BMC Veterinary Research

RESEARCH

Open Access

Identification and bioinformatics analysis of the *DUS* gene in *Eimeria media*

Yijin Zou[1], Yiyan Wang[1], Diyi Duan[1], Xun Suo[1*] and Yuanyuan Zhang[1*]

Abstract

This study aims to explore the coding sequence (CDS) of the putative *DUS* gene in *Eimeria media* and assess its potential biological functions during the parasite's lifecycle. Initially, oocysts were isolated from fecal samples of rabbits infected with *E. media*, from which DNA and RNA were extracted. These extractions were used as templates for PCR to successfully amplify the CDS of the *DUS* gene, confirming its presence within the *E. media* genome. Further analysis using quantitative PCR (qPCR) demonstrated significantly higher expression of the *DUS* gene in the precocious line (PL) compared to the wild type (WT). This differential expression highlights a potential functional role for the *DUS* gene in influencing the development and sporulation processes of *E. media*, which may contribute to the precocious phenotype. Additionally, bioinformatics analysis provided insights into the evolutionary trends and structural characteristics of the *DUS* gene across different species, suggesting a broader biological significance. The elevated expression of the *DUS* gene in the PL suggests its critical involvement in the growth and reproductive processes of the parasite. This finding opens new avenues for research aimed at controlling *E. media* infection through targeted interventions in the *DUS* gene expression pathways.

Keywords *DUS* gene, *Eimeria media*, Precocious phenotype, Bioinformatics

Background

Coccidiosis is a disease caused by the single-celled protozoan *Eimeria* spp. and occurs globally. This disease has been documented in various livestock including poultry, significantly affecting their growth and development [1, 2]. The primary mode of coccidia infection is through the ingestion of food and water contaminated with oocysts. Upon infection, coccidia replicate and reproduce in the host's digestive tract, causing symptoms such as enteritis and diarrhea [3]. They rapidly multiply within the host, forming many oocysts that are excreted in the feces, thereby spreading to other animals [4].

Rabbits are economically significant animals in many countries, valued for their meat and fur. One major disease affecting rabbit farming is coccidiosis, with *Eimeria media* as one of the most common *Eimeria* species infecting rabbits. *E. media* is a medium-sized coccidian species. It primarily infects domestic rabbits, leading to a range of clinical symptoms including diarrhea, loss of appetite, and weight loss, with severe cases resulting in death. Due to its high infectivity and drug resistance, coccidiosis is easily transmitted in intensive farming environments, causing significant economic losses to agriculture.

In our laboratory, we previously selected an *E. media* line with a pronounced early-maturing trait (Gu 2019), making it an ideal mutant model for studying developmental patterns of coccidia. The precocious line (PL) of *E. media* exhibits characteristics such as a shortened

*Correspondence:
Xun Suo
suoxun@cau.edu.cn
Yuanyuan Zhang
yyzhang@cau.edu.cn
[1] National Key Laboratory of Veterinary Public Health Security, Key Laboratory of Animal Epidemiology and Zoonosis of Ministry of Agriculture, National Animal Protozoa Laboratory & College of Veterinary Medicine, China Agricultural University, Beijing 100193, China

lifecycle, reduced pathogenicity, and weakened reproductive capability, providing a unique tool for controlling coccidiosis. The PL begins oocyst output at 70 h post-inoculation, peaking at 120 h, while the wild-type (WT) begins oocyst output at 108 h post-inoculation, peaking at 196 h. Investigating the molecular mechanisms of these characteristics can offer biological insights, contributing to a more in-depth understanding of the life cycle of coccidia and host interactions. We believe that using *E. media* as a research model allows the findings to be extended to other coccidia species. Additionally, *E. media* is relatively unexplored. Studying this species can fill knowledge gaps in evolutionary mechanisms and interspecies relationships of protozoan parasites, enhancing our understanding of parasitic biology and advancing parasitology.

In the early genetic screening, we identified genetic differential regions between the PL and WT genomes. *DUS* gene is one of the candidate genes that is functionally important in development. Further functional predictions indicate that the *DUS* gene, along with its surrounding region, exhibits abundant variations and is likely a key gene involved in the developmental process. Therefore, conducting an in-depth analysis of the sequence, structure, and function of the *DUS* gene is of significant research value. Structural analysis and homology comparisons have shown that the gene structure is highly conserved across most coccidia species, suggesting that the *DUS* gene's function is both conserved and crucial across different coccidia species. By studying the expression and regulation of the *DUS* gene and other key genes, we can uncover the molecular mechanisms of precocious phenomena, contributing new knowledge to coccidia biology research.

The *DUS* gene, encoding dihydrouridine synthase, is involved in RNA modification, particularly in tRNA, and is found in bacteria, eukaryotes, and other organisms [5]. Dihydrouridine contributes to maintaining the stable three-dimensional structure of tRNA, which is crucial for successful protein synthesis [6]. It is believed that the expression level of the *DUS* gene is closely related to cell proliferation and development. Overexpression of *DUS* may impact tRNA modification, and subsequently, gene transcription and translation, influencing cell growth and differentiation processes [7, 8]. Although the specific functions of the *DUS* gene may vary among different organisms, the role of *DUS* in RNA modification and cellular function regulation is ubiquitous.

This study focuses on identifying and validating the coding sequence (CDS) of *DUS*. PCR technology and specifically designed primers were used to amplify the gene, followed by sequencing alignment to confirm its sequence. Bioinformatics tools and methods were employed to conduct a comprehensive analysis of the *DUS* gene sequence to understand its potential functions and roles in coccidia. Comparative analysis, phylogenetic tree construction, domain analysis, and secondary and tertiary structure predictions were performed using online tools. Based on experimental data, we explored the biological significance of the *DUS* gene in coccidiosis, particularly its potential role in the PL of *E. media*.

Materials and methods

Rabbit coccidia inoculation and feces collection

Eight healthy New Zealand white rabbits were provided by the Zhuozhou Teaching Experimental Farm of China Agricultural University. The rabbits were 9 weeks old with an average body weight of 2778 (± 187) g. They were acclimatized in our laboratory for 7 days before the experiment. The rabbits were housed in ammonia-prefumigated isolators, and fed coccidian-free food and water (baked at 80 °C for 2 h). During the period, rabbit feces were collected and tested for coccidian oocysts three times using the saturated saline method and optical microscopy to ensure these animals have no coccidia infection. Subsequently, the rabbits were randomly divided into two groups: one group was inoculated with the precocious line (PL) of *E. media*, and the other with the wild type (WT).

Parasites and infection

Oocysts of WT and PL strains of *E. media* were stored in the National Animal Protozoa Lab [9]. Each rabbit in the two experimental groups received 20,000 coccidia oocysts by oral gavage at Day 0. Fecal samples were collected during the period of oocyst shedding, i.e. 72–120 h for the PL strain and 108–156 h for the WT strain (Gu et al. 2019) [9]. All animal experiments were approved by the China Agricultural University Laboratory Animal Welfare and Ethics Committee.

Oocyst collection and purification

Oocysts were collected and purified following the protocol outlined by Long et al. [10]. The process involved mixing feces with water, followed by filtration and centrifugation. The precipitated oocysts were sporulated in potassium dichromate at 28 °C with shaking at 120 rpm for 48 h. After sporulation, oocysts were treated with sodium hypochlorite, washed, and stored in potassium dichromate solution at 4 °C.

DNA and RNA extraction

Approximately 10^7 purified oocysts were used for genomic DNA extraction using the CTAB method. The quality and concentration of DNA were validated

Table 1 PCR reaction system (Taq)

Component	Volume/μL
RNA extraction liquid	10
10 × RT Mix	2
HiScript III Enzyme Mix	2
Oligo(dT)20VN	1
Random hexamers	1
RNase-free ddH2O	4

Table 2 Primers used for amplifying the target fragment

Primer Name	Sequence (5′ to 3′)
DUS-F	ATGACACCTACAGATGGGAGCAATG
DUS-R	TCATGAGAGGTCAATGTTTTTATA GTAATTGAGAGT

Table 3 PCR Reaction system

Component	Volume/μL
Water	7
2 × Taq Mix	10
DUS-F	1
DUS-R	1
DNA Template	1
RNase-free ddH2O	4

Table 4 PCR Reaction conditions

Step	Temperature (°C)	Time	Cycles
Initial Denaturation	95	10 min	1
Denaturation	95	30 s	
Annealing	50–65	30 s	35
Extension	72	15 s/kb	
Final Extension	72	10 min	1
Hold	16	∞	

by nanodrop. Another 10^7 oocysts were used for RNA extraction and cDNA synthesis.

DNA Extraction: Sporocysts were obtained from approximately 10^7 purified oocysts through mechanical disruption of the oocyst wall. The sporocysts were then washed with PBS and centrifuged. DNA extraction was performed using the CTAB lysis buffer with proteinase K digestion, followed by RNase treatment. The DNA was purified using the phenol: chloroform: isoamyl alcohol method and precipitated with isopropanol. The DNA was then washed with 75% ethanol, dried, and resuspended in distilled water. DNA concentration and quality were assessed using a spectrophotometer.

RNA Extraction and cDNA Synthesis: The TRIzol method was used to extract RNA from purified oocysts. For cDNA synthesis, the purified RNA was processed using an RNA kit according to the manufacturer's protocol as shown in Table 1.

PCR amplification and verification of the *DUS* gene

The full length of the cDNA sequences were obtained using PCR amplification and TA cloning. Using the SnapGene tool, specific primers were designed based on the CDS sequence of the *DUS* gene provided by the laboratory. This CDS sequence has not been published and does not have a GenBank accession number in NCBI. After multiple optimizations, the optimal PCR reaction conditions were selected to achieve the highest amplification efficiency. The primers used for amplifying the target fragment are shown in Table 2, the PCR reaction system is shown in Table 3 and the PCR reaction conditions are shown in Table 4. After amplification, gel electrophoresis was performed to confirm the correct target band, and the PCR product was sequenced. The synthesis of primers and sequencing of the PCR product were performed by Beijing Qianke Biological Technology Co., Ltd.

Using cDNA as the template, PCR was performed again with 35 cycles. Initially, only the precocious line exhibited a faint band, indicating that the primers used might have a stronger affinity for the target gene sequence of the PL and a weaker affinity for the WT. In addition, the relatively faint band of the target gene was also possibly due to insufficient amplification of the target fragment. To address this issue, the PCR program was optimized by increasing the number of cycles. In the second attempt, the cycle number was increased to 40, and the product of the first PCR was used as the template, with the same program and system as that in the previous PCR run. After optimization, a band was clearly visible for the PL. To ensure the accuracy of the results, the sample was sequenced to verify the PCR product.

Real-time fluorescence quantitative PCR analysis

Regular PCR amplification was performed using the obtained cDNA samples to ensure that the template and primers were correct before qPCR. The primers used for amplifying the target fragment are shown in Table 5, the PCR reaction system is shown in Table 6, and the PCR reaction conditions are shown in Table 4. The primers were independently designed by the authors and synthesized by Beijing Qingke Biotechnology Co., Ltd.

After observing a clear target band, the qPCR experiment was conducted. The primers are listed in Table 5,

Table 5 Primers used for amplifying the target fragment

Primer Name	Sequence (5′ to 3′)
DUS-F	ATGACACCTACAGATGGGAGCAATG
DUS-R	TCATGAGAGGTCAATGTTTTTATA GTAATTGAGAGT
Actin-F	AAAGGACCTCTACGGCAACG
Actin-R	GATGCTTGGTCCAGACTCGT

Table 6 PCR Reaction system

Component	Volume/μL
Water	8
2×Taq Mix	10
DUS-F/ Actin-F	0.5
DUS-R/ Actin-R	0.5
cDNA Template	1

Table 7 PCR reaction system

Component	Volume/μL
Water	8
2×Taq Mix	10
DUS-F/ Actin-F	0.5
DUS-R/ Actin-R	0.5
PCR Template	1

and the system is detailed in Table 7. A lower CT value indicates a higher initial quantity of the target nucleic acid in the sample.

T-Vector cloning and positive clone screening

The purified PCR product (4 μl) was combined with 1 μl of T-vector, ligase, and buffer, and incubated for 15 min at 30 °C to enable ligation. This mixture was transformed into *Escherichia coli* competent cells by the heat shock method at 42 °C followed by ice cooling. After incubation in LB medium without antibiotics and centrifugation, about 150 μl of cell pellet was obtained. Transformed cells were plated on agar with antibiotics and incubated for 16 h, using blue-white screening for colony selection. White or light-colored, indicating successful insertion, underwent colony PCR with M13F and M13R primers for verification. Gel electrophoresis of PCR products identified clear bands, from which three were sequenced. The most promising sequence was further analyzed with universal primers.

DUS gene bioinformatics analysis and protein function exploration

Using the CDS sequence of the *DUS* gene as a query, a BLAST [11] comparison was conducted on NCBI, and sequences with 46% or more coverage were selected. Eight sequences met the selection criteria. The CDS sequences were translated into protein sequences using the "batch translate CDS to Protein" function in TBtools [12], and multiple sequence alignment was performed using the CLUSTALW online tool [13], with visualization achieved through the espript3 tool [14]. Further analysis of the protein sequences was conducted using MEGA11 software [15], where the JTT + G + F model was chosen, and the Neighbor-Joining method was used to construct the evolutionary tree. Motif searching was performed using the MEME online tool [16], and the distribution of motifs was visualized with TBtools. Structural feature analysis involved identifying domains using the NCBI Batch CDD tool [17] and visualizing the results with TBtools. The physicochemical properties of the proteins were analyzed using the ProtParam tool from Expasy [18], which provided parameters such as amino acid count, molecular weight, and isoelectric point, with all data integrated for a detailed overview. Secondary structure prediction was performed using the PRABI tool [19], and tertiary structure prediction was performed using the SWISS-MODEL tool on the Expasy homepage, based on known structures of similar proteins in the database.

Tools and software utilized in this study

During this project, a variety of tools and software were employed to facilitate different tasks and enhance productivity. The tools and software are listed in Table 8. The primary reagents and their suppliers are listed in Table 9.

Statistics methods

The t-test was used to compare the expression levels of the *DUS* gene between the PL and the WT *E. media* to determine if there is a significant difference between the two groups. Analysis of Variance (ANOVA) was used to determine the statistical significance of differences among the groups. The relative quantification method (ΔΔCT method) was used to calculate the changes in expression levels by comparing the difference in expression levels between the target gene and a reference gene (such as Actin). In the bioinformatics analysis part, probability models were utilized to predict the likelihood of the presence of specific motifs or structural domains.

Results

Oocysts isolation and extraction of DNA and RNA

Approximately 2000 g of fecal samples were collected from inoculated rabbits, from which coccidian oocysts were isolated, sporulated, and treated to stabilize their structure. DNA was successfully extracted from *E. media* oocysts, with a concentration of 115.690 ng/μL confirmed by UV–visible spectrophotometry. PCR using DNA templates verified the presence and accuracy of the *DUS* gene sequence. RNA was also extracted, reverse-transcribed to cDNA, and used for PCR to assess gene expression at the transcriptional level, providing insights into the gene's regulatory mechanisms.

DUS gene sequence

The coding sequence (CDS) of the *DUS* gene in *E. media* was determined to be 1512 bp in length (Fig. 1). It encodes proteins directly and does not include promoters or introns. In this 1512 base pair DNA sequence, the GC

Table 8 Utilized tools and software

Tool/Software Name	Function Description	Reference
NCBI BLAST	Performs BLAST comparison using the CDS sequence of the DUS gene as a query	[11]
TBtools	Batch translates CDS sequences into corresponding protein sequences	[12]
CLUSTALW	Performs multiple sequence alignment	[13]
ESPript	Generates easily readable alignment visualization diagrams	[14]
MEGA11	Uses the JTT + G + F model and the Neighbor-Joining (NJ) method to construct evolutionary trees	[15]
MEME	Searches for potential structural domains or functional modules	[16]
NCBI Batch CDD	Domain identification and analysis	[17]
Expasy	Analyzes basic physicochemical properties of proteins	[18]
PRABI	Predicts the protein's secondary structure	[19]
SnapGene	Used for primer design and graphical representation of DNA sequences	[20]
GraphPad Prism 9	A graphics and statistics software used for scientific research, particularly suited for graphing qPCR expression levels	[21]

Table 9 Primary reagents

Reagent	Company
2×Taq Mix	Beijing Zoman Biotechnology Co., Ltd
Q5® High-Fidelity DNA Polymerase	New England Biolabs, USA
pEASY®-T1 Simple Cloning Kit	Beijing TransGen Biotech Co., Ltd
pEASY®-Uni Seamless Cloning and Assembly Kit	Beijing TransGen Biotech Co., Ltd
pEASY®-Blunt Simple Cloning Kit	Beijing TransGen Biotech Co., Ltd
Trans1-T1 Phage Resistant Chemically Competent Cells	Beijing TransGen Biotech Co., Ltd
Transetta (DE3) Chemically Competent Cells	Beijing TransGen Biotech Co., Ltd
Large Quantity Plasmid Extraction Kit	Beijing Biomed Gene Technology Co., Ltd
Large Quantity Plasmid Extraction Kit	TaKaRa Biotechnology (Dalian) Co., Ltd

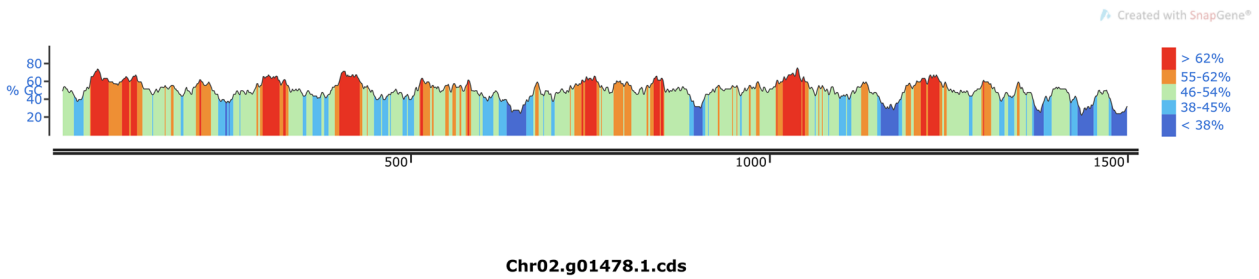

Fig. 1 *E. media* DUS CDS sequence, which is 1512 bp

content is 50.40%, accounting for approximately half of the nucleotide composition, indicating moderate thermal stability. The analysis revealed multiple repeat sequences, including di-nucleotide, tri-nucleotide, and tetra-nucleotide repeats. These repeat sequences may influence gene regulatory functions or structural stability, particularly tri-nucleotide and tetra-nucleotide repeats, which are closely associated with specific gene expression regulation or protein functions.

The amino acid distribution in the protein sequence reveals several high-frequency amino acids: serine (S, 48 occurrences), alanine (A, 45 occurrences), leucine (L, 43 occurrences), and threonine (T, 37 occurrences). These amino acids are closely related to the structural stability and function of the protein, suggesting the presence of significant hydrophilic and hydrophobic regions, which may be involved in critical functions such as transmembrane segments or protein folding.

An analysis of codon usage preference in encoding specific amino acids was conducted. Common codons include ATG (methionine, M, 13 occurrences), GAT and GAC (aspartic acid, D, 16.5 times each), AAA and AAG (lysine, K, 15.5 times each), and GAA and GAG (glutamic acid, E, 16.5 times each). The analysis of codon usage preference indicates that specific codons are frequently used in the encoding of aspartic acid, lysine, and glutamic acid, reflecting specific demands during gene expression or regulatory mechanisms formed during protein evolution. Additionally, the even distribution of synonymous codons implies that the protein may be adaptable across diverse expression systems. This CDS analysis is essential for understanding the role of the *DUS* gene in *E. media* and provides a critical foundation for further functional research.

On the electrophoresis gel, a clear band around 1500 bp was observed, indicating the successful amplification of the target gene sequence. For the PL strain of *E. media*, the E-value was 0 with a query cover of 86%. For the WT *E. media*, the E-value was 2e-14, and the query cover was 87% (Table 10). This indicates that the *DUS* gene exists in the genome of *E. media*, and the base alignment results are shown in Additional File 1 and Additional File 2. In comparative analyses of the *DUS* gene's CDS sequences with the PL-DNA and WT-DNA, numerous highly conserved regions (highlighted in red) were observed indicating a high degree of sequence consistency between the CDS and both PL-DNA and WT-DNA. Additionally, both comparisons commonly exhibited single nucleotide polymorphisms (SNPs) and insertion/deletion (indel) events, which could play crucial roles in the functional expression of proteins, particularly mutations within known functional domains. However, mutations patterns varied between the two comparisons, reflecting the genetic diversity between the two *E. media* strains. Specifically, the comparison with PL-DNA showed more conserved regions, while the comparison with WT-DNA revealed more frequent mutations, indicating a higher degree of genetic stability in certain gene sequences of the PL strain.

cDNA Template sequencing and optimization

The PCR amplification product of the PL strain was visible on the electrophoresis gel, while the band from the

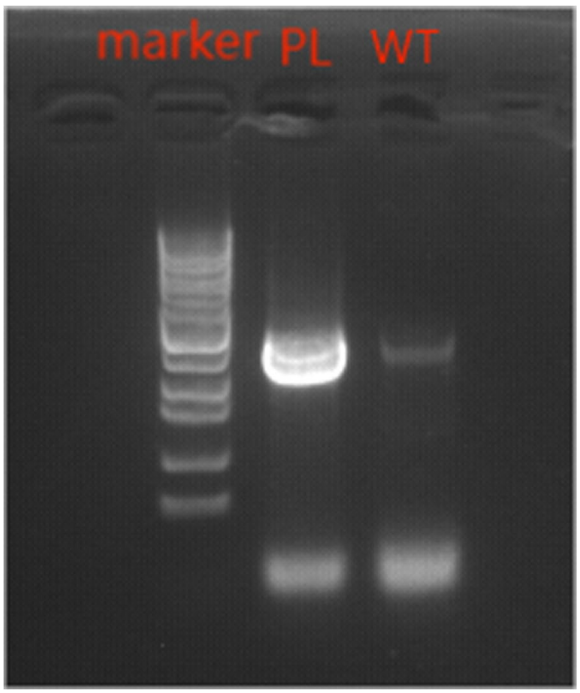

Fig. 2 PCR Bands of PL and WT cDNA

Table 10 Sequence alignment results for *E. media* PCR products

Description	Max Score	Total Score	Query Cover	E value	Per. Ident
E. media PL Vector PCR seq	2680	2680	100%	0.0	98.44%
E. media PL cDNA PCR seq	2556	2556	94%	0.0	99.86%
E. media WT DNA PCR seq	950	2374	87%	0.0	99.81%
E. media PL DNA PCR seq	603	1822	86%	2e-175	83.13%

WT strain was relatively faint (Fig. 2). The sequence alignment results are shown in Additional File 3. To verify the identity of this PL band and its match with the expected target, the PCR product was sequenced. The sequencing results were further analyzed using BLAST, yielding an E-value of 0 and a query coverage of 94% (Table 2). The sequencing and comparative analysis results confirm a high consistency between the sequence of the PL PCR product and the expected target sequence, suggests successful amplification of the target gene. Since cDNA is obtained through reverse transcription of RNA, our findings provide direct evidence of *DUS* gene transcription in *E. media*.

High-fidelity enzyme PCR and bacterial liquid PCR analysis

High-fidelity PCR was conducted to isolate the target *DUS* gene for cloning. A bright 1500 bp band was obtained, which was recovered from the electrophoresis gel, and then cloned into the T vector. To verify that the expected gene was successfully inserted into the T vector, bacterial liquid PCR was conducted. The brightest band on the electrophoresis gel was selected for sequencing. Initially, sequencing was conducted using specific primers designed for this experiment. To obtain more comprehensive sequence information on the target gene, subsequent sequencing was performed using universal primers. Comparing the sequencing results with the target gene sequence useing BLAST, a sequence with E-value 0 and query cover 100% was obtained (Table 2). The findings confirmed the successful cloning of the target gene into the T vector, and the sequence is highly consistent with the known sequence of *DUS*. The alignment results are shown in Additional File 4.

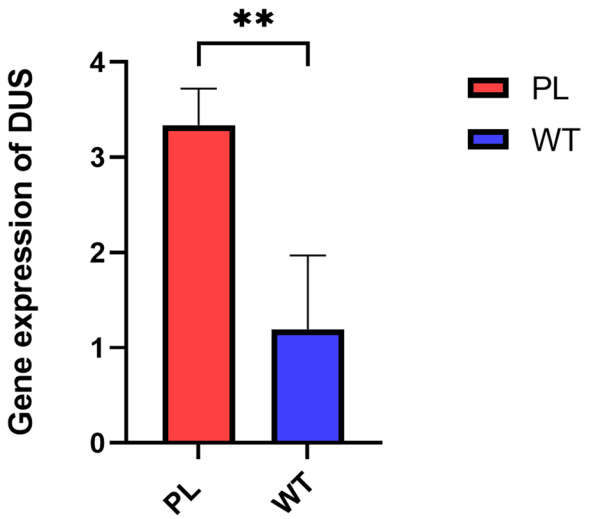

Fig. 3 *DUS* Expression in PL and WT oocysts validated by qPCR (**$P < 0.01$)

DUS expression

qPCR analysis showed the expression levels of the *DUS* gene in PL and WT oocysts (Fig. 3). The red bar represents the expression of the *DUS* gene in PL oocysts of *E. media*, while the blue bar represents its expression in WT oocysts of *E. media*. As shown in the figure, the expression of the *DUS* gene in PL oocysts is significantly higher than in WT oocysts ($P < 0.01$). The high expression of the *DUS* gene in the PL *E. media* may provide some biological advantages, enabling it to develop and reproduce rapidly under specific environmental or ecological conditions.

qPCR analysis showed that the expression level of the *DUS* gene in oocysts of the PL group (red bar) was significantly higher than that in WT oocysts (blue bar) ($P < 0.01$). Asterisks indicate statistically significant differences between the two groups.

Bioinformatics analysis

NCBI alignment of DUS gene CDS sequences

After integrating the original fasta format query, nine sequence segments were obtained. The chosen sequences span various species of protozoan parasites. The NCBI alignments for the selected protozoan parasites are listed in Table 11. An initial comparison of DNA

Table 11 NCBI sequence alignment results for selected protozoan Parasites

Source Species	Max Score	Query Cover	E value	Per. Ident	Accession number
E. tenella	566	66%	1e-155	73.05%	JN987528.1
C. cayetanensis	542	66%	1e-148	72.15%	XM_026336677.1
E. acervulina	534	66%	7e-146	71.94%	XM_013391865.1
E. tenella	425	66%	3e-113	69.36%	XM_013372616.1
E. necatrix	415	66%	5e-110	69.13%	XM_013577463.1
E. maxima	334	46%	4e-86	74.23%	XM_013481802.1
T. gondii	180	46%	2e-39	66.43%	XM_002371431.2
B. besnoiti	169	46%	3e-36	65.82%	XM_029364201.1

sequeneces(Additional file 5) and protein sequences (Additional file 6) was conducted to assess their similarities and differences.

Evolutionary tree analysis

The evolutionary tree was constructed using the Neighbor-Joining method [22], illustrating the evolutionary relationship between the target gene "*Eimeria media* Chr02.g01478.1" and eight other amino acid sequences. This method focuses on analyzing the differences in gene sequences to infer the relationships between species. Values on the tree represent evolutionary distances, calculated according to the JTT matrix method [23], reflecting the frequency of amino acid substitutions at each site. The evolutionary distance between "*E. media*" and "*E. tenella*" is 0.29, indicating a relatively close evolutionary relationship. Conversely, the distance between "*E. media*" and "*T. gondii*" is 0.97, suggesting significant divergence and a more distant genetic relationship. The length of each branch is related to evolutionary time, with shorter branches indicating a closer common ancestor and longer branches indicating an older common ancestor. The analysis considered definite positions in all sequence pairs, with ambiguous positions being pairwise deleted to ensure accurate calculation of evolutionary distances. The final dataset involved 1026 positions, and the analysis was completed on the MEGA11 software platform [15]. The genetic relationships among different *Eimeria* species and related parasites are depicted in Fig. 4 through a molecular phylogenetic tree.

Motif distribution reveals potential differences in protein function and structure

The *DUS E. media* Chr02.g01478.1 sequence contains five important motifs. The motifs in *DUS* Like *E. tenella* and *DUS* Partial *E. acervulina* are similar, suggesting potentially similar functions. The *E. maxima* sequence has four motifs, which may indicate a functional difference from *E. media*. The *T. gondii* sequence lacks any motifs, suggesting it may have a completely different function and structure. The *B. besnoiti* sequence has four motifs, but they differ from those in *DUS E. media* Chr02.g01478.1 (Fig. 5). The protein motif visualization displays the most frequently occurring amino acids in a set of amino acid sequences. Each row represents a specific motif, composed of amino acids. The vertical size of the letters indicates the frequency of that amino acid at a given position in the sequence; the larger the letter, the higher its occurring frequency (Fig. 6).

Motif 1, Motif 2, and Motif 5 are all associated with the tRNA-dihydrouridine synthase-like (DUS-like) family, which is involved in tRNA modification by catalyzing the formation of dihydrouridine to maintain the three-dimensional structural stability of tRNA and promote protein synthesis. Their primary molecular function is tRNA dihydrouridine synthase activity (GO:0017150), and the biological process is tRNA dihydrouridine synthesis (GO:0002943). These functions primarily occur in the cytoplasm (GO:0005737).

Despite these commonalities, each motif also has its unique characteristics. Motif 1 mainly contains the

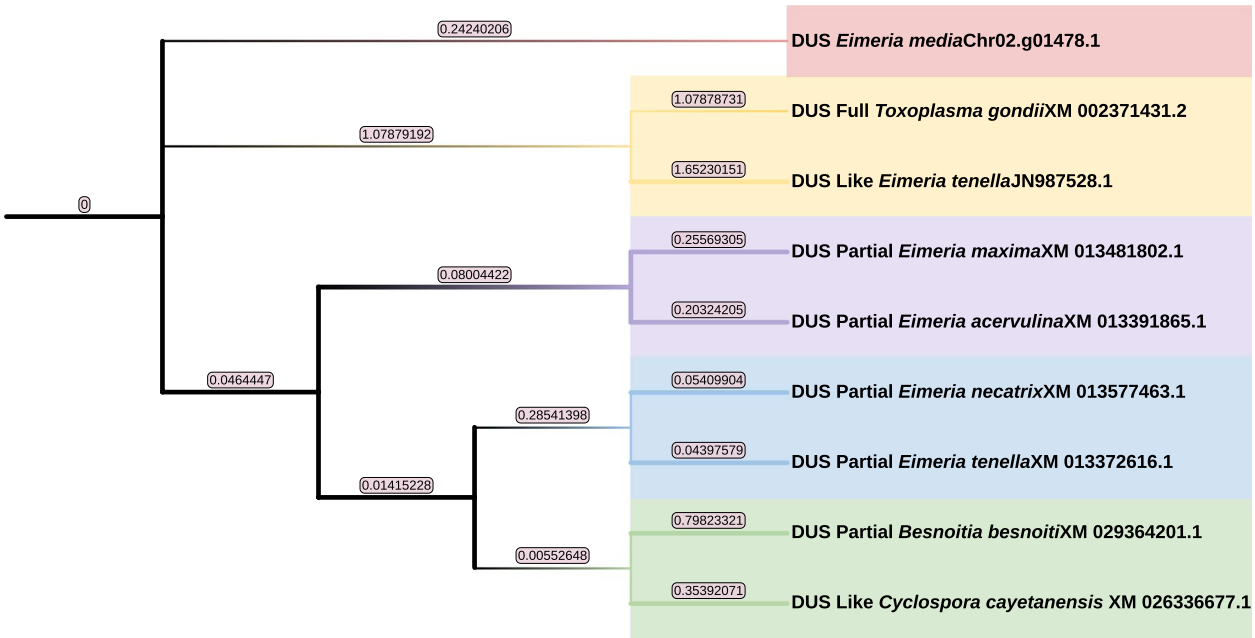

Fig. 4 Molecular phylogenetic tree showing genetic relationships among different *Eimeria* species and related parasites

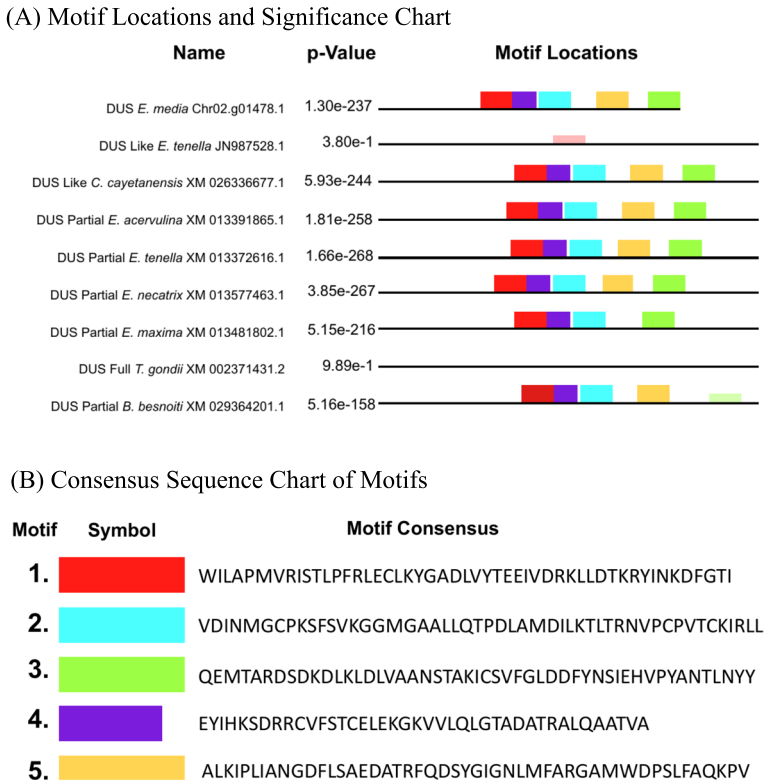

Fig. 5 *DUS* Motif distribution. **A** Motif locations and significance chart. **B** Consensus sequence chart of motifs

DUS-like_FMN-bd domain, indicating its close association with FMN (flavin mononucleotide) binding, which is crucial for redox reactions. In addition to the DUS-like_FMN-bd domain, Motif 2 is also associated with the Aldolase_TIM domain (IPR013785), typically found in enzymes involved in glycolysis and other metabolic pathways. Additionally, Motif 2 contains the tRNA_hU_synthase_CS (IPR018517) conserved site, demonstrating its significant role in the structure and function of tRNA dihydrouridine synthase and its flavin adenine dinucleotide (FAD) binding function (GO:0050660). Motif 5 also includes the DUS-like_FMN-bd and Aldolase_TIM domains, further supporting its important role in redox reactions and metabolic processes, and is associated with FMN-linked oxidoreductases, underscores its crucial role in cellular metabolism.

Structural domain analysis of *DUS* genes in *Eimeria* and related species

A structural domain analysis was conducted to reveal differences in the distribution of domains among various protein sequences. The distribution of domains is shown in Fig. 7, which illustrates the distribution and length of DUS-like FMN domains (green blocks) in multiple *DUS* genes across different *Eimeria* species and related organisms, highlighting the high degree of conservation of these genes among different species. The *DUS* genes in *E. media* (Chr02.g01478.1) and *E. tenella* (JN987528.1, XM 013372616.1) among other species, show similar DUS-like FMN domains, indicating a high level of conservation in specific functional regions across species. The position and length of these domains are largely consistent across species, suggesting the importance of this region in the gene's function and its preservation through evolution.

For instance, in *E. media* and *E. tenella*, the domain is located in the anterior midsection of the sequence, while in *C. cayetanensis*, the domain is slightly shifted backward. The length of the *DUS* gene sequences varies among species; for example, the sequences are shorter in *E. maxima* (XM 013481802.1) and *T. gondii* (XM 002371431.2) compared to the longer sequence in *B. besnoiti* (XM 029364201.1). This variation may reflect specific regulatory mechanisms or additional functional regions in different species.

The *DUS* gene encodes dihydrouridine synthase, which plays a crucial role in the modification of tRNA by catalyzing the reduction of uridine to dihydrouridine. This modification of tRNA is essential for the folding and stability of tRNA, which is critical for protein synthesis and,

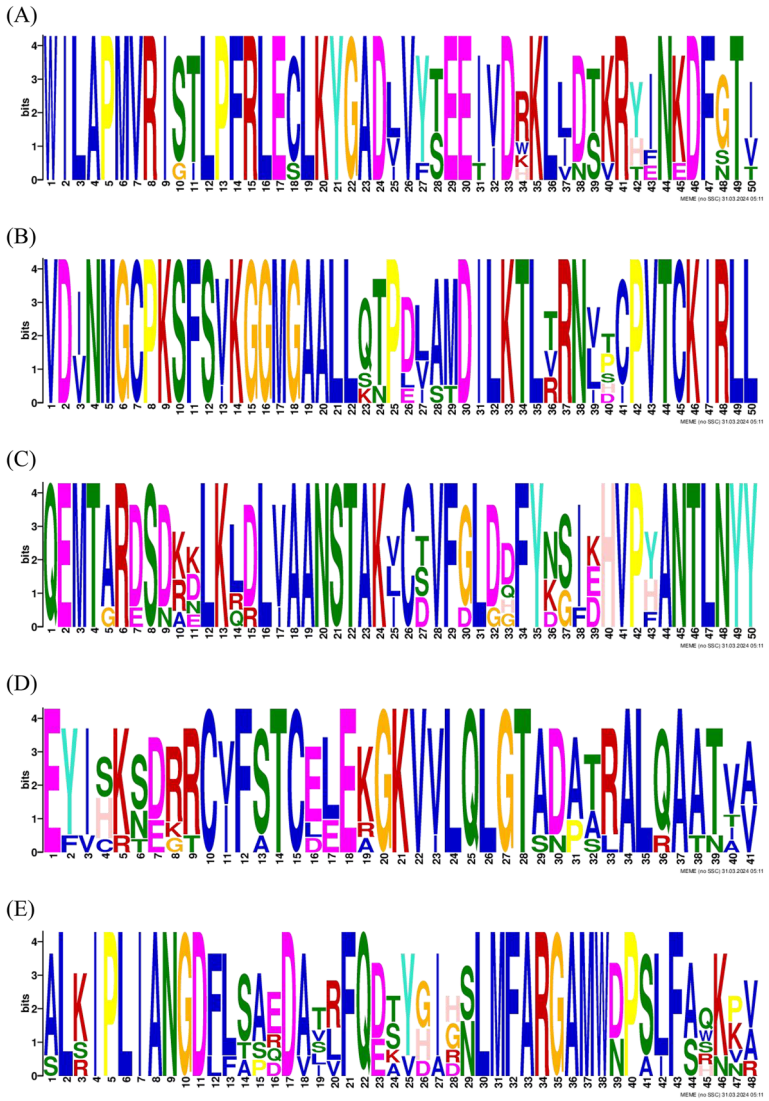

Fig. 6 Protein motif visualization

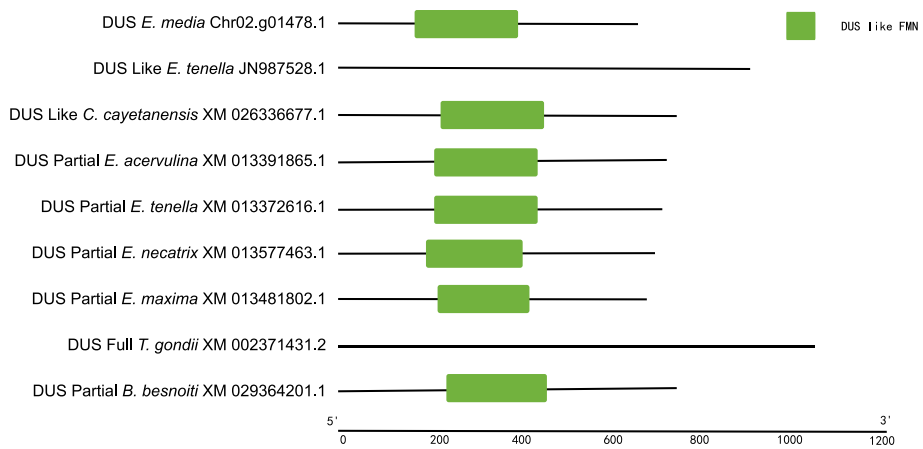

Fig. 7 Domain distribution

consequently, for the survival and reproduction of the parasite. The conservation of the DUS-like FMN domain across multiple species indicates its importance in the evolutionary adaptation of parasites. Understanding the function of these conserved domains can aid in the development of new antiparasitic drugs by targeting the essential tRNA modification process to inhibit parasite growth and proliferation.

Diversity of protein physicochemical properties and their potential functional relevance

The *DUS* proteins exhibit noticeable differences in physicochemical properties (Table 12). Most proteins have isoelectric points (pI) within the range of 5–7, but two exceptions (DUS *E. tenella* JN987528.1 and DUS *T.gondii* XM 002371431.2) have pI values of around 10. The instability index of most proteins exceeds 40, suggesting potential instability under laboratory conditions. Notably, DUS *T. gondii* XM 002371431.2 has an instability index of 70.42, significantly higher than those of the other proteins. Negative hydropathy values indicate the overall hydrophilicity of proteins. Although all the proteins have negative hydropathy values, DUS *B. besnoiti* XM 029364201.1 has the lowest negative hydropathy value, suggesting it is the most hydrophilic among these proteins. These differences may be related to the proteins' functions, locations, and behaviors within cells. For instance, proteins with high hydrophobicity may be more prone to function near or within the cell membrane, while hydrophilic proteins may predominantly function in the cytoplasm.

The protein encoded by the DUS *E. media* Chr02.g01478.1 is 503 amino acids long, representing a moderately sized protein with a molecular weight of 55,160.16 Daltons. With a pI value of 5.61, the DUS protein may carry a positive charge in slightly acidic environments. An instability index of 44.51 suggests potential protein instability. The hydropathy value of -0.392 indicates the DUS protein is hydrophilic.

Secondary and tertiary structures of the DUS proteins

The predicted secondary structure of the protein is presented in Table 13. DUS *B. besnoiti* XM 029364201.1 has the highest α-helix content at 39.55%, while DUS *T.*

Table 12 Predicted physicochemical properties

Sequence ID	Number of Amino Acid	Molecular Weight	Theoretical pI	Total number of atoms	Instability Index	Aliphatic Index	Grand Average of Hydropathicity
DUS *E. media* Chr02.g01478.1	503	55,160.16	5.61	7678	44.51	77.61	-0.392
DUS *E. tenella* JN987528.1	656	74,280.85	10.01	10,228	51.49	54.79	-0.728
DUS *C. Cayetanensis* XM 026336677.1	559	61,706.83	6.34	8570	51.47	70.61	-0.468
DUS *E. acervulina*XM 013391865.1	546	59,815.98	8.07	8360	42.42	74.58	-0.379
DUS *E. tenella* XM 013372616.1	537	58,638.25	5.96	8196	47.92	80.35	-0.368
DUS *E. necatrix* XM 013577463.1	511	55,746.1	5.58	7797	47.88	82.54	-0.285
DUS *E. maxima* XM 013481802.1	493	53,225.82	5.03	7425	47.63	80.18	-0.341
DUS *T. gondii* XM 002371431.2	1007	111,872.72	10.02	15,578	70.42	71.4	-0.374
DUS *B. besnoiti* XM 029364201.1	622	64,585.61	6.52	9015	53.45	78.31	-0.127

Table 13 Predicted protein secondary structures

Proteins	Percentage/%			
	α-Helix	Extended Strand	β-Turn	Random Coil
DUS *E. media* Chr02.g01478.1	32.01	4.97	0	40.56
DUS *E. tenella* JN987528.1	22.87	12.04	0	63.11
DUS *C. cayetanensis* XM 026336677.1	33.45	8.05	0	56.53
DUS *E. acervulina* XM 013391865.1	36.81	8.24	0	53.66
DUS *E. tenella* XM 013372616.1	36.87	9.68	0	52.14
DUS *E. necatrix* XM 013577463.1	35.62	10.76	0	52.84
DUS *E. maxima* XM 013481802.1	29.61	11.16	0	57.20
DUS *T. gondii* XM 002371431.2	19.38	13.45	0	67.17
DUS *B. besnoiti* XM 029364201.1	39.55	6.11	0	52.73

gondii XM 002371431.2 has the lowest at only 19.38%. Extended strands are often associated with β-sheet structures, contributing to protein stability. DUS *T. gondii* XM 002371431.2 has the highest content of extended strands at 13.45%. Random coils are typically linked with the flexibility of proteins. DUS *T. gondii* XM 002371431.2 has the highest content of random coils at 67.17%, suggesting this protein may be more flexible and adaptable in function.

Regarding the predicted *E. media* DUS protein, it has an α-helix content of 32.01% and a random coil content of 40.56%. This indicates that the protein structure has both stability (contributed by α-helices) and a degree of flexibility (contributed by random coils). Such structural features are likely directly related to the function of the DUS protein, enabling it to perform effectively in specific biological processes.

Comparative analysis shows that the *DUS* gene in *E. media* exhibits a compact and complex tertiary structure with multiple interacting regions (Fig. 8). The *DUS* gene structures in several *Eimeria* species, such as *E. tenella* and *E. acervulina*, are similar to that of *E. media*, indicating a high degree of conservation in the critical functional regions of the *DUS* gene across these species. These structures highlight different functional domains of the protein, and the arrangement and shape of these domains are likely related to their specific roles in tRNA modification. The preservation of these similar domains' relative positions and shapes across different species suggests their importance in gene function and their conservation through evolution.

However, the *DUS* gene structure in *T. gondii* is notably different, being more elongated and less compact, which may reflect differences in substrate specificity or interaction mechanisms. The structures of *E. maxima* and *B. besnoiti* are similar to that of *E. media*, but with minor differences, potentially indicating species-specific regulatory mechanisms and adaptations in the evolution of the *DUS* gene function.

The compactness and complexity of these tertiary structures imply that protein folding plays a crucial role in their functional execution. Compact structures are generally associated with efficient enzyme activity and stable protein conformations. Species with more similar structures, such as *E. media* and *E. tenella*, may have closer evolutionary relationships since the tertiary structure of proteins directly reflects their amino acid sequences and the chemical bonds formed.

Overall, these tertiary structures demonstrate the conservation and adaptability of the *DUS* gene across different species, highlighting its critical role in tRNA modification and providing important insights for further research into the function of these genes in parasite biology and the development of antiparasitic drugs.

Discussion

Dihydrouridine Synthase (DUS) is a key RNA modification enzyme, specifically responsible for catalyzing the transformation of uridine (U) to dihydrouridine (D) in tRNA [24]. This transformation occurs through an NADPH-dependent reduction process, which is crucial for the structure and function of tRNA, affecting the efficiency and accuracy of protein synthesis [25]. In-depth studies on the structure and function of DUS enzymes in specific organisms provide insights into their stability in different environments and their catalytic substrate recognition and mechanisms [5]. Computer simulations and structure predictions offer insights into the three-dimensional structure of DUS enzymes and its potential as a new target for future antibiotic drug design [26]. Additionally, recent studies have discovered the presence of dihydrouridine modifications in mRNA, expanding the known functional scope of *DUS* [24] including potential impacts on cell proliferation, protein translation rates, and chromosome separation. In addition, certain DUS enzymes are associated with the progression of human cancers [27]. In the field of microbiology, the establishment of an in vitro *DUS* activity testing system for *Thermus thermophilus* has enabled direct observation of how *DUS* modifications affect the structure and function of tRNA [28]. The study confirmed the importance of such modifications in maintaining the correct folding of tRNA and its involvement in the protein synthesis process. In terms of interactions with hosts, identified secretory proteins and their interactions with hosts provide new insights, especially in infection and survival strategies [6]. Finally, phylogenetic analysis offers valuable insights into the evolution and diversity of DUS enzymes across the tree of life [29]. Collectively, these studies comprehensively reveal the complex functions and roles of dihydrouridine and DUS enzymes in living organisms.

Previously, our laboratory generated PL of *E. media*, which exhibits a shortened lifecycle, reduced pathogenicity, and weakened reproductive capability. In this study, we identified *DUS* in both the WT and PL strains of *E. media*. This is the first time that *DUS* has been identified in a coccidia species. Alignment of the coding sequence (CDS) of the *DUS* gene revealed subtle base differences between the precocious and WT *Eimeria DUS* genes, and qPCR analysis showed a markedly higher expression of *DUS* by the PL strain than by the WT strain. These findings suggest that *DUS* may play an important role in the reproduction and development of *E. media*. The *DUS* gene may play a more significant role in the lifecycle of the PL strain than the WT strain. Such differences in expression provide crucial insights into the biological function of the *DUS* gene in different growth states of *Eimeria*.

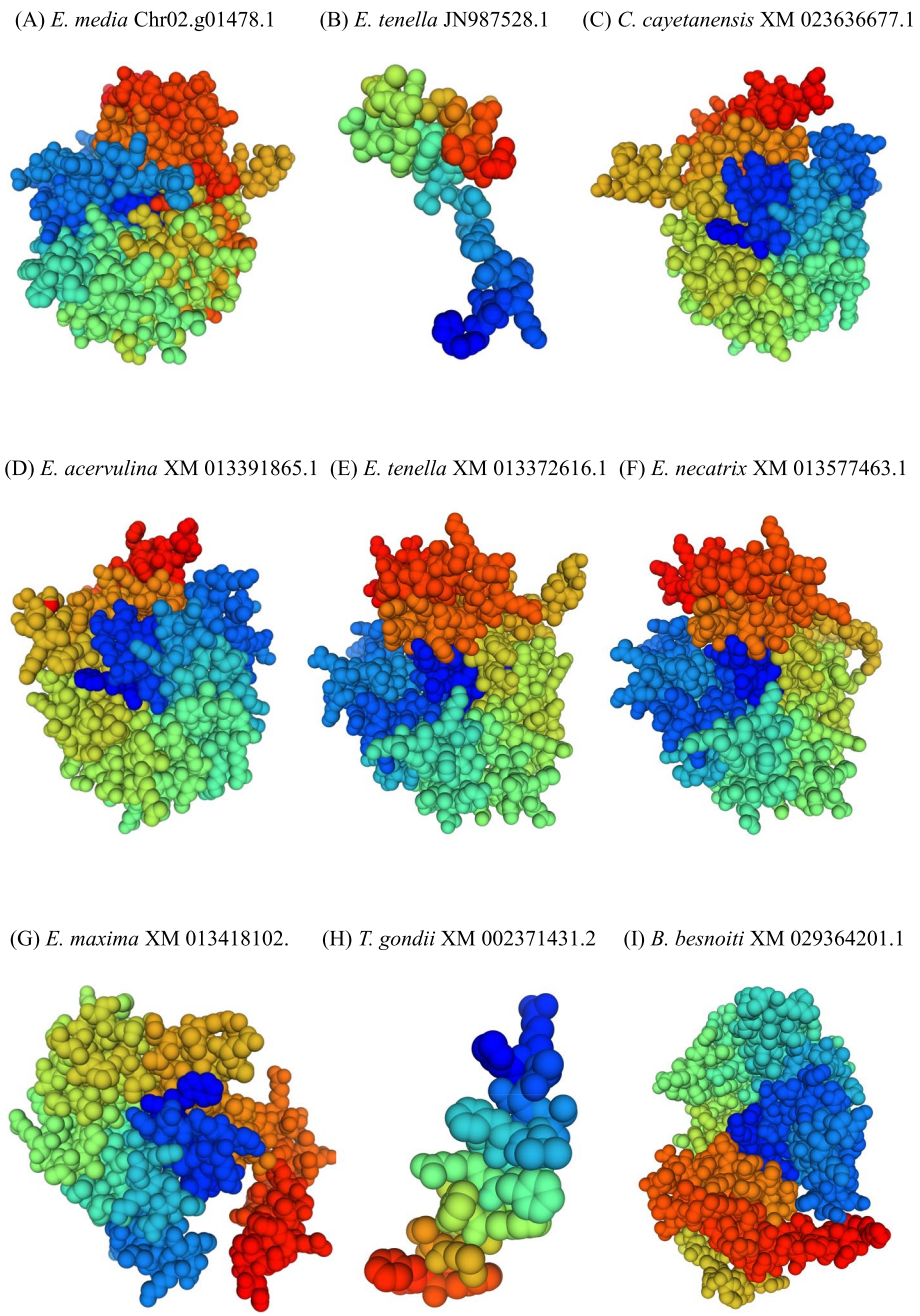

Fig. 8 Predicted protein tertiary structures

Through multiple sequence alignments, genes highly homologous to the reference genome were selected, and highly conserved amino acid positions and variant positions were highlighted. Phylogenetic tree construction and motif analysis presented the evolutionary and functional patterns of the *DUS* gene and its encoded protein in different organisms. Particularly, based on the physicochemical properties of the protein, the DUS protein appears to be of medium size, hydrophilic, and exhibits specific isoelectric point and instability, providing crucial clues for understanding its function and behavior within cells. Finally, combined with secondary structure predictions and tertiary structure modeling, we gained a comprehensive understanding of the structural characteristics of the DUS protein and its potential function in RNA editing.

This research has some limitations. Primarily, the samples mainly come from rabbits infected with *E. media*,

which may imply that the results primarily reflect the specific interaction between this host and pathogen, potentially limiting their applicability to other hosts. Additionally, our study focused on the structural and sequence analysis of the *DUS* gene, with limited exploration of its biological functions. Future research could further enhance understanding of the *DUS* by introducing a broader range of samples, employing diverse analytical techniques, conducting in-depth functional experiments, and considering additional environmental factors.

Conclusion

This study confirmed the presence of the *DUS* gene in *Eimeria* oocysts and increased expression in the PL of *E. media* compared to the WT parasites. The finding suggests a possible correlation between *DUS* gene expression levels and the precocious phenotype. No sequences matching the identified *DUS* gene were found in the NCBI database, though the sequence showed the highest homology with the DUS-like gene of *E. tenella*. Bioinformatics analysis provided new perspectives on the evolution and structural–functional aspects of the *DUS* gene in diverse organisms.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12917-024-04362-8.

Additional file 1.

Additional file 2.

Additional file 3.

Additional file 4.

Additional file 5.

Additional file 6.

Additional file 7.

Acknowledgements

Special thanks to Professor Jin Zhu for the valuable assistance in manuscript revision. We are grateful for the academic training platform provided by the College of Veterinary Medicine, China Agricultural University, and the unwavering support from the National Animal Protozoa Laboratory. Our appreciation extends to the rabbit industry, and we offer my utmost respect to the rabbits involved in this study. Finally, We would like to thank our own perseverance and passion for scientific research.

Authors' contributions

Yijin Zou conducted the experiments, collecting samples, analyzing data, conducting bioinformatics analysis, writing the manuscript, and participating in manuscript revision. Yiyan Wang and Diyi Duan assisted in the experiment. Xun Suo and Yuanyuan Zhang conceived the study, designed the experiments, and contributed to the revision of the manuscript. All authors reviewed the manuscript.

Funding

This study was supported by the National Key Research and Development Program of China (2023YFD1802403). Additionally, this research was funded by a university-level research project at China Agricultural University (Project No. X2022100190082).

Data availability

Data is provided within the manuscript or supplementary information files.

Declarations

Ethics approval and consent to participate

The China Agricultural University's Committee for Laboratory Animal Welfare and Ethical Examination of Animal Experiments approved the experiments. All animal tests complied with the Chinese National Laboratory Animal Guidelines (GB 14925–2010/XG1-2011).

Consent for publication

All animal procedures were performed in accordance with the ethical standards of the institution at which the studies were conducted (China Agricultural University). This study did not involve any animal owners as all animals were bred and maintained under the control of the university.

Competing interests

The authors declare no competing interests.

Received: 26 June 2024 Accepted: 28 October 2024
Published online: 06 January 2025

References

1. Bould JG, Elsheikha HM, Morsy TA. Avian coccidiosis: the basic pathology to control. J Egypt Soc Parasitol. 2009;39(1):85–98.
2. Stuart BP, Lindsay DS. *Coccidiosis* in swine. Vet Clin North Am Food Anim Pract. 1986;2(2):455–68. https://doi.org/10.1016/s0749-0720(15)31256-1.
3. Madlala T, Okpeku M, Adeleke MA. Understanding the interactions between *Eimeria* infection and gut microbiota, towards the control of chicken coccidiosis: a review. Parasite. 2021;28:48. https://doi.org/10.1051/parasite/2021047.
4. Burrell A, Tomley FM, Vaughan S, Marugan-Hernandez V. Life cycle stages, specific organelles and invasion mechanisms of *Eimeria* species. Parasitology. 2020;147(3):263–78. https://doi.org/10.1017/S0031182019001562.
5. Goyal N, Chandra A, Qamar I, Singh N. Structural studies on dihydrouridine synthase A (DusA) from *Pseudomonas aeruginosa*. Int J Biol Macromol. 2019;132:254–64. https://doi.org/10.1016/j.ijbiomac.2019.03.209.
6. Yu F, Tanaka Y, Yamashita K, Suzuki T, Nakamura A, Hirano N, Suzuki T, Yao M, Tanaka I. Molecular basis of dihydrouridine formation on tRNA. Proc Natl Acad Sci U S A. 2011;108(49):19593–8. https://doi.org/10.1073/pnas.1112352108.
7. Dewe JM, Whipple JM, Chernyakov I, Jaramillo LN, Phizicky EM. The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. RNA. 2012;18(10):1886–96. https://doi.org/10.1261/rna.033654.112.
8. Kato T, Daigo Y, Hayama S, Ishikawa N, Yamabuki T, Ito T, Miyamoto M, Kondo S, Nakamura Y. A novel human tRNA-dihydrouridine synthase involved in pulmonary carcinogenesis. Cancer Res. 2005;65(13):5638–46. https://doi.org/10.1158/0008-5472.CAN-05-0600.
9. Gu X, Liu H, Li C, Fang S, Cui P, Liao Q, Zhang S, Wang S, Duan C, Yu F, Suo X, Liu X. Selection and characterization of a precocious line of *Eimeria media*. Parasitol Res. 2019;118(10):3033–41. https://doi.org/10.1007/s00436-019-06422-7.
10. Long PL, Millard BJ, Joyner LP, Norton CC. A guide to laboratory techniques used in the study and diagnosis of avian coccidiosis. Folia Vet Lat. 1976;6(3):201–17.
11. Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, Comeau DC, Funk K, Kim S, Klimke W, Marchler-Bauer A, Landrum M, Lathrop S, Lu Z, Madden TL, O'Leary N, Phan L, Rangwala SH, Schneider VA, Skripchenko Y, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2021;49(D1):D10–7. https://doi.org/10.1093/nar/gkaa892.
12. Chen C, Wu Y, Li J, Wang X, Zeng Z, Xu J, Liu Y, Feng J, Chen H, He Y, Xia R. TBtools-II: A "one for all, all for one" bioinformatics platform for biological

big-data mining. Mol Plant. 2023;16(11):1733–42. https://doi.org/10.1016/j.molp.2023.09.010.
13. Li KB. ClustalW-MPI: ClustalW analysis using distributed and parallel computing. Bioinformatics. 2003;19(12):1585–6. https://doi.org/10.1093/bioinformatics/btg192.
14. Gouet P, Robert X, Courcelle E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. 2003;31(13):3320–3. https://doi.org/10.1093/nar/gkg556.
15. Tamura K, Stecher G, Kumar S. MEGA11: Molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38(7):3022–7. https://doi.org/10.1093/molbev/msab120.
16. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37(Web Server issue):W202–8. https://doi.org/10.1093/nar/gkp335.
17. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 2011;39(Database issue):D225–9. https://doi.org/10.1093/nar/gkq1189.
18. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003;31(13):3784–8. https://doi.org/10.1093/nar/gkg563.
19. Deléage G. ALIGNSEC: viewing protein secondary structure predictions within large multiple sequence alignments. Bioinformatics. 2017;33(24):3991–2. https://doi.org/10.1093/bioinformatics/btx521.
20. Zaib S, Akram F, Liaqat ST, Altaf MZ, Khan I, Dera AA, Uddin J, Khan A, Al-Harrasi A. Bioinformatics approach for the construction of multiple epitope vaccine against omicron variant of SARS-CoV-2. Sci Rep. 2022;12(1):19087. https://doi.org/10.1038/s41598-022-23550-w.
21. Poskus T, Sabonyte-Balsaitiene Z, Jakubauskiene L, Jakubauskas M, Stundiene I, Barkauskaite G, Smigelskaite M, Jasiunas E, Ramasauskaite D, Strupas K, Drasutiene G. Preventing hemorrhoids during pregnancy: a multicenter, randomized clinical trial. BMC Pregnancy Childbirth. 2022;22(1):374. https://doi.org/10.1186/s12884-022-04688-x.
22. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4(4):406–25. https://doi.org/10.1093/oxfordjournals.molbev.a040454.
23. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci. 1992;8(3):275–82. https://doi.org/10.1093/bioinformatics/8.3.275.
24. Brégeon D, Pecqueur L, Toubdji S, Sudol C, Lombard M, Fontecave M, de Crécy-Lagard V, Motorin Y, Helm M, Hamdane D. Dihydrouridine in the transcriptome: new life for this ancient RNA chemical modification. ACS Chem Biol. 2022;17(7):1638–57. https://doi.org/10.1021/acschembio.2c00307.
25. Rider LW, Ottosen MB, Gattis SG, Palfey BA. Mechanism of dihydrouridine synthase 2 from yeast and the importance of modifications for efficient tRNA reduction. J Biol Chem. 2009;284(16):10324–33. https://doi.org/10.1074/jbc.M806137200.
26. Whelan F, Jenkins HT, Griffiths SC, Byrne RT, Dodson EJ, Antson AA. From bacterial to human dihydrouridine synthase: automated structure determination. Acta Crystallogr D Biol Crystallogr. 2015;71(Pt 7):1564–71. https://doi.org/10.1107/S1399004715009220.
27. Li Z, Yin C, Li B, Yu QY, Mao WJ, Li J, Lin JP, Meng YQ, Feng HM, Jing T. DUS4L silencing suppresses cell proliferation and promotes apoptosis in human lung adenocarcinoma cell line A549. Cancer Manag Res. 2020;12:9905–13. https://doi.org/10.2147/CMAR.S265671.
28. Kusuba H, Yoshida T, Iwasaki E, Awai T, Kazayama A, Hirata A, Tomikawa C, Yamagami R, Hori H. In vitro dihydrouridine formation by tRNA dihydrouridine synthase from *Thermus thermophilus*, an extreme-thermophilic eubacterium. J Biochem. 2015;158(6):513–21. https://doi.org/10.1093/jb/mvv066.
29. Lombard M, Reed CJ, Pecqueur L, Faivre B, Toubdji S, Sudol C, Brégeon D, de Crécy-Lagard V, Hamdane D. Evolutionary diversity of Dus2 enzymes reveals novel structural and functional features among members of the RNA dihydrouridine synthases family. Biomolecules. 2022;12(12):1760. https://doi.org/10.3390/biom12121760.

Publisher's Note