Identification and Comparison of Hyaluronic Acid Biosynthetic Genes from Different Capsular Types of Pasteurella Multocida

TEERASAK E-KOBON (✉ fscitse@ku.ac.th)
Kasetsart University Faculty of Science  https://orcid.org/0000-0002-3919-9841

Pailin Pasomboon
Kasetsart University Faculty of Science

Pramote Chumnanpuen
Kasetsart University Faculty of Science

Research article

Keywords: Pasteurella multocida, hyaluronic acid, capsular biosynthetic pathway, capsular polysaccharides, bioinformatics

DOI: https://doi.org/10.21203/rs.3.rs-22706/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

*Pasteurella multocida* produces a capsule composed of different polysaccharides according to the capsular serotype (A, B, D, E, and F). Hyaluronic acid (HA) is a component of certain capsular types of this bacterium, especially capsular type A. Previously, two HA biosynthetic genes from a capsular type A strain were studied for the industrial-scale improvement of HA production. Molecular comparison of these genes across different capsular serotypes of *P. multocida* has not been reported. This study aimed to compare nine HA biosynthetic genes (*glck*, *pgi*, *pgm*, *galU*, *hyaC*, *glmS*, *glmM*, *glmU*, and *hyaD*) of eleven *P. multocida* strains (A:B:D:F = 8:1:1:1) with those of other organisms using sequence and structural bioinformatics analyses.

Results

These nine genes showed a high level of within-species similarity (98–99%) compared to other organisms. Only the last gene of two strains with capsular type A:3 (PM70 and CRIMBP-0884) and one capsular type F strain (HN07) significantly differed from those of other strains (82%). Analysis of amino acid patterns together with phylogenetic results showed that the HA biosynthetic genes of the type A and D strains were closely related compared to those of the type B and F strains. However, the genes in the capsular type F strain were notably similar to those of the capsular type A:3 strain. Protein structural analysis supported structural similarities of the encoded enzymes between the strains of capsular types A, B, D, and F, except for the Glck, Pgm, GlmU and HyaD proteins.

Conclusion

Our bioinformatics analyses proposed that variations observed within these genes could be useful for genetic engineering-based improvement of hyaluronic acid production.

Background

Hyaluronic acid (HA) or hyaluronan is a linear repeat of glucuronic acid and N-acetylglucosamine found in various animal parts, i.e., rooster combs, vitreous humour in eyes, umbilical cords, skin, and cartilage [1]. HAs are high molecular mass molecules, usually more than a million Daltons, with viscoelastic properties that can maintain elasticity and moisture, reduce inflammation, and lubricate the movement of various body parts. It has been used for various biomedical applications; for example, as a diagnostic marker for cancer, rheumatoid arthritis and liver pathologies, reducing inflammation of the wound, and in drug delivery [2, 3]. HAs are also applied in certain ophthalmological and ontological surgeries, cosmetic regeneration, and soft tissue reconstruction as biocompatible and non-immunogenic materials [4]. Moreover, low molecular mass HAs are involved in wound healing, angiogenesis, cell differentiation, tumour cell migration and apoptosis [5, 6].
Some bacteria, including *Streptococcus* sp. and *Pasteurella multocida*, produce HAs as part of their capsule and slime [1]. *P. multocida* is a gram-negative bacterium that causes various diseases in livestock, including avian cholera, respiratory diseases, septicaemia, and atrophic rhinitis [7]. This bacterium produces a capsule consisting of different polysaccharide compositions according to their capsular serotypes, including hyaluronic acid in serotype A, heparin in serotype D, and chondroitin sulfate in serotype F [8–10]. The capsule of type B contains arabinose, mannose, and galactose, while the content of capsular type E remains unclear [11]. *P. multocida* serotype A produces a capsule containing HA similar to *Streptococcus* sp., algae, viruses and vertebrates [12, 13]. HA biosynthesis in these organisms involves nine genes, which encode glucokinase (*glck*), glucose-6-phosphate isomerase (*pgi*), phosphoglucomutase (*pgm*), UDP-glucose-1-phosphate uridylyltransferase (*galU*), UDP-glucose 6-dehydrogenase (*hyaC*), L-glutamine:D-fructose-6-phosphate aminotransferase (*glmS*), phosphoglucoamine mutase (*glmM*), bifunctional N-acetylglucosamine-1-phosphate uridylyltransferase/glucosamine-1-phosphate acetyltransferase (*glmU*), and hyaluronan synthase (*hyaD*) [12, 14].

HA is usually synthesized by two distinct pathways that synthesize HA precursors; these pathways begin with the phosphorylation of glucose by glucokinase to produce two precursors (UDP-glucuronic acid and N-acetylglucosamine). In the first reaction, phosphoglucomutase (Pgm) converts glucose-6-phosphate to glucose-1-phosphate, after which the phosphate group from UTP is transferred to glucose-1-phosphate via UTP-glucose-1-phosphate uridylyltransferase (GalU) to produce UDP-glucose. UDP-glucose is oxidized by UDP-glucose 6-dehydrogenase (HyaC), deriving the first HA precursor, UDP-glucuronic acid. In the second pathway, glucose-6-phosphate is converted to fructose-6-phosphate by glucose-6-phosphate isomerase (Pgi) and then changed to fructose-6-phosphate by adding an amino group from a glutamine residue via L-glutamine:D-fructose-6-phosphate aminotransferase (GlmS) to produce glucosamine-6-phosphate, which is later modified by phosphoglucoamine mutase (GlmM) to yield glucosamine-1-phosphate. Glucosamine-1-phosphate is acetylated and phosphorylated by bifunctional N-acetylglucosamine-1-phosphate uridylyltransferase/glucosamine-1-phosphate acetyltransferase (GlmU) to the second precursor, UDP-N-acetylglucosamine. Finally, hyaluronan synthase (HyAD) combines and polymerizes the two precursors, generating the HA polymer [15, 16]. Recent reviews by Peng *et al.* (2019) showed that the capsule biosynthesis genes in *cap* loci varied among different capsular types, particularly between types A, D, and F and type B. Comparative analysis of the HA biosynthetic genes from various organisms showed that the hyaluronan synthase from *P. multocida* was the only class II enzyme, while other organisms produced the class I hyaluronan synthase [17]. Class II hyaluronan synthases are different from class I hyaluronan synthases in terms of structural topology and biosynthesis mechanism [18, 19]. In this step, the HA synthesized by class II hyaluronan synthases is connected to the cytoplasmic membrane and transported through the membrane directly to the extracellular matrix during chain elongation and not sulfated or chemically modified after biosynthesis [17, 20, 21].

Previous research improved HA production by genetical modification of the *has* operon in *Streptococcus* sp. However, production has a risk of pathogenic bacterial contamination [22]. HA synthesis has been
studied in many other microorganisms (Bacillus sp., Lactococcus sp., and Escherichia. coli) for the production of high-quality and safe-to-use HAs [16, 19]. Industrial-scale HA production was developed with recombinant E. coli and Bacillus sp. strains harbouring the hyaD genes from Streptococcus sp. and P. multocida [23, 24]. The recombinant B. subtilis strain produced HasA from Streptococcus equisimilis together with the co-expression of TuaD from B. subtilis. This recombinant bacterium produced a high level of HAs and had been used industrially [14]. Another study optimized conditions for HA production in recombinant E. coli based on the expression of the hyaD gene of P. multocida ATCC 15742 (capsular type A:3) and the hyac gene of E. coli strain K5 [1]. The level of HA production in the recombinant bacterium increased (2.7 to 3.7 g/L (37%)) after varying the amount of supplemented oxygen and glucosamine. The synthesis also increased by up to 70% when fosfomycin was added to inhibit cell wall synthesis [1]. Co-expression of the hyaC and hyaD genes from P. multocida gave the highest level of HAs (5.4 g/L) and stability compared to the co-expression of the hyaC gene from E. coli and hyaD from S. pyogenes [13]. Although variations in the capsule biosynthesis genes of P. multocida have been studied, HA biosynthetic genes have not been explored and could be useful for further improving HA production. Therefore, this study aimed to compare the HA biosynthesis genes in P. multocida using sequence, pattern and structural analyses. The HA genes were compared across different capsular types of P. multocida and to their orthologues in other bacteria and humans. We proposed that variations observed within these genes could be useful targets for genetic engineering-based improvement of HA production.

Results

Identification of the hyaluronic acid biosynthetic genes of Pasteurella multocida

Eight HA biosynthetic genes were identified in 11 strains of Pasteurella multocida, while hyaluronan synthase (hyaD) was absent in strain HN06 of capsular type D. These nine HA biosynthetic genes shared a high percentage of identity (between 82–99%) and percentage of sequence coverage (between 99–100%) with the query sequences from Mao et al. (2009) (Fig. 1). The lowest percentage of identity (82%) was observed in the hyaC and hyaD genes. Comparison of these genes with those of other organisms showed a lower percentage of identity and coverage: 29–100% and 0–79% in H. sapiens, 25–100% and 0–94% in S. pyogenes strain M1, and 35–100% and 1–100% in S. thermophilus. Among these, the enzymes from E. coli shared the highest percentage of identity (32–93%) and coverage (0–100%) with those of P. multocida, and the hyaD gene was absent in E. coli. The results were used to reconstruct the HA biosynthetic pathway of P. multocida, as depicted in Fig. 1.

Sequence similarities and amino acid patterns of the nine HA biosynthetic genes
Comparison of the nucleotide and amino acid sequences of these nine HA biosynthetic genes from 11 strains of *P. multocida* showed a high level of within-species identity (98–99%) when using the ATCC43137 strain of capsular type A as a reference. Only the *hyaD* gene of strains with capsular type A:3 (PM70 and CRIMBP-0884) and capsular type F (HN07) differed significantly from those of other strains (82%), as shown in Table 1.

| Strain       | Capsular type | Host    | Sequence comparison | Nucleotide | Protein | E-value | % Identity | E-value | % Identity |
|--------------|---------------|---------|---------------------|------------|---------|---------|------------|---------|------------|
| ATCC43137    | A             | Swine   |                     | 0          | 100.00  | 0       | 100.00     |         |            |
| 3480         | A             | Swine   |                     | 0          | 99.79   | 0       | 99.79      |         |            |
| 36950        | A             | Bovine  |                     | 0          | 99.28   | 0       | 99.28      |         |            |
| HB01         | A             | Bovine  |                     | 0          | 100.00  | 0       | 100.00     |         |            |
| HB03         | A             | Swine   |                     | 0          | 100.00  | 0       | 100.00     |         |            |
| FDAARGOS_218 | A:3           | Turkey  |                     | 0          | 99.76   | 0       | 99.90      |         |            |
| PM70         | A:3           | Avian   |                     | 0          | 82.91   | 0       | 82.91      |         |            |
| CIRMBP-0884  | A:3           | Rabbit  |                     | 0          | 82.91   | 0       | 82.91      |         |            |
| PMTB2.1      | B:2           | Bovine  |                     | 0          | 99.79   | 0       | 99.79      |         |            |
| HN06         | D             | Swine   | N/A                 | N/A        | N/A     | N/A     | N/A        |         |            |
| HN07         | F             | Swine   |                     | 0          | 82.91   | 0       | 82.91      |         |            |

N/A = Not available.

Pattern analysis of the amino acid sequences of these nine HA biosynthetic genes in *P. multocida* showed that capsular types A and D had a close relationship compared to capsular types B and F. Particularly, capsular type A strain 3480 had the most different pattern among other type A strains and had five proteins (Pgi, GalU, GlmS, GlmM, and GlmU) closely related to the capsular type D strain HN06 (Fig. 2). The PMTB2.1 strain of capsular type B had four enzymes that were different from the enzymes of other strains, except GalU and HyaC, which were shared in some type A strains, GlmM, which was shared with type A:3 (FDAARGOS_218) and A:3 (PM70) strains, and HyaD, which was similar to the type A strain (Fig. 2). Seven proteins of the capsular type F strain were shared with the capsular type A:3
strain, and three were shared with the type A. This result showed that the HA biosynthetic enzymes across the examined *P. multocida* strains, even within the same serotype, had different variation patterns.

**Phylogenetic relationship of the nine HA biosynthetic genes**

Phylogenetic analysis of the HA biosynthetic genes in Fig. 3 shows the separation between these nine enzymes from *P. multocida* and those from other organisms. The *glcK* and *pgi* genes of capsular types A:3 and F were closely clustered, while the A:3 and F strains were only grouped in the trees of the *pgm*, *glmM*, and *hyaD* genes. The *galU* and *glmU* genes of the A:3 strains were clustered with those of the capsular type B strains. All these genes from most of the remaining capsular type A strains were grouped, except the *pgm*, *gluA*, *glmS*, *glmM*, and *glmU* genes of strain 3480, which were together with those of the capsular types B or D strains and the *hyaC* gene of strain 3480, which was grouped with the type A:3 strains.

The patterns were changed in the protein tree of these nine enzymes (Fig. 4). The phylogenetic tree from the amino acid sequences of the capsular type A strains remained similar to the nucleotide sequence phylogenetic tree. The Pgi, Pgm, and GalU proteins of the capsular type D strains were grouped with those of the capsular type A:3 strains. In contrast, the GlmS, GlmM, and GlmU proteins were grouped with capsular type A. Six proteins (Pgi, Pgm, HyaC, GlmM, GlmU, and HyaD) of capsular type F were clustered together with capsular type A:3, while GlcK and GlmS were grouped with capsular type A:3 strain FDAARGOS_218. GalU was grouped together with the other capsular type A strain (HB03). Five enzymes (Pgi, GalU, HyaC, GlmU, and HyaD) of capsular type B:2 were closely related to capsular type A, while GlmM was closely related to capsular type A:3. Phylogenetic trees from the concatenated nucleotide and protein sequences of the HA biosynthetic genes in Fig. 5 showed the same clustering pattern of the capsular type A:3 and F strains and the four type A strains (HB01, HB03, 36950, and ATCC43137). The patterns of the gene and protein trees of the capsular type A (3480), A:3 (FDAARGOS_218), and B:2 strains differed (Fig. 5A and 5B). These phylogenetic results depicted sequence variation and possibly mutational history within the HA biosynthetic genes of different *P. multocida* strains.

**Structural comparison of the hyaluronic acid biosynthetic enzymes**

The structures of nine HA biosynthetic enzymes from different *P. multocida* strains were successfully predicted and compared. Predicted structures of the Pgi, GalU, HyaC, GlmS, and GlmM proteins were highly similar between capsular types A, B, D, and F (Fig. 6). Structural comparison of the GlcK, Pgm, GlmU, and HyaD proteins showed variations, particularly GlcK of the type A strain HB01, Pgm of the type A:3 strain FDAARGOS_218, GlmU of the types A strain 3480, type A:3 strain FDAARGOS_218, and type D strain HN06, and HyaD of all strains. The structural variations of the HyaD proteins were divided into three patterns: (i) capsular types A and B, (ii) capsular types A and A:3, including strains 36950, 3480, and FDAARGOS_218, and (iii) capsular types A:3 and F, including strains CIRMP-0884, PM70, and HN07.
Comparing the binding site pocket of the predicted structures in Fig. 7 could provide clues into the molecular functions. The binding site pockets were predicted, and different pocket surfaces were observed in the structures of Glck from strain HB01; Pgm from strain FDAARGOS_218; GlmU from strains 3480, FDAARGOS_218, and HN06; and HyaD from strains 3480, 36950, FDAARGOS_218, PM70, CRIMBP-0884, and HN70. Certain positions correlated with the active site of the enzymes determined by a sample score greater than 0.50 (Table S1) A score of less than 0.50 showed no relation to the active site of the enzymes. From all predicted binding pockets, the catalytic pocket of GlmU at the P6 position (Fig. 7) was observed only in the capsular type A strain ATCC43137. The catalytic pocket of HyaD at the P5 position was unique to the capsular type A strains ATCC43137, 3480, and 36950 and capsular type A:3 strain FDAARGOS_218.

Analysis of the conserved domains of these nine enzymes was performed by comparing the domain size and architecture of these proteins using the domains of strain ATCC43137 as a reference. Differences were found in four enzymes, as shown in Fig. 8. The size of the ROK family domain of the Glck protein (Fig. 8A) of strain HB01 was longer than those of other strains. The nucleotidyl transferase (NTP_transferase) and fucokinase domains of the GlmU protein (Fig. 8B) of strain ATCC43137 were different in length (number of amino acid residues) when compared with those of other strains (Table S3), while the hexapeptide repeat of the succinyl-transferase domain (Hexapep_2), the unknown Duf4954, and fucokinase domain of strain FDAARGOS_218 differed from those of other strains in terms of position (Table S3). The domains of strain FDAARGOS_218 were slightly different in length and position compared to those of the reference strain ATCC43137 and other strains (including Hexapep_2, Duf4954, and fucokinase domain). Several changes were observed in the domains of the HyaD proteins (Fig. 8C). The domains of the HyaD proteins of strains 3480 and 36950 were similar to those of other strains; only the domain sizes were different. The strain FDAARGOS_218 had an additional domain of unknown function (Duf2536), while other strains lacked this domain. The TPR_2 (tetratricopeptide repeat) repeat (white box in Fig. 8C) was also absent in the strains PM70, CRIMBP-0884, and HN70. Glycosyl transferase-like family 2 (Glycos_transd_2_2) and glycosyl transferase family 2 (Glycos_transd_2_4) were more abundant in the PM70, CIRMBP-0884, and HN07 strains than in the other strains. Glycos_transd_2_2 also showed differences in sizes and positions between strains (Table S4).

Discussion

The hyaluronic acid biosynthesis pathway of P. multocida involves nine enzyme-encoding genes, including glck, pgi, pgm, galU, hyaC, glmS, glmM, glmU and hyaD, similar to the capsular polysaccharide biosynthesis pathway in S. pyogenes [19, 25–27]. HAs are known to be primary capsular components of type A strains of P. multocida. This study found that P. multocida strains of capsular types A, B, D, and F shared eight-core HA biosynthetic enzymes that produce UDP-D-glucuronic acid and UDP-N-acetyl-D-glucosamine, which are modified to hyaluronic acid (hyaD), chondroitin sulfate, and heparin, and disaccharide repeats of other components that vary across different strains [28]. The hyaluronan synthase (hyaD) of capsular type A is essential in the joining and polymerization of the two precursors for HA biosynthesis. In the capsular type F strain, N-acetyl-D-galactosamine was used instead of N-acetyl-D-
glucosamine to generate chondroitin sulfate (CS) by chondroitin synthase [29, 30]. Heparin (HP) was made from three precursors (D-glucuronic acid, D-glucosamine, and L-iduronic acid (IdoA)). Members of the HA biosynthetic pathway synthesize two of these precursors. Only L-iduronic acid could be synthesized by an epimerization reaction of D-glucuronic acid and polymerization of these precursors to generate heparin by heparin synthase [31, 32]. This result suggested that the first eight enzymes of the HA biosynthesis pathway could be the core pathway for synthesizing other types of capsular polysaccharides. Modification of these enzymes and the pathway allowed the cells to diversify their capsular polysaccharide components.

Comparative sequence analysis of the HA biosynthetic genes from different capsular types showed a high level of within-species similarity because these genes were a member of the capsule biosynthesis locus (cap locus), a gene cluster for capsular polysaccharide biosynthesis [25, 28, 33]. The absence of the hyaD gene would not affect heparin and chondroitin sulfate biosynthesis because the enzyme is not required. Only the hyaD gene of the capsular types A:3 and F differed significantly from that of the other strains, suggesting a functional difference between HA and chondroitin biosynthesis. The disappearance of hyaD in type D could be due to the requirement of a different enzyme to synthesize the heparin capsular component [34]. The protein pattern and phylogenetic analysis supported the sequence alignment of these genes. These results indicated that the HA biosynthetic genes of types A and D had a closer relationship compared to those of types B and F, except hyaD when compared between capsular type A strain 3480 and capsular type D. Although strains with the same capsular type evidently have highly similar genes in the pathway, this study found that the type A strains had diverse patterns of HA biosynthetic genes. The genes of the capsular type F strain shared more similarity to those of the type A:3 strain (previously used in genetically engineered E. coli) than those of the others of type A strains. However, this study was unable to subtype some of the type A samples, and the deviation of type A:3 could potentially affect their enzymatic activities. The HyaD pattern of the type B strain was similar to that of type A, which might indicate the ability to synthesize HA-containing capsules in this capsular type.

Structures of the Glck, Pgm, GlmU, and HyaD proteins showed individual variations across different strains of P. multocida. Analysis of the binding pockets of Glck and Pgm by comparison with the structures of these enzymes in other organisms found that the sites were irrelevant to the catalytic pockets. In contrast, some of the GlmU and HyaD binding pockets were at the catalytic sites. HyaD also contains two active sites: one site has glucuronic acid-transferase activity, which elongates UDP-glucuronic acid to oligosaccharides, and the other site has glucosamine-1-P-transferase activity, which prolongs UDP-N-acetyl-D-glucosamine to oligosaccharides [35]. Surface variation in the binding sites of the HyaD protein in this study was consistent with previous studies that identified the enzymatic binding site in the glucuronic acid-transferase domain of HyaD [35]. Our study found that the arrangement of DUF2536 and TPR_2 repeats in the HyaD protein differed in these P. multocida strains, suggesting the reason for the structural variation of the binding site and perhaps different molecular activities. HyaD and HyaC were also reported to be rate-limiting enzymes for HA biosynthesis in bacteria and humans due to the limited substrate (UDP-D-glucuronic acid) produced by HyaC and the specificity of HyaD to the polymerization of UDP-D-glucuronic acid and UDP-N-acetyl-D-glucosamine [36–38]. Therefore, it is
possible that the HyaD protein of the *P. multocida* strains in this study could have different substrate specificities and catalytic rates.

The protein structure of GlmU was previously studied in *E. coli*, and the catalytic site of this protein was also matched to the variable binding regions predicted from the *P. multocida* strains in this study. These variable regions correlated with N-terminal pyrophosphorylase and C-terminal acetyltransferase catalytic sites [39, 40]. GlmU has two active sites: (i) acetyltransferase, which is responsible for CoA-dependent acetylation of D-glucosamine-1-P to N-acetyl-D-glucosamine-1-P at the C-terminal domain, and (ii) pyrophosphorylase, which catalyses the transfer of uridylic acid from UTP to D-glucosamine-1-P, forming UDP-N-acetyl-D-glucosamine and pyrophosphate at the N-terminal domain [40]. However, compared to previous studies in *E. coli* and *Salmonella enterica*, the binding regions of GlcK and Pgm were not related to the active sites. Nevertheless, the binding regions of these proteins might help to regulate enzyme activity [41, 42]. Consequently, structural variations of these four enzymes could potentially affect molecular function, substrate specificity, and catalytic rate in HA biosynthesis.

The results from all analyses in this study emphasized differences within the enzymes of capsular types A:3, A:1, and A strains. These three types had HA-containing capsules with different somatic serotypes important to virulence [43–45]. These three serotypes may evolve different virulence factor-associated genes as well as different capsule and lipopolysaccharide (LPS) compositions. The somatic serotype was found to be more specific to the disease than the capsule type and quite diverse between strains [46]. Further investigation of the HyaD proteins in these type A strains might help improve industrial HA production, allowing different alternatives to HyaD enzymes, as well as commercial biosynthesis of heparan and chondroitin, where other enzymes related to the HA biosynthetic pathway can be used.

**Conclusion**

This study identified nine genes involved in the HA biosynthetic pathway of *P. multocida* strains. Eight genes had a high level of within-species similarity (98–99%), while *hyaD* of strains with capsular type A:3 and F differed significantly those of other strains. Sequence, phylogenetic and pattern analysis of these genes showed that capsular types A and D had a close relationship compared to capsular types B and F. Structural comparison revealed variations in the GlcK, Pgm, GlmU, and HyaD proteins. The variations within the active sites and domains of HyaD and GlmU could influence enzymatic function and substrate specificity. Thus, these enzymes will be potential targets for improving HA production.

**Materials And Methods**

**Identification of hyaluronic acid biosynthetic genes**

The hyaluronic acid biosynthetic genes from eleven complete genomes of *Pasteurella multocida*, including ATCC43137 (accession number CP008918), 3480 (accession number CP001409), 36950 (accession number CP003022), HB01 (accession number CP006976), HB03 (accession number
CP003328), FDAARGOS_218 (accession number CP020405), PM70 (accession number AE004439), CIRMBP-0884 (accession number CP020345), PMTB2.1 (accession number CP007205), HN06 (accession number CP003313), and HN07 (accession number CP007040), were downloaded from the NCBI genome database. Nine HA biosynthetic genes (glck, pgi, pgm, galU, hyaC, glmS, glmM, glmU, and hyaD) previously described from the study of Mao et al. (2009) [1] were used to search for orthologous nucleotide sequences in the obtained P. multocida genomes using BLASTn [47]. A similar search was performed against the genomes of Homo sapiens, E. coli strain K-12 MG1655, Streptococcus pyogenes strain M1, and S. thermophilus to identify orthologous genes according to previous reports [48, 49], which studied the HA production of these strains, except E. coli, which is often used for the expression of recombinant HA biosynthesis enzymes. Identification of orthologous genes was considered at a percentage identity cut-off above 35, an e-value higher than 10e-10, and relevant functional information. Protein sequences encoded from these genes were obtained from the NCBI protein database.

Sequence comparison and amino acid pattern analysis

Nucleotide and amino acid sequences of the HA biosynthesis genes were edited and aligned by using the MAFFT program version 7 [50] to identify conserved regions across different capsular types of P. multocida, and the sequences were compared to those of other organisms using iterative refinement methods (G-INS-i) and the Needleman-Wunsch algorithm, which extracted the conserved sequences and truncate flanking sequences [51, 52]. Amino acid patterns were analysed from the aligned protein sequences to identify possible variations that could impact the enzymatic function. The pattern was considered different if there was at least one position within the aligned sequence block that differed from other samples at the same position. Numerical patterns were assigned to depict sequence variations of the proteins. This pattern analysis assumed that the orthologous proteins would have a similar molecular function if their amino acid sequences were more similar. A numerical pattern of nine digits was obtained for each sample to represent variations within these nine HA biosynthetic enzymes. These numeric sets were then compared within different P. multocida strains and across different organisms.

Phylogenetic analysis of the hyaluronic acid biosynthetic genes

Phylogenetic analysis was employed to analyse the evolutionary relationship of these nine HA biosynthetic genes from different P. multocida strains and other organisms, including H. sapiens, E. coli, S. pyogenes strain M1 GAS, and S. thermophilus as an outgroup. The phylogenetic relationships were constructed from the aligned nucleotide and amino acid sequences based on the Tamura-Nei and Dayhoff models using the maximum likelihood (ML) and neighbour-joining (NJ) methods with 1,000 bootstraps performed with the MEGA program version 7.0 [53]. The phylogenetic relationships were compared to the amino acid sequence patterns previously described to understand the relatedness of the HA biosynthetic enzymes. The nucleotide and amino acid sequences of these nine genes were edited and concatenated using the AliView program version 3.0 (GPLv3) [54] before the sequence alignment and phylogenetic analysis as previously described.
Structural comparison of the hyaluronic acid biosynthetic enzymes

Protein structures of these nine HA biosynthetic enzymes from different *P. multocida* strains were predicted using SWISS-MODEL [55]. The best-predicted protein structures were selected based on Qualitative Model Energy Analysis (QMEAN) [56] and Global Model Quality Estimation (GMQE) values [57]. Scores closer to 1 indicated a high level of reliability of the structural prediction. Protein structural domains were also determined by the Pfam program using amino acid sequences as the input [58]. The program annotated and identified functional domains that could be used to explain the structural differences and perhaps different biochemical activities of these nine enzymes in the HA biosynthetic pathway. Multiple comparisons of the predicted protein structures were performed by calculating the root mean square deviation (RMSD) values and Q scores between the models with the PDBeFold program [59] and visualization with the SWISS-pdb viewer program [60, 61] to identify different structural positions on the aligned structures [62]. Binding site pockets of the predicted protein structures were also compared by the DoGSiteScorer program [63], which used a grid-based method and Gaussian filters to detect potential binding pockets and functional groups present in the pockets. The binding site pockets were analysed in terms of distance-dependent histograms between atomic pairs, size, shape complexity, hydrophobicity, hydrogen bonds, metal co-ordinations, and lipophilic contacts. These various parameters were used to estimate the size and shape descriptors of the binding site pocket (volume (Å³), surface (Å²), and depth (Å)) [64]. These parameters were summarized based on the sample score, which described the properties and capabilities of the structural area. The sample score ranged between 0 for the nonbinding pocket and 1 for the potential binding pocket within the protein structure. Strains that had a high level of structural changes were hypothesized to have different HA biosynthesis capabilities.

**Abbreviations**

*hyaC*, HYAC, UDP-glucose 6-dehydrogenase; *hyaD*, HYAD, Hyaluronan synthase; IdoA, L-iduronic acid; ML, Maximum likelihood methods; NJ, Neighbour-joining methods; *pgi*, PGI, Glucose-6-phosphate isomerase; *pgm*, PGM, Phosphoglucomutase; QMEAN, Qualitative Model Energy Analysis; RMSD, Root Mean Square Deviation

**Declarations**

**Supplementary information**

**Additional file 1: Table S1.** Analysis of the protein pockets from the HA biosynthetic enzymes.

**Additional file 2: Table S2.** Domain description of the Glck protein in *P. multocida*.

**Additional file 3: Table S34.** Domain description of the GlmU protein in *P. multocida*. 

Page 11/25
Acknowledgements

We greatly thank the Department of Genetics and Department of Zoology, Faculty of Science, Kasetsart University, Bangkok, Thailand, and the Bioinformatics and Systems Biology Unit, Computational Biomodelling Laboratory for Agricultural Science and Technology (CBLAST), Faculty of Science, Kasetsart University, Thailand.

Author contribution

Pailin Pasomboon planned the project, carried out the analyses, and wrote the manuscript.
Pramote Chumnanpuensupervised the project, discussed the results, and revised the manuscript.
Teerasak E-kobonplanned the project, verified the analytical methods, discussed the results, and edited the manuscript.

Funding

This research was financially supported by a research grant from Do Day Dream Pub. Co. Ltd.

Availability of data and materials

All data generated or analysed during this study are included in the manuscript and its supplementary information files.

Ethics approval and consent to participate

Not applicable.

Authors’ information

Not applicable.

Consent for publication

Not applicable.
Competing interests

The authors declare that they have no competing interests.

Author details

1 Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand 2 Department of Zoology, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand 3 Computational Biomodelling Laboratory for Agricultural Science and Technology, Kasetsart University, Bangkok 10900, Thailand *Corresponding author

References

1. Mao Z, Shin HD, Chen R. A recombinant E. coli bioprocess for hyaluronan synthesis. Appl Microbiol Biotechnol. 2009;84(1):63–9.
2. Fakhari A, Berkland C. Applications and emerging trends of hyaluronic acid in tissue engineering, as a dermal filler and in osteoarthritis treatment. Acta Biomater. 2013;9(7):7081–92.
3. Tsepilov RN, Beloded AV. Hyaluronic Acid—an "Old" Molecule with "New" Functions: Biosynthesis and Depolymerization of Hyaluronic Acid in Bacteria and Vertebrate Tissues Including during Carcinogenesis. Biochemistry. 2015;80(9):1093–108.
4. Kogan G, Soltes L, Stern R, Gemeiner P. Hyaluronic acid: a natural biopolymer with a broad range of biomedical and industrial applications. Biotechnol Lett. 2007;29(1):17–25.
5. Maharjan AS, Pilling D, Gomer RH. High and low molecular weight hyaluronic acid differentially regulate human fibrocyte differentiation. PLoS One. 2011;6(10):e26078.
6. Rayahin JE, Buhrman JS, Zhang Y, Koh TJ, Gemeinhart RA. High and low molecular weight hyaluronic acid differentially influence macrophage activation. ACS Biomater Sci Eng. 2015;1(7):481–93.
7. Wilson BA, Ho M. Pasteurella multocida: from zoonosis to cellular microbiology. Clin Microbiol Rev. 2013;26(3):631–55.
8. Wilkie IW, Harper M, Boyce JD, Adler B. Pasteurella multocida: diseases and pathogenesis. Curr Top Microbiol Immunol. 2012;361:1–22.
9. Okay S, Kurt Kizildogan A. Comparative genome analysis of five Pasteurella multocida strains to decipher the diversification in pathogenicity and host specialization. Gene. 2015;567(1):58–72.
10. Peng Z, Wang X, Zhou R, Chen H, Wilson BA, Wu B. Pasteurella multocida: Genotypes and Genomics. Microbiol Mol Biol Rev 2019, 83(4).
11. Harper M, Boyce JD, Adler B. The key surface components of Pasteurella multocida: capsule and lipopolysaccharide. Curr Top Microbiol Immunol. 2012;361:39–51.
12. Chong BF, Blank LM, McLaughlin R, Nielsen LK. Microbial hyaluronic acid production. Appl Microbiol Biotechnol. 2005;66(4):341–51.

13. Chu X, Han J, Guo D, Fu Z, Liu W, Tao Y. Characterization of UDP-glucose dehydrogenase from Pasteurella multocida CVCC 408 and its application in hyaluronic acid biosynthesis. Enzyme Microb Technol. 2016;85:64–70.

14. Widner B, Behr R, Von Dollen S, Tang M, Heu T, Sloma A, Sternberg D, Deangelis PL, Weigel PH, Brown S. Hyaluronic acid production in Bacillus subtilis. Appl Environ Microbiol. 2005;71(7):3747–52.

15. Zhang Y, Luo K, Zhao Q, Qi Z, Nielsen LK, Liu H. Genetic and biochemical characterization of genes involved in hyaluronic acid synthesis in Streptococcus zooepidemicus. Appl Microbiol Biotechnol. 2016;100(8):3611–20.

16. Liu L, Liu Y, Li J, Du G, Chen J. Microbial production of hyaluronic acid: current state, challenges, and perspectives. Microb Cell Fact. 2011;10:99.

17. Weigel PH. Functional characteristics and catalytic mechanisms of the bacterial hyaluronan synthases. IUBMB Life. 2002;54(4):201–11.

18. Weigel PH. Hyaluronan Synthase: The Mechanism of Initiation at the Reducing End and a Pendulum Model for Polysaccharide Translocation to the Cell Exterior. Int J Cell Biol. 2015;2015:367579.

19. Sze JH, Brownlie JC, Love CA. Biotechnological production of hyaluronic acid: a mini review. 3 Biotech. 2016;6(1):67.

20. Jing W, DeAngelis PL. Dissection of the two transferase activities of the Pasteurella multocida hyaluronan synthase: two active sites exist in one polypeptide. Glycobiology. 2000;10(9):883–9.

21. Weigel PH, DeAngelis PL. Hyaluronan synthases: a decade-plus of novel glycosyltransferases. J Biol Chem. 2007;282(51):36777–81.

22. Yu H, Stephanopoulos G. Metabolic engineering of Escherichia coli for biosynthesis of hyaluronic acid. Metab Eng. 2008;10(1):24–32.

23. Brown SH, Pummill PE. Recombinant production of hyaluronic acid. Curr Pharm Biotechnol. 2008;9(4):239–41.

24. Jia Y, Zhu J, Chen X, Tang D, Su D, Yao W, Gao X. Metabolic engineering of Bacillus subtilis for the efficient biosynthesis of uniform hyaluronic acid with controlled molecular weights. Bioresour Technol. 2013;132:427–31.

25. Chung JY, Zhang Y, Adler B. The capsule biosynthetic locus of Pasteurella multocida A:1. FEMS Microbiol Lett. 1998;166(2):289–96.

26. Boyce JD, Chung JY, Adler B. Pasteurella multocida capsule: composition, function and genetics. J Biotechnol. 2000;83(1–2):153–60.

27. Wessels MR. Capsular Polysaccharide of Group A Streptococcus. Microbiol Spectr 2019, 7(1).

28. DeAngelis PL, Gunay NS, Toida T, Mao W-j, Linhardt RJ. Identification of the capsular polysaccharides of Type D and F Pasteurella multocida as unmodified heparin and chondroitin, respectively. Carbohyd Res. 2002;337(17):1547–52.
29. Schiraldi C, Cimini D, De Rosa M. Production of chondroitin sulfate and chondroitin. Appl Microbiol Biotechnol. 2010;87(4):1209–20.

30. Mikami T, Kitagawa H. Biosynthesis and function of chondroitin sulfate. Biochim Biophys Acta. 2013;1830(10):4719–33.

31. DeAngelis PL, White CL. Identification and molecular cloning of a heparosan synthase from Pasteurella multocida type D. J Biol Chem. 2002;277(9):7209–13.

32. Prydz K, Dalen KT. Synthesis and sorting of proteoglycans. J Cell Sci. 2000;113(Pt 2):193–205.

33. Boyce JD, Chung JY, Adler B. Genetic organisation of the capsule biosynthetic locus of Pasteurella multocida M1404 (B:2). Vet Microbiol. 2000;72(1–2):121–34.

34. Otto NJ, Green DE, Masuko S, Mayer A, Tanner ME, Linhardt RJ, DeAngelis PL. Structure/function analysis of Pasteurella multocida heparosan synthases: toward defining enzyme specificity and engineering novel catalysts. J Biol Chem. 2012;287(10):7203–12.

35. Kooy FK, Beeftink HH, Eppink MHM, Tramper J, Eggink G, Boeriu CG. Structural and functional evidence for two separate oligosaccharide binding sites of Pasteurella multocida hyaluronan synthase. Advances in Enzyme Research. 2013;01(04):97–111.

36. Edwards JC. The nature and origins of synovium: experimental approaches to.

37. The study of synoviocyte differentiation. Journal of anatomy 1994, 184:493–501.

38. Choi SB, Lew LC, Hor KC, Liong MT. Fe²⁺ and Cu²⁺ increase the production of hyaluronic acid by Lactobacilli via affecting different stages of the pentose phosphate pathway. Appl Biochem Biotechnol. 2014;173(1):129–42.

39. Baggenstoss BA, Harris EN, Washburn JL, Medina AP, Nguyen L, Weigel PH. Hyaluronan synthase control of synthesis rate and hyaluronan product size are independent functions differentially affected by mutations in a conserved tandem B-X7-B motif. Glycobiology. 2017;27(2):154–64.

40. Olsen LR, Roderick SL. Structure of the Escherichia coli GlmU pyrophosphorylase and acetyltransferase active sites. Biochemistry. 2001;40(7):1913–21.

41. Olsen LR, Vetting MW, Roderick SL. Structure of the E. coli bifunctional GlmU acetyltransferase active site with substrates and products. Protein Sci. 2007;16(6):1230–5.

42. Lunin VV, Li Y, Schrag JD, Iannuzzi P, Cygler M, Matte A. Crystal structures of Escherichia coli ATP-dependent glucokinase and its complex with glucose. J Bacteriol. 2004;186(20):6915–27.

43. Mehra-Chaudhary R, Mick J, Tanner JJ, Henzl MT, Beamer LJ. Crystal structure of a bacterial phosphoglucomutase, an enzyme involved in the virulence of multiple human pathogens. Proteins. 2011;79(4):1215–29.

44. Lu YS, Pakes SP, Stefanu C. Capsular and Somatic Serotypes of Pasteurella multocida Isolates Recovered from Healthy and Diseased Rabbits in Texas. JOURNAL OF CLINICAL MICROBIOLOGY. 1983;18(2):292–5.

45. Dillehay DL, Paul KS, DiGiacomo RF, Chengappa MM. Pathogenicity of Pasteurella multocida A:3 in Flemish giant and New Zealand white rabbits. Lab Anim. 1991;25(4):337–41.
46. Chung JY, Wilkie I, Boyce JD, Townsend KM, Frost AJ, Ghoddusi M, Adler B. **Role of capsule in the pathogenesis of fowl cholera caused by Pasteurella multocida serogroup A.** Infect Immun. 2001;69(4):2487–92.

47. Peng Z, Liang W, Wang F, Xu Z, Xie Z, Lian Z, Hua L, Zhou R, Chen H, Wu B. **Genetic and Phylogenetic Characteristics of Pasteurella multocida Isolates From Different Host Species.** Front Microbiol. 2018;9:1408.

48. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.

49. Crater DL, van de Rijn I. **Hyaluronic acid synthesis operon (has) expression in group A Streptococci.** J Biol Chem. 1995;270(31):18452–8.

50. Izawa N, Serata M, Sone T, Omasa T, Ohtake H. **Hyaluronic acid production by recombinant Streptococcus thermophilus.** J Biosci Bioeng. 2011;111(6):665–70.

51. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.

52. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059–66.

53. Katoh K, Kuma K, Miyata T, Toh H. Improvement in the accuracy of multiple sequence alignment program MAFFT. Genome Inform. 2005;16(1):22–33.

54. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33(7):1870–4.

55. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30(22):3276–8.

56. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. 2003;31(13):3381–5.

57. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics. 2011;27(3):343–50.

58. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46(W1):W296–303.

59. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, et al. The Pfam protein families database in 2019. Nucleic Acids Res. 2019;47(D1):D427–32.

60. Krissinel E, K H: **Multiple Alignment of Protein Structures in Three Dimensions.** In: *Computational Life Sciences: Lecture Notes in Computer Science.* Edited by Berthold M. GRC, Diederichs K., Kohlbacher O., Fischer I. Berlin, Heidelberg.: Springer; 2005: 67–78.

61. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis. 1997;18(15):2714–23.
Proposed hyaluronic acid biosynthetic pathway of Pasteurella multocida consisting of nine genes. D-Glucose is first converted to D-glucose-6-phosphate by glucokinase (glk) and then processed through two distinct pathways to form UDP-D-glucuronic acid (by pgm, galU, and hyaC) and UDP-N-acetyl-D-glucosamine (by pgi, glmS, glmM, and glmU). These two precursors are combined by hyaluronan synthase (hyaD) to form hyaluronic acids. Heatmaps represent the presence of these enzymes in 11 strains of P. multocida, H. sapiens, E. coli, S. pyogenes strain M1 GAS, and S. thermophilus. Coloured boxes show the percentage of sequence identity, and N/A indicates the absence.
Figure 2

Pattern analysis of the amino acid sequences of nine hyaluronic acid biosynthetic proteins from different P. multocida strains. The patterns were derived from a single amino acid difference on each enzyme based on the multiple aligned sequences. Pattern numbers and colour shades were compared on the same row, representing variations in the same enzyme.
Figure 3

Phylogenetic trees constructed by comparing the nucleotide sequences of nine HA biosynthetic genes from 11 strains of P. multocida with those of other organisms. Coloured boxes represent different capsular serotypes. The genes from H. sapiens, E. coli, S. pyogenes, and S. thermophilus were considered as the outgroup.
Figure 4

Phylogenetic trees constructed by comparing amino acid sequences of nine HA biosynthetic enzymes from 11 strains of P. multocida with those of other organisms. Coloured boxes represent different capsular serotypes. The enzymes from H. sapiens, E. coli, S. pyogenes, and S. thermophilus were considered as the outgroup.
Figure 5

Phylogenetic trees constructed by concatenated nucleotide and amino acid sequences of genes (A) and proteins (B) in the HA biosynthetic pathway from 11 strains of P. multocida compared to those of other organisms. Coloured boxes represent different capsular serotypes, and H. sapiens, E. coli, S. pyogenes, and S. thermophilus were considered the outgroup.
Figure 6

Superimposed predicted structures of nine hyaluronic acid biosynthetic enzymes of 11 P. multocida strains. Red arrows indicate structural positions that had different conformations across the compared strains. The predicted structures are shown in different colours according to the strains.
Figure 7

Predicted structures of the binding pockets of the Glck, Pgm, GlmU, and HyaD proteins using the structure of strain ATCC4317 as a reference. The binding pockets that were different from those of other strains are highlighted and displayed in red circles (labelled P1, P2, ...).
Figure 8

Comparative domain size and architecture of the Glck (A), GlmU (B), and HyaD (C) proteins in P. multocida by using strain ATCC43137 as a reference. Red dot shapes represent the domains with different amino acid residues. Non-coloured domains indicate the absence of the domain in that strain.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS4.docx
- TableS2.docx
- TableS3.docx
- TableS1.docx