Comparison of Hyaluronic Acid Biosynthetic Genes From Different Strains of Pasteurella multocida

Pailin Pasomboon1,2, Pramote Chumnanpuen2,3 and Teerasak E-kobon1,2

1Department of Genetics, Faculty of Science, Kasetsart University, Bangkok, Thailand. 2Omics Center for Agriculture, Bioresources, Food and Health, Kasetsart University (OmiKU), Bangkok, Thailand. 3Department of Zoology, Faculty of Science, Kasetsart University, Bangkok, Thailand.

Introduction
Hyaluronic acid (HA) or hyaluronan is a linear repeat of glucuronic acid and N-acetylglucosamine found in various animal parts, ie, rooster combs, vitreous humour in eyes, umbilical cords, skin, and cartilage.1 Hyaluronic acids 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: eg, as a diagnostic marker for cancer, rheumatoid arthritis, and liver pathologies; reducing inflammation of the wound; and in drug delivery.2,3 Hyaluronic acids 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 a component of their capsule and mucus.1 P multocida is a gram-negative bacterium that causes various diseases in livestock, including fowl cholera, respiratory diseases, septicaemia, and atrophic rhinitis.7 This bacterium produces a capsule consisting of different polysaccharide compositions according to their capsular serotypes, including HA in serotype A, heparin in serotype D, and chondroitin sulphate in serotype F which shared similar enzymes in the pathways.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 is known to produce a capsule containing HA similar to Streptococcus sp., algae, viruses, and vertebrates.12,13 Hyaluronic acid biosynthesis in these organisms involves 8 genes, which encode glucose-6-phosphate isomerase (pfi), phosphoglucomutase (pgm), UTP-glucose-1-phosphate uridytransferase (galU), UDP-glucose 6-dehydrogenase (hyaC), l-glutamine:d-fructose-6-phosphate aminotransferase (glmS), phosphoglucosamine mutase (glmM), bifunctional N-acetylglucosamine-1-phosphate uridytransferase/glucosamine-1-phosphate acetyltransferase (glmU), and hyaluronan synthase (hyaD).12,14

Hyaluronic acid is usually synthesized by 2 distinct pathways that synthesize HA precursors; these pathways begin with the phosphorylation of glucose to produce glucose-6-phosphate for 2 precursors of HA biosynthesis (UDP-glucuronic acid and N-acetylglucosamine).12,15 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 uridytransferase (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 glucos-
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 amidotransferase (GlmS) to produce glucosamine-6-phosphate, which is later modified by phosphoglucosamine 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 2 precursors, generating the HA polymer.16,17 Recent reviews by Peng et al10 showed that the capsule biosynthesis genes in _cap_ loci varied among different capsular types, particularly between types A, D, and F and type B. However, sequence and structural comparison of these enzymes have not been explicitly explained. 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.18 Class II hyaluronan synthases are different from class I hyaluronan synthases of structural topology and biosynthesis mechanism.15,19 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 sulphated or chemically modified after biosynthesis.18,20,21

Previous research improved HA production by genetical improvement of the _has_ operon of _Streptococcus_ sp. which is composed of 3 genes including _hasA_, _hasB_, and _hasC_ to encode hyaluronan synthase, UDP-glucose dehydrogenase, and UTP-glucose-1-phosphate uridyltransferase, respectively. However, production has a risk of pathogenic bacterial contamination.22 HA synthesis has been studied by the genetical modification of microorganisms ( _Bacillus_ sp., _Lactococcus_ sp., and _Escherichia coli_) to harbour hyaluronan synthases from _Streptococcus_ sp. and _P. multocida_ for the production of high-quality and safety-to-use HAs.15,17 An example of industrial-scale HA production was developed with recombinant _E. coli_ and _Bacillus_ sp. strains harbouring the _hasA_ or _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 improved and 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 _ugd_ gene of _E. coli_ strain K1.5 The level of HA production in the recombinant bacterium increased (2.7–3.7 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 with the co-expression of the _ugd_ gene from _E. coli_ and _hasA_ from _Streptococcus pyogenes_.13 HA production by co-overexpression of HA synthase genes ( _hasA_ from _S. equi_ ( _sHasA_), _S. pyogenes_ ( _spHasA_), _S. uberis_ ( _suHasA_), and _P. multocida_ ( _pmHasA_)) in _Corynebacterium glutamicum_ found _sHasA_ to be better suited to _C. glutamicum_ with the high titre of HA production (74.1 g L⁻¹), and deletion of putative glycosyltransferases ( _gg0424_ and _gg0420_) could reduce cell-surface polysaccharides and enhanced HA production.25 Although differences in the capsule biosynthesis genes of _P. multocida_ have been studied, the HA biosynthetic enzymes have not been molecularly and structurally compared among _P. multocida_ strains 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 biosynthesis genes were compared across different capsular types of _P. multocida_ and to their orthologues in other bacteria and humans. We proposed that structural variations observed within these enzymes could be useful targets for the improvement of HA production.

**Materials and Methods**

**Identification of HA biosynthetic genes**

The HA biosynthetic genes from 22 strains contain 11 complete genomes and contig sequences from 11 strains of _P. multocida_ with known capsular type were obtained from the NCBI genome database including ATCC43137 (accession number CP008918), 36950 (accession number CP003022), HB01 (accession number CP006976), HB03 (accession number CP003328), FDAARGOS-218 (accession number CP020405), PMT12.1 (accession number CP007205), NCTC10323 (accession number LR134532), Razi_Pn0001 (accession number CP017961), HN06 (accession number CP003313), HN07 (accession number CP007040), PM70 (accession number AE004439), Pesh (accession number JQAC00000000), PMTB (accession number AWTD00000000), THA (accession number JQAE00000000), THD (accession number JQAF00000000), HND03 (accession number PPWE00000000), HND16 (accession number PPWR00000000), HND18 (accession number PPWT00000000), HND20 (accession number PPWV00000000), HND21 (accession number PPWY00000000), HNF01 (accession number PPWX00000000), and HNF02 (accession number PPWY00000000). Eight HA biosynthetic genes ( _pgi_ , _pgm_ , _galU_ , _hyaC_ , _glmS_ , _glmM_ , _glmU_ , and _hyaD_ ) previously described from the study of Mao et al1 were used to search for orthologous nucleotide sequences in the obtained _P. multocida_ genomes using BLASTn.26 A similarity search was performed against the genomes of _Homo sapiens_ , _E. coli_ strain K-12 MG1655, _S. pyogenes_ strain M1, and _S. thermophilus_ to identify orthologous genes according to previous reports,27,28 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 programme version 7.30 to identify conserved regions across different capsular types of *P. multocida*, and the sequences were compared with 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.30,31 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 1 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 8 digits was obtained for each sample to represent variations within these 8 HA biosynthetic enzymes. These numeric sets were then compared within different *P. multocida* strains and across different organisms.

**Phylogenetic analysis of the HA biosynthetic genes**

Phylogenetic analysis was used to analyse the evolutionary relationship of these 8 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 1000 bootstraps performed with the MEGA programme version 7.0.32 The phylogenetic relationships were compared with the amino acid sequence patterns previously described to understand the relatedness of the HA biosynthetic enzymes. The nucleotide and amino acid sequences of these 8 genes were edited and concatenated using the AliView programme version 3.0 (GPLv3)33 before the sequence alignment and phylogenetic analysis as previously described.

**Structural comparison of the HA biosynthetic enzymes**

As the protein crystal structures of the 8 HA biosynthesis enzymes of *P. multocida* are not available, the protein structures of these enzymes from different *P. multocida* strains were homology modelled using SWISS-MODEL programme and homologous protein structural templates of related bacterial species from the PDB database (Supplementary Table S1).34

The best-predicted protein structures were selected based on Qualitative Model Energy Analysis (QMEAN)35 and Global Model Quality Estimation (GMQE) values.36 Scores closer to 1 indicated a high level of reliability of the structural prediction. Protein structural domains were also determined by the Pfam programme using amino acid sequences as the input.37 The programme annotated and identified functional domains that could be used to explain the structural differences and perhaps different biochemical activities of these 8 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 programme38 and visualization with the SWISS-pdb viewer program39,40 to identify different structural positions on the aligned structures.41 Binding site pockets of the predicted protein structures were also compared by the DoGSiteScorer program,42 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 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]).43 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. The enzyme structures of the examined *P. multocida* strains that had a high level of structural changes were hypothesized to have different HA biosynthesis capabilities.

**Results**

**Identification of the HA biosynthetic genes of *P. multocida* strains**

Eight HA biosynthetic genes were identified in 22 strains of *P. multocida*, while UDP-glucose dehydrogenase (hyaC) was absent in capsular type B and hyaluronan synthase (hyaD) were absent in capsular type B, D, and F strains, the similar chondroitin synthase (fcbD) was found in type F instead. These 8 *P. multocida* biosynthetic genes shared a high percentage of identity (between 82% and 100%) and percentage of sequence coverage (between 99% and 100%) with the query sequences from Mao et al. (Figure 1). The lowest percentage of identity (82%) was observed in the hyaD similar to fcbD genes which encode chondroitin synthase in capsular type F strains (Table 1). Comparison of these genes with those of other organisms showed a lower percentage of identity and coverage: 34% to 100% and 0% to 13% in *H. sapiens, 25% to 100% and 0% to 24%
Bioinformatics and Biology Insights

Figure 1. Diagram representing steps in the comparison to 8 HA biosynthetic genes using sequence and structural bioinformatics analyses. Each step represented by the block and parameters were described inside the block. HA indicates hyaluronic acid.

Table 1. Sequence comparison of the hyaluronan synthase (hyaD) gene in 10 strains of *P. multocida* using strain ATCC43137 as a reference.

| STRAINS     | GENBANK ACCESSION NUMBERS | CAPSULAR TYPES | HOST  | SEQUENCE COMPARISON |
|-------------|----------------------------|----------------|-------|---------------------|
|             |                            |                |       | NUCLEOTIDE          | PROTEIN |
|             |                            |                |       | E-VALUE % IDENTITY  | E-VALUE % IDENTITY |
| ATCC43137   | CP008918                   | A              | Swine | 0 100               | 0 100   |
| 36950       | CP003022                   | A              | Bovine| 0 99                | 0 99    |
| HB01        | CP006976                   | A              | Bovine| 0 100               | 0 100   |
| HB03        | CP003328                   | A              | Swine | 0 100               | 0 100   |
| PMTB2.1     | CP007205                   | A              | Bovine| 0 99                | 0 99    |
| FDAARGOS_218| CP020405                   | A:3            | Turkey| 0 99                | 0 99    |
| HN07        | CP007040                   | F*             | Swine | 0 82                | 0 82    |
| PM70        | AE004439                   | F*             | Avian | 0 82                | 0 82    |
| HNF01       | PPWX000000000              | F*             | Swine | 0 82                | 0 82    |
| HNF02       | PPWY00000000               | F*             | Swine | 0 82                | 0 82    |

Abbreviation: *P. multocida*, Pasteurella multocida.
*Strains that harbour the fcbD gene, possible the hyaD ortholog.

Comparison of the nucleotide and amino acid sequences of these 8 HA biosynthetic genes from 22 strains of *P. multocida* showed a high level of within-species identity (98%-100%) when using the ATCC43137 strain of capsular type A as a reference. Only the fcbD gene of strains with capsular type F differed significantly from hyaD of the capsular type A strains (82%), as shown in Table 1.

Sequence similarities and amino acid patterns of the 8 HA biosynthetic genes of *P. multocida*

Comparison of the nucleotide and amino acid sequences of 8 HA biosynthetic genes of strains *P. multocida* considered at single amino acid differences on the sequence that showed the capsular types A and D had a close relationship. Particularly, capsular type A:3 strain FDAARGOS_218 had the most different pattern among other type A strains and had 3 proteins shared with other strains. The pattern of GalU was shared with some strains of type A and D, GlmS was shared with some strains of type A, D, and F, and GlmM of capsular type A strains PMTB2.1 and FDAARGOS_218 were closely related to capsular type F strain PM70 (Figure 3). The PMTB2.1 strain of capsular type A had the pattern of 3 enzymes that were different from the enzymes of other strains, while GalU, HyaC,
Figure 2. Proposed hyaluronic acid (HA) biosynthetic pathway of *P. multocida* consisting of 8 genes. HAs biosynthesis processed through 2 distinct pathways to form UDP-β-D-glucuronic acid (by *pgm*, *galU*, and *hyaC*) and UDP-N-acetyl-β-D-glucosamine (by *pgi*, *glmS*, *glmM*, and *glmU*). These 2 precursors are combined by hyaluronan synthase (*hyaD*) to form hyaluronic acids. Heatmaps represent the presence of these enzymes in 22 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.

and HyaD were shared in some type A strains, GlmS was shared with type B strains, and GlmM was shared with type A:3 (FDAARGOS_218) and F (PM70) strains (Figure 3). Capsular type B had the pattern of 3 proteins that were different from other strains while the pattern of GalU was closely related to type F and GlmU was shared with type D strain. Three proteins of the capsular type F strain had shared patterns with the capsular type D strain, and GalU was shared with type B. GlmM of Pm70 strain had shared pattern with some type A (PMTB2.1 and FDAARGOS_218) strains. 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 8 HA biosynthetic genes of *P. multocida***

Phylogenetic analysis of the HA biosynthetic genes in Figure 4 shows the separation between these 8 enzymes from *P. multocida* and those from other organisms. The capsular type A was grouped together with type A:3 strain in the tree of *glmS*, *glmM*, *glmU*, and *hyaD* while the *pgi* of PMTB2.1 was grouped in capsular type D strain. The *glmM*, *pgi*, *pgm*, and *hyaC* genes of capsular types A:3 and F were closely clustered, while the capsular type F and A (PMTB2.1) strain were only grouped in the trees of the *galU*, *glmS*, and also *glmM*. These genes from the remaining capsular type A, B, and D strains were well grouped together.

The patterns were changed in the protein tree of these 8 enzymes (Figure 5). The phylogenetic tree from the amino acid sequences of the capsular type A strains remained similar to the nucleotide sequence phylogenetic tree in some enzymes. The GalU proteins of the capsular type D strains were grouped with those of the capsular type A and A:3 strains. In contrast, the GlmS proteins were grouped with capsular type A and F. Three proteins (Pgi, Pgm, and GlmS) of capsular type F were clustered together with capsular type D, while GlmM of PM70 strain were grouped with capsular type A and A:3 strain FDAARGOS_218. For type B strains, Pgm was closely related to the capsular type A:3, and GalU was mixed with the proteins of capsular type F strains, GlmU was clustered to type D strains, and GlmS proteins of capsular type B were closely related to
capsular type A. Phylogenetic trees from the concatenated nucleotide and protein sequences of the HA biosynthetic genes in Figure 6 showed the same clustering pattern of the capsular type A, A:3, B, D, and F strains (Figure 6A and 6B). These phylogenetic results depicted within-gene variation from the single sequence analysis and within-serotype similarity from the analysis of concatenated sequences, and possibly different mutational history of these genes in the HA biosynthetic pathway of *P. multocida* strains.

**Structural comparison of the HA biosynthetic enzymes of *P. multocida***

The structures of 8 HA biosynthetic enzymes from different *P. multocida* strains were successfully predicted and compared (Supplementary Tables S1, S2). Predicted structures of the Pgi, GalU, HyaC, and GlmM proteins were highly similar between capsular types A, B, D, and F (Figure 7). Structural comparison of the Pgm, Glms, Glmu, and HyaD or FcbD proteins showed variations, particularly Pgm of the type A:3 strain FDAARGOS_218, capsular type A strain PMTB2.1, and capsular type B, D, and F strains; Glms of capsular type A and A:3 strains (FDAARGOS_218 and PMTB2.1); capsular type B, D, and F strains; HyaD from capsular type A and A:3 (36950 and FDAARGOS_218) strains; and FcbD from capsular type F strains (PM70, HN70, HNF02, and HNF02) (Supplementary Tables S4-S6). Certain positions of these were correlated with the active site of the enzymes determined by a sample prediction score greater than 0.50 (Supplementary Table S3). 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 Pgm (Figure 8A) was at the P3 position of all strains. The P1 position of Glmu (Figure 8C) was predicted to be the active site only in the capsular type A strain ATCC43137 while position P4 of capsular type A strain ATCC43137, capsular type B, D, and F strains was predicted to be involved with the active site. The catalytic site at the P3 position of all strains except PMTB2.1 was predicted. The catalytic pockets of HyaD (Figure 8D) at the P2 position of all strains and the P5 position were unique to the capsular type A strain 36950 and capsular type A:3 strain FDAARGOS_218 which covered 3 pocket sites (P4, P5, and P6) of type A strain ATCC43137. The P5 position was also absent in capsular type F strains. Analysis of the conserved domains of these 8 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 2 enzymes, as shown in Figure 9. The fucokinase domains of the Glmu protein (Figure 9A) of strain ATCC43137 were different in length (number of amino acid residues) when compared with those of other
Figure 4. Phylogenetic trees constructed by comparing the nucleotide sequences of 8 hyaluronic acid (HA) biosynthetic genes from 22 strains of *P. multocida* with those of other organisms. A to H were pgi, pgm, galU, hyaC, glmS, glmM, glmU, and hyaD, respectively. Coloured boxes represent different capsular serotypes. The genes from *H. sapiens*, *E. coli*, *S. pyogenes*, and *S. thermophilus* were considered as the outgroup.
Figure 5. Phylogenetic trees constructed by comparing amino acid sequences of 8 hyaluronic acid (HA) biosynthetic enzymes from 22 strains of *P. multocida* with those of other organisms. A to H were Pgi, Pgm, GaU, HyaC, GlmS, GlmM, GlmU, and HyaD, respectively. Coloured boxes represent different capsular serotypes. The enzymes from *H. sapiens*, *E. coli*, *S. pyogenes*, and *S. thermophilus* were considered as the outgroup.
Figure 6. Phylogenetic trees constructed by concatenated nucleotide and amino acid sequences of (A) genes and (B) proteins in the hyaluronic acid (HA) biosynthetic pathway from 22 strains of *P. multocida* compared with those of other organisms. Coloured boxes represent different capsular serotypes, and *H. sapiens, E. coli, S. pyogenes*, and *S. thermophilus* were considered as the outgroup.

Figure 7. Superimposed predicted structures of 8 hyaluronic acid biosynthetic enzymes of 22 *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. Structures A to H belonged to Pgi, Pgm, GalU, HyaC, GlmS, GlmM, GlmU, and HyaD, respectively.
strains (Supplementary Table S7), while the 2C-methyl-D-erythritol-4-phosphate cytidylyltransferase (IspD) domain length of strain FDAARGOS_218 differed from those of other strains (Supplementary Table S7). The domains of capsular type F strains (HNF01 and HNF02) were slightly different in length and position compared with those of the reference strain ATCC43137 and other strains (including NTP_transfase and NTP_transf_3 domains) (Figure 9A). Several changes were observed in the domains of the HyaD proteins (Figure 9B). The HyaD domains of strains 36950 were similar to those of other strains; only the domain sizes and positions 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 and domain of unknown function (DUF5828) (white box in Figure 9B) were absent in FcbD protein of the capsular type F strains. Glycosyl transferase-like family 2 (Glycos_transd_2_2) was present in capsular type F strains more than other strains and also showed differences in sizes and positions between strains. The glycosyl-hydrolase 97N-terminal (GH97_N) domain was only present in FcbD protein of capsular type F strain (Supplementary Table S8).

**Discussion**

The HA biosynthesis pathway of *P. multocida* involves 8 enzyme-encoding genes, including *pgi*, *pgm*, *galU*, *hyaC*, *glmS*, *glmM*, *glmU*, and *hyaD*, similar to the HA biosynthesis pathway in *S. pyogenes*.15,44-46 Hyaluronic acids 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, D, and F shared 7-core HA biosynthetic enzymes that produce UDP-d-glucuronic acid and UDP-N-acetyl-d-glucosamine, which are modified to HA, chondroitin sulphate, and heparin, and disaccharide repeats of other components that vary across different strains.47 Capsular type B strains do not have *hyaC* and *hyaD* genes to produce glucuronic acid and HA as they have

---

**Figure 8.** Predicted structures of the binding pockets of the (A) Pgm, (B) GlmS, (C) GlmU, and (D) HyaD or FcbD 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, . . .).
$N$-acetyl-mannosaminuronic acid as a major component of capsular type B.45 The hyaluronan synthase ($hyaD$) of capsular type A strain is essential in the joining and polymerization of the 2 precursors for HA biosynthesis. In the capsular type F strain, $N$-acetyl-$\alpha$-galactosamine was used instead of $N$-acetyl-$\alpha$-glucosamine to generate chondroitin sulphate (CS).46,49
Chondroitin used d-glucuronic acid with N-acetylgalactosamine which generated from N-acetylgalactosamine by acetylgalactosamine-4-epimerase (GalE) and polymeric chondroitin chain growth by chondroitin synthase (FcbD) before the modification to chondroitin sulphate A, C, D, and E units by sulfotransferase. Heparin (HP) consists of repeating disaccharide units of uronic acid residues (d-glucuronic acid [GlcA] or l-iduronic [IdoA]) and N-acetyl-d-glucosamine produced by heparosan synthase (pmHS) and modification through N-deacetylation and N-sulfation by C5-epimerase and O-sulfotransferase. Members of the HA biosynthetic pathway synthesize 2 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. Seven HA biosynthesis genes were found in capsule type D and F strains except hyaD while type B strains did not have hyaC and hyaD because the type B strains have arabinose, galactose, mannose, and especially N-acetylmannosaminuronic acid as a major capsular component which might require different enzymatic pathways. The production of each glycosaminoglycan of P. multocida was different, especially the genes (hyaD, pmHS, and fcbD) used to catalyse the formation of the glycosidic linkage to produce HA, heparin, and chondroitin. This result suggested that the first 7 enzymes of the HA biosynthesis pathway could be the part of carbohydrate metabolism for synthesizing the core 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 carbohydrate metabolism including nucleotide sugar biosynthesis and UDP-N-acetyl-d-glucosamine biosynthesis. The absence of the hyaD gene would not affect heparin and chondroitin sulphate biosynthesis because the enzyme is not required. Only the hyaD gene of the capsule type A (36950 strain), A:3 and fcbD of capsule type 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. The protein pattern and phylogenetic analysis supported the sequence alignment of these genes. These results indicated that the HA biosynthetic genes of types A, D, and F had a closer relationship compared with those of types B, except hyaD which was absent in capsule type B, D, and F strains and was similar to fcbD for chondroitin biosynthesis. Although strains with the same capsule type evidently have highly similar genes in the pathway, this study found that the type A and A:3 strains had diverse patterns of HA biosynthetic genes (previously used in genetically engineered E. coli). These variations could diversify catalytic activities of the HA biosynthetic enzymes among the capsule 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.

Structures of the Pgm, GlmS, GlmU, and HyaD proteins showed individual variations across different strains of P. multocida. Analysis of the binding pockets by comparison with the structures of these enzymes in other organisms found that the sites were relevant to the catalytic pockets. The Pgm active site was studied in Salmonella enterica serovar Typhimurium LT2 and Pseudomonas aeruginosa and was consistent with the predicted binding site pockets in this study (4 domains consisted of residues 1-153, residues 154-256, residues 257-368, and residues 369-463). The highly conserved catalytic residues in these active sites involved in binding of phosphosugar substrate for the catalysis of glucose-6-P to glucose-1-P. GlmS contains 2 active sites including glutamine amidotransferases (GATase) domain to catalyse the removal of the ammonia group from a glutamine molecule and its subsequent transfer to a specific substrate and fructose-6-P isomerization activity that were also revealed in the predicted binding site pockets of GlmS in this study (residues 71-123). The HyaD also contains 2 active sites: 1 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. 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 (residues 427-688) and glucosamine-1-P-transferase domain (residues 135-426) of HyaD. The FcbD protein for chondroitin biosynthesis of capsular type F was similar to HyaD in type A and also contains 2 active sites including glucuronic acid-transferase has a function to elongate UDP-glucuronic acid to oligosaccharides like HyaD and UDP-N-acetylgalactosamine transferase domain to extent UDP-N-acetyl-d-galactosamine chains. The predicted binding site pocket in this study was correlated to the active site published in the previous study from E. coli strain K4. Our study found that the arrangement of Glycos_tranf_2, Glycos_tranf_2_2, Glycos_tranf_2_3, Glycos_tranf_2_4, and DUF2536 domains in the HyaD protein differed in these P. multocida strains related to glycosyltransferase activity to catalyse the formation of the glycosidic linkage, suggesting the reason for the structural variation of the binding site and perhaps different molecular activities. FcbD is different from HyaD and only harbours the GH97_N domain in the capsular type F strain that involves hydrolyzing the glycosidic bond between carbohydrates or carbohydrate and a non-carbohydrate moiety. 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-p-glucosamine. 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 were correlated with the C-terminal acetyltransferase catalytic sites relating to the nucleotidyl transferase (NTP_transferase) domain. GlmU has 2 active sites: (1) 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 (residues 260-437) and (2) pyrophosphorylase, which catalyses the transfer of uridyl from UTP to d-glucosamine-1-P, forming UDP-N-acetyl-d-glucosamine and pyrophosphate at the N-terminal domain (residues 3-229). Consequently, structural variations of these 4 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 and A:3 strains that had HA-containing capsules, especially HyaD, the hyaluronan synthase type II, that found only in P. multocida and is a unique cytosolic glycosyltransferase compared with Type I transmembrane enzyme. Because HyaD is the key enzyme in the HA biosynthesis, identification of the HyaD variants in these type A strains could presumably result in different production of the HA-containing capsular materials. Certain HyaD variants might be responsible for the watery mucoid colony phenotype which had more capsular materials than the typical mucoid colony. As only one HyaD variant of P. multocida has been used in the HA synthesis commercially, the authors could obtain new additional variants which might outperform the currently used HyaD in terms of catalytic velocity, enzymatic stability, and substrate specificity. The finding could be supported by the study of Mandawe et al which randomly mutated P. multocida hyaD gene by the error-prone PCR and found 3 mutant positions distant from the catalytic site could help increasing the enzyme flexibility to extend the HA chain-length to approximately 4.7 MDA. Experimental investigation of the HyaD variants in this study could improve industrial HA production by expression of these new HyaD variants in recombinant bacteria, investigating regulation of this key enzyme for metabolic engineering as well as the production of other glycosaminoglycans including heparin and chondroitin in capsular types D and F.

Conclusions
This study identified 8 genes involved in the HA biosynthetic pathway of P. multocida strains. Seven genes had a high level of within-species similarity (98% to 100%), while hyaD of strains with capsular type A (36950), A:3, and fcbD of capsular type F differed significantly to those of other strains. Sequence, phylogenetic, and pattern analysis of these genes showed that capsular types A, D, and F had a close relationship compared with capsular type B. Structural comparison revealed variations in the Pgm, GlmS, GlmU, and HyaD or FcbD proteins. The variations within the active sites and domains of Pgm, GlmS, GlmU, and HyaD could influence enzymatic function and substrate specificity. Thus, these enzymes, particularly new HyaD variants, will be potential targets for improving HA production.

Acknowledgements
We greatly thank the Department of Genetics and Department of Zoology, Faculty of Science, Kasetsart University, Bangkok, Thailand, and the Omics Center for Agriculture, Bioresources, Food and Health, Kasetsart University (OmiKU), Kasetsart University, Thailand.

Author Contributions
PP planned the project, carried out the analyses, and wrote the manuscript. PC supervised the project, discussed the results, and revised the manuscript. TE planned the project, verified the analytical methods, discussed the results, and edited the manuscript.

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

ORCID iD
Teerasak E-kobon https://orcid.org/0000-0002-3919-9841

Supplemental material
Supplemental material for this article is available online.

REFERENCES
1. Mao Z, Shin HD, Chen R. A recombinant E. coli bioprocess for hyaluronan synthesis. Appl Microbiol Biotechnol. 2009;84:63-69.
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;9:7081-7092.
3. Tseplikov RN, Beloded AV. Hyaluronic acid – an ‘old’ molecule with ‘new’ functions: biosynthesis and depolymerization of hyaluronic acid in bacteria and vertebrates including during carcinogenesis. Biochemistry (Mosc). 2015;80:1093-1108.
4. Kogan G, Sohtés L, Stern R, Gemeineier P. Hyaluronic acid: a natural biopolymer with a broad range of biomedical and industrial applications. Biotechnol Lett. 2007;29:17-25.
5. Mahajan AS, Pilling D, Gomer RH. High and low molecular weight hyaluronic acid differentially regulate human fibrocyte differentiation. PLoS ONE. 2011;6:e26078.
6. Rayahin JE, Buhrman JS, Zhang Y, Koh TJ, Gemeinhardt RA. High and low molecular weight hyaluronic acid differentially influence macrophage activation. ACS Biomater Sci Eng. 2015;1:481-493.
7. Wilson BA, Ho MA. Pasturella multocida from zoonosis to cellular microbiology. Clin Microbiol Rev. 2013;26:631-655.
8. Wilkie IW, Harper M, Boyce JD, Adler B. Pasturella multocida diseases and pathogenesis. Curr Top Microbiol Immunol. 2012;361:1-22.
9. Okay S, Kurt Kuzildogan A. Comparative genome analysis of five Pasturella multocida strains to decipher the diversification in pathogenicity and host specialization. Gene. 2015;567:58-72.
10. Peng Z, Wang X, Zhou R, Chen H, Wilson BA, Wu B. Pasturella multocida: genotypes and genomics. Microbial Mol Biol Rev. 2019;83:e00014-19.
11. Harper M, Boyce JD, Adler B. The key surface components of Pasturella multocida: capsule and lipopolysaccharide. Curr Top Microbiol Immunol. 2012;361:39-51.
12. Chong BF, Blank LM, McLauglin R, Nielsen LK. Microbial hyaluronic acid production. Appl Microbiol Biotechnol. 2005;66:341-351.
