 54%, limiting their diagnostic usefulness (e.g., LAA, *hes*) [19,28].

In this study, the final selection of candidates was guided not only by their differential prevalence between STEC and non-STEC genomes, but also by biological criteria relevant to their potential diagnostic application. Accordingly, proteins predicted to be cytoplasmic were excluded due to their limited accessibility, which reduces their suitability as diagnostic targets [27]. Shiga toxin subunits were also excluded to avoid redundancy with well-established STEC markers and to prioritize the identification of alternative targets that may complement existing diagnostic approaches. In addition, proteins of phage origin or annotated as hypothetical in public databases were excluded because of their uncertain functional relevance, limited annotation support, and potential variability across strains, which could compromise their robustness and specificity as diagnostic markers. Collectively, these selection criteria allowed the analysis to focus on proteins with greater biological coherence and higher potential as reliable diagnostic markers for STEC.

A comprehensive in silico analysis identified two putative hydrolase genes, strongly associated with STEC strains. The statistical association observed, characterized by odds ratios of 23.72 and 23.81, and a χ^2 value of 9924.59 and 5454.483 ($p < 0.0001$), demonstrates that the presence of these genes is a highly reliable indicator of the STEC pathotype. Although the hydrolase-encoding genes were identified separately, single hits were detected per positive genome, indicating that each strain harbors a single copy. These results, along with the high discriminatory power demonstrated in the ROC analysis, establish this protein as a reliable diagnostic target in STEC. Notably, this protein was present in 81.6% of STEC genomes, a statistically significant frequency compared to other potential markers analyzed in this study (such as *eae* and *hes*), regardless of the presence or absence of the LEE locus. When Galarce et al. [45] analyzed STEC genomes from South America, they reported marked variability in the prevalence of classical virulence-associated markers, with *eae* detected in only 30.8% of strains, *ehxA* in 60%, *saa* in 40.8%, and the LAA pathogenicity island in 43.1%. In this context, the markedly higher distribution of the hydrolase identified in the present study across diverse STEC lineages underscores its potential relevance as a broadly represented genomic marker.

The presence of the hydrolase was analyzed in the non-STEC group, where it was identified in 15.8% of the sequences. These results may be attributed to the significant genomic plasticity of *E. coli* strains, which is reflected in their highly diverse virulence profiles [46]. Although the hydrolase is not exclusively restricted to STEC, this does not diminish our findings, as the detection of this candidate marker is proposed as a complementary tool alongside detection of other genetic markers. The gene that codifies this hydrolase could be incorporated into existing multiplex PCR systems that differentiate *stx1* (180 bp), *stx2* (255 bp), and *eae* (384 bp) by amplicon length [42]. Preliminary primer designs (~100 bp) have already been generated for this purpose (data not published), which could facilitate the inclusion of its gene in epidemiological fast surveillance. Moreover, the hydrolase was detected across the six STs analyzed, indicating that the identified protein is not specific to a single lineage but is instead present in multiple genetic backgrounds within

STEC. Functional and structural predictions obtained through multiple bioinformatic tools indicated that this protein belongs to the Hydrolase family, suggesting an enzymatic role related to bacterial metabolism or host interaction. The hydrolase class encompasses a wide range of enzymes that catalyze hydrolytic reactions using water to break chemical bonds, including proteases, lipases/phospholipases, glycosidases, and esterases [47].

The analysis performed with AlphaFold indicated a 99% structural association between WP_249926694.1 and the α/β -hydrolase AF-A0A085NW94-F1-v6, confirming the prediction obtained with the other software tools. Previous studies have reported that proteins belonging to the α/β -hydrolase family share a highly conserved, simple, and stable structural fold that supports the formation of diverse catalytic triads capable of mediating hydrolytic reactions [48]. While the general function, hydrolysis catalysis, can be inferred from the fold architecture and the presence of the catalytic triad, substrate specificity cannot be reliably predicted through in silico analysis alone. This limitation arises because minimal variations in the residues of the substrate-binding pocket can profoundly alter specificity. Therefore, accurate identification of the physiological substrate requires experimental validation or, alternatively, the analysis of conserved structural motifs within the binding pocket [48,49]. Moreover, STRING analysis associated the potential marker with several extracellular or membrane-related proteins, including OmpT, AidB and ApaH, which could be implicated in host interaction, stress responses, and virulence processes (Figure 1) [25,47,50,51].

Previous studies have demonstrated the role of hydrolases as virulence factors in other pathogens such as *Clostridium perfringens*, *Streptococcus pyogenes*, and *Staphylococcus aureus* [52–54]. The identified hydrolase in STEC could therefore be involved in virulence mechanisms, although its role remains to be experimentally validated.

The high-scoring epitope (364–376), together with other predicted epitopic regions identified in this protein, represents a promising target for the development of peptide-based or recombinant vaccines. The presence of multiple B-cell epitope candidates further suggests that combining these peptides may broaden immune coverage and enhance vaccine efficacy [55]. In addition to its potential as a vaccine antigen, the strong predicted humoral immunogenicity of these epitopic regions indicates that this protein could also serve as a valuable marker for serological assay development, improving both the sensitivity and specificity of current diagnostic approaches.

Despite the robustness of our comparative genomics approach, this study has certain limitations. First, as an exclusively in silico analysis, our findings regarding the expression, extracellular localization, and enzymatic activity of the candidate hydrolase require further experimental validation through in vitro assays. While the identified B-cell epitopes show high prediction scores, their actual immunogenicity and ability to elicit a protective response remain to be characterized. Additionally, the presence of the hydrolase gene in a fraction of non-STEC genomes suggests that, while it is a highly prevalent marker within STEC, its use in diagnostic frameworks should be complemented with traditional markers, such as Shiga toxin genes, to ensure maximum specificity. Future investigations should focus on characterizing its biochemical activity, confirming its expression in vitro, PCR validation, protein expression, and assessing the immunogenic potential of its predicted epitopic regions.

This study identifies, for the first time, a highly prevalent and previously uncharacterized extracellular hydrolase as a robust STEC marker with significant diagnostic and vaccine potential. Our findings offer a more comprehensive approach for identifying STEC strains in complex epidemiological scenarios and lay the groundwork for future experimental studies to characterize their biological role and validate their utility in public health surveillance and disease prevention.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/vetsci13020153/s1>. Table S1: Comprehensive genomic dataset and presence/absence matrix of virulence factors across 35,828 *Escherichia coli* strains. Table S2: Comprehensive bioinformatic profiling of candidate markers: results from functional, structural, and localization analysis tools. Table S3: Specificity cross-screening: sequence identity of candidate markers against commensal and laboratory *Escherichia coli* reference strains.

Author Contributions: Conceptualization, M.V.V., A.E.J., F.D.C. and N.L.P.; Methodology M.V.V., A.E.J. and F.D.C.; Software F.D.C.; Validation, F.D.C.; Formal Analysis, A.E.J. and F.D.C.; Investigation M.V.V., R.C. and, N.L.P.; Data Curation, A.E.J., R.C., N.L.P. and F.D.C.; Writing—Original Draft Preparation, M.V.V.; Writing—Review & Editing, A.E.J., F.D.C. and N.L.P.; Visualization, R.C.; Supervision, N.L.P.; Project Administration, M.V.V.; Funding Acquisition, R.C., F.D.C. and N.L.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by PICT-2021-I-00178, CIC and SECAT from Argentina, and FONDECYT Link Project ENL13/24, Vice-Rectorate for Research and Development, University of Chile.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

<i>E. coli</i>	<i>Escherichia coli</i>
STEC	Shiga toxin-producing <i>Escherichia coli</i>
Stx	Shiga toxin
LEE	Locus of Enterocyte Effacement
HUS	Hemolytic Uremic Syndrome
OMP(s)	Outer membrane protein(s)
ST(s)	Sequence type(s)
MLST	Multilocus sequence typing
LS-BSR	Large-scale BLAST score ratio
BSR	BLAST score ratio
ROC	Receiver Operating Characteristic
AUC	Area under the curve
CI	Confidence interval
χ^2	Chi-square test
TMHMM	Transmembrane Hidden Markov Model

References

1. Shoaib, M.; Hameed, M.F.; Aqib, A.I.; Wang, W.; Wang, Q.; Wang, S.; Pu, W. Emerging Threat of Antimicrobial Resistance Determinants and Plasmid Replicon Types Acquisition by *Escherichia coli* of Poultry and Other Food-Producing Animal Origin in China: Local Findings with Global Implications. *Poult. Sci.* **2025**, *104*, 106055. [[CrossRef](#)]
2. Govindarajan, D.K.; Viswalingam, N.; Meganathan, Y.; Kandaswamy, K. Adherence Patterns of *Escherichia coli* in the Intestine and Its Role in Pathogenesis. *Med. Microecol.* **2020**, *5*, 100025. [[CrossRef](#)]
3. Leimbach, A.; Hacker, J.; Dobrindt, U. *E. coli* as an All-Rounder: The Thin Line Between Commensalism and Pathogenicity. *Curr. Top. Microbiol. Immunol.* **2013**, *358*, 3–32. [[CrossRef](#)]
4. Rahman, A.; Chowdhury, M.S.R.; Hossain, H.; Elsaid, F.G.; Almutairi, L.A.; Begum, R.; Sabrin, M.S.; Akanda, M.R.; Hossain, M.M.; Islam, M.R.; et al. Identification of Virulence Genes and Multidrug Resistance in Shiga-Toxin Producing *Escherichia coli* (STEC) from Migratory and Captive Wild Birds. *Pak. Vet. J.* **2024**, *44*, 1120–1130. [[CrossRef](#)]

5. Alhadlaq, M.A.; Aljurayyad, O.I.; Almansour, A.; Al-Akeel, S.I.; Alzahrani, K.O.; Alsalman, S.A.; Yahya, R.; Al-Hindi, R.R.; Hakami, M.A.; Alshahrani, S.D.; et al. Overview of Pathogenic *Escherichia coli*, with a Focus on Shiga Toxin-Producing Serotypes, Global Outbreaks (1982–2024) and Food Safety Criteria. *Gut Pathog.* **2024**, *16*, 57. [\[CrossRef\]](#)
6. Fernández, D.; Padola, N.L. *Escherichia coli* Verocitotoxigénico: Varias Cuestiones... y Los Tambos También. *Rev. Argent. Microbiol.* **2012**, *44*, 312–323.
7. Padola, N.L.; Sanz, M.E.; Blanco, J.E.; Blanco, M.; Blanco, J.; Etcheverría, A.I.; Arroyo, G.H.; Usera, M.A.; Parma, A.E. Serotypes and Virulence Genes of Bovine Shigatoxigenic *Escherichia coli* (STEC) Isolated from a Feedlot in Argentina. *Vet. Microbiol.* **2004**, *100*, 3–9. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Montero, D.; Orellana, P.; Gutiérrez, D.; Araya, D.; Salazar, J.C.; Prado, V.; Oñate, Á.; del Canto, F.; Vidal, R. Immunoproteomic Analysis to Identify Shiga Toxin-Producing *Escherichia coli* Outer Membrane Proteins Expressed during Human Infection. *Infect. Immun.* **2014**, *82*, 4767–4777. [\[CrossRef\]](#)
9. Colello, R.; Cáceres, M.E.; Ruiz, M.J.; Sanz, M.; Etcheverría, A.I.; Padola, N.L. From Farm to Table: Follow-up of Shiga Toxin-Producing *Escherichia coli* throughout the Pork Production Chain in Argentina. *Front. Microbiol.* **2016**, *7*, 93. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Angel Villegas, N.; Baronetti, J.; Albesa, I.; Polifroni, R.; Parma, A.; Etcheverría, A.; Becerra, M.; Padola, N.; Paraje, M. Relevance of Biofilms in the Pathogenesis of Shiga-Toxin-Producing *Escherichia coli* Infection. *Sci. World J.* **2013**, *2013*, 607258. [\[CrossRef\]](#) [\[PubMed\]](#)
11. EFSA BIOHAZ Panel; Koutsoumanis, K.; Allende, A.; Alvarez-Ordóñez, A.; Bover-Cid, S.; Chemaly, M.; Davies, R.; De Cesare, A.; Herman, L.; Hilbert, F.; et al. Pathogenicity Assessment of Shiga Toxin-Producing *Escherichia coli* (STEC) and the Public Health Risk Posed by Contamination of Food with STEC. *EFSA J.* **2020**, *18*, e05967. [\[CrossRef\]](#)
12. Krüger, A.; Lucchesi, P.M.A. Shiga Toxins and Stx Phages: Highly Diverse Entities. *Microbiology* **2015**, *161*, 451–462. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Muniesa, M.; Recktenwald, J.; Bielaszewska, M.; Karch, H.; Schmidt, H. Characterization of a Shiga Toxin 2e-Converting Bacteriophage from an *Escherichia coli* Strain of Human Origin. *Infect. Immun.* **2000**, *68*, 4850–4855. [\[CrossRef\]](#)
14. de Sablet, T.; Bertin, Y.; Vareille, M.; Girardeau, J.P.; Garrivier, A.; Gobert, A.P.; Martin, C. Differential Expression of Stx2 Variants in Shiga Toxin-Producing *Escherichia coli* Belonging to Seropathotypes A and C. *Microbiology* **2008**, *154*, 176–186. [\[CrossRef\]](#)
15. Burgán, J.; Krüger, A.; Lucchesi, P.M.A. Comparable Stx2a Expression and Phage Production Levels between Shiga Toxin-Producing *Escherichia coli* Strains from Human and Bovine Origin. *Zoonoses Public Health* **2020**, *67*, 44–53. [\[CrossRef\]](#)
16. Herold, S.; Paton, J.C.; Paton, A.W. Sab, a Novel Autotransporter of Locus of Enterocyte Effacement-Negative Shiga-Toxigenic *Escherichia coli* O113:H21, Contributes to Adherence and Biofilm Formation. *Infect. Immun.* **2009**, *77*, 3234–3243. [\[CrossRef\]](#)
17. Vélez, M.V.; Colello, R.; Etcheverría, A.I.; Padola, N.L. *Escherichia coli* Productora de Toxina Shiga: El Desafío de Adherirse Para Sobrevivir. *Rev. Argent. Microbiol.* **2022**, *55*, 100–107. [\[CrossRef\]](#)
18. Tarr, P.I.; Gordon, C.A.; Chandler, W.L. Shiga-Toxin-Producing *Escherichia coli* and Haemolytic Uraemic Syndrome. *Lancet* **2005**, *365*, 1073–1086. [\[CrossRef\]](#)
19. Montero, D.; Velasco, J.; Del Canto, F.; Puente, J.L.; Padola, N.L.; Rasko, D.A.; Farfán, M.; Salazar, J.C.; Vidal, R. Locus of Adhesion and Autoaggregation (LAA), a Pathogenicity Island Present in Emerging Shiga Toxin-Producing *Escherichia coli* Strains. *Sci. Rep.* **2017**, *7*, 7011. [\[CrossRef\]](#)
20. Spears, K.J.; Roe, A.J.; Gally, D.L. A Comparison of Enteropathogenic and Enterohaemorrhagic *Escherichia coli* Pathogenesis. *FEMS Microbiol. Lett.* **2006**, *255*, 187–202. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Krishnan, S.; Prasadara, N.V. Outer Membrane Protein A and OprF: Versatile Roles in Gram-Negative Bacterial Infections. *FEBS J.* **2012**, *279*, 919–931. [\[CrossRef\]](#)
22. James, C.E.; Mahendran, K.R.; Molitor, A.; Bolla, J.M.; Bessonov, A.N.; Winterhalter, M.; Pagès, J.M. How β -Lactam Antibiotics Enter Bacteria: A Dialogue with the Porins. *PLoS ONE* **2009**, *4*, 10–13. [\[CrossRef\]](#)
23. Krewulak, K.D.; Vogel, H.J. Structural Biology of Bacterial Iron Uptake. *Biochim. Biophys. Acta-Biomembr.* **2008**, *1778*, 1781–1804. [\[CrossRef\]](#)
24. Gentle, I.E.; Burri, L.; Lithgow, T. Molecular Architecture and Function of the Omp85 Family of Proteins. *Mol. Microbiol.* **2005**, *58*, 1216–1225. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Bishop, R.E. Structural Biology of Membrane-Intrinsic β -Barrel Enzymes: Sentinels of the Bacterial Outer Membrane. *Biochim. Biophys. Acta-Biomembr.* **2008**, *1778*, 1881–1896. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Ulett, G.C.; Webb, R.I.; Schembri, M.A. Antigen-43-Mediated Autoaggregation Impairs Motility in *Escherichia coli*. *Microbiology* **2006**, *152*, 2101–2110. [\[CrossRef\]](#)
27. Montero, D.A.; Garcia-Betancourt, R.; Vidal, R.M.; Velasco, J.; Palacios, P.A.; Schneider, D.; Vega, C.; Gómez, L.; Montecinos, H.; Soto-Shara, R.; et al. A Chimeric Protein-Based Vaccine Elicits a Strong IgG Antibody Response and Confers Partial Protection against Shiga Toxin-Producing *Escherichia coli* in Mice. *Front. Immunol.* **2023**, *14*, 1186368. [\[CrossRef\]](#)

28. Montero, D.A.; Del Canto, F.; Velasco, J.; Colello, R.; Padola, N.L.; Salazar, J.C.; Martin, C.S.; Oñate, A.; Blanco, J.; Rasko, D.A.; et al. Cumulative Acquisition of Pathogenicity Islands Has Shaped Virulence Potential and Contributed to the Emergence of LEE-Negative Shiga Toxin-Producing *Escherichia coli* Strains. *Emerg. Microbes Infect.* **2019**, *8*, 486–502. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Sayers, E.W.; Bolton, E.E.; Brister, J.R.; Canese, K.; Chan, J.; Comeau, D.C.; Connor, R.; Funk, K.; Kelly, C.; Kim, S.; et al. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **2022**, *50*, D20–D26. [\[CrossRef\]](#)
30. Qi, W.; Lacher, D.W.; Bumbaugh, A.C.; Hyma, K.E.; Lindsey, M.; Large, T.M.; Tarr, C.L.; Whittam, T.S. Ec MLST: An Online Database for Multi Locus Sequence Typing of Pathogenic *Escherichia coli*. In Proceedings of the 3rd International IEEE Computer Society Computational Systems Bioinformatics Conference (CSB 2004), Stanford, CA, USA, 16–19 August 2004; pp. 520–521. [\[CrossRef\]](#)
31. Savojardo, C.; Martelli, P.L.; Fariselli, P.; Profiti, G.; Casadio, R. BUSCA: An Integrative Web Server to Predict Subcellular Localization of Proteins. *Nucleic Acids Res.* **2018**, *46*, 459–466. [\[CrossRef\]](#)
32. Sahl, J.W.; Caporaso, J.G.; Rasko, D.A.; Keim, P. The Large-Scale Blast Score Ratio (LS-BSR) Pipeline: A Method to Rapidly Compare Genetic Content between Bacterial Genomes. *PeerJ* **2014**, *2*, e332. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [\[CrossRef\]](#)
34. Ahmad, S.; Jose, L.; Bowler-Barnett, E.H.; Rice, D.L.; Kim, M.; Wijerathne, S.; Luciani, A.; Kandasaamy, S.; Luo, J.; Watkins, X.; et al. The UniProt website API: Facilitating programmatic access to protein knowledge. *Nucleic Acids Res.* **2025**, *53*, W547–W553. [\[CrossRef\]](#)
35. Blum, M.; Andreeva, A.; Florentino, L.C.; Chuguransky, S.R.; Grego, T.; Hobbs, E.; Pinto, B.L.; Orr, A.; Pa, T.; Ponamareva, I.; et al. InterPr O: The Protein Sequence Classification Resource in 2025. *Nucleic Acids Res.* **2025**, *53*, 444–456. [\[CrossRef\]](#)
36. Biegert, A.; Lupas, A.N. The HHpred Interactive Server for Protein Homology Detection and Structure Prediction. *Nucleic Acids Res.* **2005**, *33*, 244–248. [\[CrossRef\]](#)
37. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L. Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Fleming, J.; Magana, P.; Nair, S.; Tsenkov, M.; Bertoni, D.; Pidruchna, I.; Querino, M.; Afonso, L.; Laydon, A.; Midlik, A.; et al. AlphaFold Protein Structure Database and 3D-Beacons: New Data and Capabilities Q. *J. Mol. Biol.* **2025**, *437*, 168967. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Szklarczyk, D.; Kirsch, R.; Koutrouli, M.; Nastou, K.; Mehryary, F.; Hachilif, R.; Gable, A.L.; Fang, T.; Doncheva, N.T.; Pyysalo, S.; et al. The STRING Database in 2023: Protein–Protein Association Networks and Functional Enrichment Analyses for Any Sequenced Genome of Interest. *Nucleic Acids Res.* **2023**, *51*, 638–646. [\[CrossRef\]](#)
40. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R.D.; Bairoch, A.; Servet, R.M. ExPASy: The Proteomics Server for in-Depth Protein Knowledge and Analysis. *Nucleic Acids Res.* **2003**, *31*, 3784–3788. [\[CrossRef\]](#)
41. Ponomarenko, J.; Papangelopoulos, N.; Zajonc, D.M.; Peters, B.; Sette, A.; Bourne, P.E. IEDB-3D: Structural Data within the Immune Epitope Database. *Nucleic Acids Res.* **2011**, *39*, 1164–1170. [\[CrossRef\]](#)
42. Paton, A.W.; Paton, J.C. Direct Detection and Characterization of Shiga Toxigenic *Escherichia coli* by Multiplex PCR for Stx1, Stx2, Eae, EhxA, and Saa. *J. Clin. Microbiol.* **2002**, *40*, 271–274. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Amézquita-López, B.A.; Quiñones, B.; Lee, B.G.; Chaidez, C. Virulence Profiling of Shiga Toxin-Producing *Escherichia coli* Recovered from Domestic Farm Animals in Northwestern Mexico. *Front. Cell. Infect. Microbiol.* **2014**, *4*, 7. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Rivera, F.P.; Sotelo, E.; Morales, I.; Menacho, F.; Medina, A.M.; Evaristo, R.; Valencia, R.; Carbajal, L.; Ruiz, J.; Ochoa, T.J. Short Communication: Detection of Shiga Toxin-Producing *Escherichia coli* (STEC) in Healthy Cattle and Pigs in Lima, Peru 1. *J. Dairy Sci.* **2012**, *95*, 1166–1169. [\[CrossRef\]](#)
45. Galarce, N.; Sánchez, F.; Escobar, B.; Lapierre, L.; Cornejo, J.; Alegría-Morán, R.; Neira, V.; Martínez, V.; Johnson, T.; Fuentes-Castillo, D.; et al. Genomic Epidemiology of Shiga Toxin-Producing *Escherichia coli* Isolated from the Livestock-Food-Human Interface in South America. *Animals* **2021**, *11*, 1845. [\[CrossRef\]](#)
46. Bai, X.; Wang, H.; Xin, Y.; Wei, R.; Tang, X.; Zhao, A.; Sun, H.; Zhang, W.; Wang, Y.; Xu, Y.; et al. Prevalence and Characteristics of Shiga Toxin-Producing *Escherichia coli* Isolated from Retail Raw Meats in China. *Int. J. Food Microbiol.* **2015**, *200*, 31–38. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Zhang, J.-W.; Xuan, C.-G.; Lu, C.-H.; Guo, S.; Yu, J.-F.; Asif, M.; Jiang, W.-J.; Zhou, Z.-G.; Luo, Z.-Q.; Zhang, L.-Q. AidB, a Novel Thermostable N-Acylhomoserine Lactonase from the Bacterium *Bosea* sp. *Appl. Environ. Microbiol.* **2019**, *85*, e02065-19. [\[CrossRef\]](#)
48. Ozhelvaci, F.; Steczkiewicz, K. α/β Hydrolases: Toward Unraveling Entangled Classification. *Proteins Struct. Funct. Bioinform.* **2025**, *93*, 855–870. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Ollis, D.L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S.M.; Harel, M.; Remington, S.J.; Silman, I.; Schrag, J.; et al. The α/β Hydrolase Fold. *Protein Eng. Des. Sel.* **1992**, *5*, 197–211. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Grodberg, J.; Dunn, J.J. OmpT Encodes the *Escherichia coli* Outer Membrane Protease That Cleaves T7 RNA Polymerase during Purification. *J. Bacteriol.* **1988**, *170*, 1245–1253. [\[CrossRef\]](#) [\[PubMed\]](#)

51. Farr, S.B.; Arnosti, D.N.; Chamberlin, M.J.; Amest, B.N. An ApaH Mutation Causes AppppA to Accumulate and Affects Motility and Catabolite Repression in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5010–5014. [[CrossRef](#)] [[PubMed](#)]
52. Gohari, I.M.; Navarro, M.A.; Li, J.; Shrestha, A.; Uzal, F.; McClane, B.A. Pathogenicity and Virulence of *Clostridium Perfringens* ABSTRACT. *Virulence* **2024**, *12*, 723–753. [[CrossRef](#)] [[PubMed](#)]
53. Collin, M.; Olse, A. Extracellular Enzymes with Immunomodulating Activities: Variations on a Theme in *Streptococcus Pyogenes*. *Infect. Immun.* **2003**, *71*, 2983–2992. [[CrossRef](#)]
54. Michino, H.; Araki, K.; Minam, S.; Takaya, S.; Sakai, N.; Miyazaki, M.; Ono, A.; Yanagawa, H. Massive Outbreak of *Escherichia coli* 0157:H7 Infection in Schoolchildren in Sakai City, Japan, Associated with Consumption of White Radish Sprouts. *Am. J. Play* **1999**, *150*, 787–796.
55. Seadawy, M.G.; Zekri, A.R.N.; Saeed, A.A.; San, E.J.; Ageez, A.M. Candidate Multi-Epitope Vaccine against Corona B.1.617 Lineage: In Silico Approach. *Life* **2022**, *12*, 1715. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.