Heliyon 9 (2023) e16141

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

# Differential domains and endoproteolytic processing in dominant surface proteins of unknown function from *Mycoplasma hyopneumoniae* and *Mycoplasma flocculare*

Priscila Souza dos Santos [a,b], Jéssica Andrade Paes [a,b], Lais Del Prá Netto Machado [a,b], Gabriela Prado Paludo [a,b], Arnaldo Zaha [a,b,c,d], Henrique Bunselmeyer Ferreira [a,b,c,d,*]

[a] *Laboratório de Genômica Estrutural e Funcional, Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil*
[b] *Programa de Pós-Graduação em Biologia Celular e Molecular, Centro de Biotecnologia, UFRGS, Porto Alegre, Brazil*
[c] *Laboratório de Biologia Molecular de Cestódeos, Centro de Biotecnologia, UFRGS, Porto Alegre, Brazil*
[d] *Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, UFRGS, Porto Alegre, Brazil*

ARTICLE INFO

*Keywords:*
Porcine enzootic pneumonia
Proteins of unknown function
Endoproteolytic processing
Differential proteoforms
Pathogenicity

ABSTRACT

*Mycoplasma hyopneumoniae* causes porcine enzootic pneumonia (PEP), a chronic respiratory disease that leads to severe economic losses in the pig industry. Swine infection and PEP development depend on the adhesion of the pathogen to the swine respiratory tract and the host immune response, but these and other disease determinants are not fully understood. For instance, *M. hyopneumoniae* has a large repertoire of proteins of unknown function (PUFs) and some of them are abundant in the cell surface, where they likely mediate so far unknown pathogen-host interactions. Moreover, these surface PUFs may undergo endoproteolytic processing to generate larger repertoires of proteoforms to further complicate this scenario. Here, we investigated the five PUFs more represented on the surface of *M. hyopneumoniae* pathogenic strain 7448 in comparison with their orthologs from the nonpathogenic *M. hyopneumoniae* J strain and the closely related commensal species *Mycoplasma flocculare*. Comparative *in silico* analyses of deduced amino acid sequences and proteomic data identified differential domains, disordered regions and repeated motifs. We also provide evidence of differential endoproteolytic processing and antigenicity. Phylogenetic analyses were also performed with ortholog sequences, showing higher conservation of three of the assessed PUFs among *Mycoplasma* species related to respiratory diseases. Overall, our data point out to *M. hyopneumoniae* surface-dominant PUFs likely associated with pathogenicity.

## 1. Introduction

*Mycoplasma hyopneumoniae* causes porcine enzootic pneumonia (PEP), a chronic respiratory disease that leads to major economic

*Abbreviations:* CDS: Coding DNA sequence, EC: epitope coverage; PEP: porcine enzootic pneumonia, PUF: protein of unknown function.

* Corresponding author. Laboratório de Genômica Estrutural e Funcional, Centro de Biotecnologia, UFRGS, Caixa Postal 15005, 91501–970, Porto Alegre, Brazil.
*E-mail address:* henrique.bunselmeyer@ufrgs.br (H.B. Ferreira).

https://doi.org/10.1016/j.heliyon.2023.e16141
Received 19 September 2022; Received in revised form 28 April 2023; Accepted 6 May 2023
Available online 18 May 2023

losses in the pig industry [1]. This bacterium adheres to the epithelium of the porcine respiratory tract inducing cilia loss and epithelial cell death [2]. Infected pigs present clinical signs such as cough, fever and respiratory distress, and their average daily gain and feed conversion ratios are severely affected [3]. Moreover, *M. hyopneumoniae*-infected swine are also more susceptible to coinfections with viruses and other bacteria, which may lead to the development of porcine respiratory disease complex (PRDC) [4].

*M. hyopneumoniae* pathogenicity mechanisms form a complex scenario, in which bacterial and host molecules are involved in PEP establishment [5]. The virulence degree is variable among *M. hyopneumoniae* strains, ranging from the virtually avirulent strain J to high virulent strains, such as 232, 7448, 7422 and ES-2 [6–9]. Several *M. hyopneumoniae* surface proteins have been associated with relevant features for host-pathogen interactions, such as mycoplasma adhesion to the porcine respiratory tract, cytotoxicity and host immunomodulation [10]. Moreover, the occurrence of post-translational endoproteolytic processing among surface proteins from different *M. hyopneumoniae* strains have been addressed by some previous proteomic studies [11–18]. Differential endoproteolytic processing was also demonstrated for adhesion-related proteins between *M. hyopneumoniae* and *M. flocculare*, which is a closely related *Mycoplasma* species that coinhabits the porcine respiratory tract but is nonpathogenic [19]. The endoproteolytic processing is assumed to generate a combinatorial library of proteoforms, diversifying the bacterial surface [2,19]. Such diversification would have many implications for pathogen-host interactions.

The *M. hyopneumoniae* cell surface contains many proteins with assigned functions, such as adhesins, membrane transporters, enzymes, and immunomodulatory antigens [10]. However, about 35% of the genes coding for surface proteins in the genome of *M. hyopneumoniae* have been originally annotated as hypothetical proteins, with no assigned functions [7–9]. More recently, efforts have been carried out to improve coding DNA sequence (CDS) validation and genome annotation in *M. hyopneumoniae* 7448 [20,21], but many CDSs still code for proteins of unknown function (PUFs) and require further characterization, especially those enriched on the cell surface.

Interestingly, the comparative proteomics analyses demonstrated quali-quantitative differences among the repertoires of detected PUFs from *M. hyopneumoniae* pathogenic and nonpathogenic strains and from the commensal *M. flocculare* [21]. These strains and species have orthologs for more than 90% of the repertoire of predicted surface proteins [8,9]. Thus, we hypothesized that *M. hyopneumoniae* 7448 dominant surface PUFs differ in at least some aspects from their orthologs in nonpathogenic counterparts, such as *M. hyopneumoniae* J and *M. flocculare*, and that these differences are relevant for *M. hyopneumoniae* pathogenicity.

Here, we investigated the five PUFs more represented on the surface of *M. hyopneumoniae* 7448 in comparison with their orthologs from *M. hyopneumoniae* J and *M. flocculare*. The aim of the study was to identify differential features among these orthologs, including distinct structural and functional domains, differential endoproteolytic processing, and epitope repertoires. Overall, the performed *in silico* analyses, provided evidence of structural and immunological differences, whose implications for *M. hyopneumoniae* virulence and pathogenicity are discussed. Moreover, phylogenetic analyses were performed to provide information on the evolutive history of these proteins among mycoplasmas and other bacteria.

## 2. Materials and methods

### 2.1. *M. hyopneumoniae and M. flocculare PUFs and associated LC-MS/MS data*

LC-MS/MS data from proteins detected in *M. hyopneumoniae* and *M. flocculare* surface or cytoplasm-enriched protein extracts (from now on referred just as 'surface samples' or 'cytoplasm samples', respectively) were previously generated [21]. A total of 59 *M. hyopneumoniae* 7448 PUFs were identified in the LC-MS/MS data from surface samples and had the corresponding LC-MS/MS data retrieved, along with that of their *M. hyopneumoniae* J and *M. flocculare* orthologs. Surface ortholog PUF sets of *M. hyopneumoniae* (7448 and J) and *M. flocculare* were defined as orthogroups (OGs) for the subsequent analyses. From the 59 OGs, 54 were assessed in the performed studies (Table S1); the other 5 OGs, have been previously related to adhesion and were addressed in a previous study [19], and were not included in this study. The available normalized spectral abundance factor (NSAF) values [21] were then used for quantitative comparisons and to rank PUFs according to their abundances in surface samples. The top five most abundant *M. hyopneumoniae* 7448 surface PUFs and their orthologs of *M. hyopneumoniae* J and *M. flocculare*, from now on referred as MHPn, MHJn and MFn, respectively, being 'n' the OG number, were then selected for further analyses. Information regarding detected peptide sequences in surface and cytoplasmic samples, their corresponding spectral counts and sequence coverages for these 5 OGs were also retrieved from our previously published data [21]. The full-length deduced amino acid sequences of each assessed PUF and associated information were recovered from UniProt (https://www.uniprot.org/).

### 2.2. *Identification of differential domains between ortholog PUFs from M. hyopneumoniae strains and M. flocculare*

Pairwise Sequence Alignment using Needleman-Wunsch algorithm (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) was utilized to compare ortholog PUFs within each assessed OG. GeneDoc version 2.7 (https://genedoc.software.informer.com/2.7/) was used to customize the alignments. Differential domains were defined based on previously established criteria [22], being considered as such insertions/deletions of at least 5 aa in one sequence and absent in their orthologs, and/or segments with at least 10 aa with $\leq$30% of identity between orthologs with overall identities $\geq$50%.

**Table 1**
Orthogroups corresponding to the top 5 most represented surface PUFs in *M. hyopneumoniae* 7448 and their orthologs from *M. hyopneumoniae* J and *M. flocculare*.

| Orthogroup (OG)[a] | *M. hyopneumoniae* 7448 | | *M. hyopneumoniae* J | | *M. flocculare* | | PUFs rank[d] | | | Overall rank[e] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ID[b] | NCBI accession number[c] | ID[b] | NCBI accession number[c] | ID[b] | NCBI accession number[c] | *M. hyopneumoniae* 7448 | *M. hyopneumoniae* J | *M. flocculare* | *M. hyopneumoniae* 7448 | *M. hyopneumoniae* J | *M. flocculare* |
| OG1 | MHP1 | MHP7448_0489 | MHJ1 | MHJ_0486 | MF1 | MFC_00856 | 1 | 8 | 6 | 34 | 103 | 24 |
| OG2 | MHP2 | MHP7448_0352 | MHJ2 | MHJ_0347 | MF2 | MFC_01239 | 2 | 1 | 5 | 39 | 28 | 22 |
| OG3 | MHP3 | MHP7448_0538 | MHJ3 | MHJ_0540 | MF3 | MFC_00551 | 3 | 3 | 23 | 47 | 42 | 77 |
| OG4 | MHP4 | MHP7448_0138 | MHJ4 | MHJ_0134 | MF4 | WP_231237874.1 | 4 | 7 | ND | 49 | 68 | ND |
| OG5 | MHP5 | MHP7448_0483 | MHJ5 | MHJ_0480 | MF5 | MFC_01450 | 5 | 11 | ND | 50 | 109 | ND |

[a] Groups of ortholog PUFs from *M. hyopneumoniae* 7448, *M. hyopneumoniae* J and *M. flocculare*.
[b] Identification of the PUFs from *M. hyopneumoniae* (7448 and J) and *M. flocculare*, according to the OG number.
[c] Accession numbers in the NCBI database (https://www.ncbi.nlm.nih.gov/protein).
[d] According to the abundance (NSAF values) of the selected PUFs in the corresponding surface samples [21].
[e] According to the abundance (NSAF values), considering all proteins detected in surface samples [21].

### 2.3. Topology predictions, and identification and comparison of disordered regions and repeated motifs in ortholog PUFs from *M. hyopneumoniae* strains and *M. flocculare*

The full-length deduced amino acid sequences of the assessed ortholog PUFs from *M. hyopneumoniae* 7448 and J, and *M. flocculare* were used for topology predictions by the Protter software (https://wlab.ethz.ch/protter/start/) and for predictions of disordered regions by the Predictor of Natural Disordered Regions (PONDR®) algorithm VSL2 (http://www.pondr.com/) using a PONDR Score $\geq$0.5.

The search for fixed-length motifs was initially performed with the full-length deduced amino acid sequences of the *M. hyopneumoniae* 7448 assessed PUFs using Multiple Em for Motif Elicitation (MEME) with default parameters (https://meme-suite.org/meme/tools/meme). The MEME-defined consensus of the repeated motifs identified in *M. hyopneumoniae* 7448 PUFs were then used to identify matches in ortholog PUFs from *M. hyopneumoniae* 7448 and J strains and *M. flocculare* with the Motif Alignment & Search Tool (MAST) with default parameters (https://meme-suite.org/meme/tools/mast).

### 2.4. Validation and mapping of peptides detected by LC-MS/MS, and identification of endoproteolytic processing events in ortholog PUFs from *M. hyopneumoniae* and *M. flocculare*

Peptides detected by LC-MS/MS in surface and cytoplasmic samples corresponding to the PUFs within each assessed OG were recovered from our previously published data [21]. They were then validated based on the detection in at least two out of three biological replicates, as previously described [19]. The relative abundance of each validated peptide for a given PUF was inferred from the correspondent average of the peptide's exclusive unique spectral counts in the biological replicates. Validated peptides were physically mapped to the cognate PUF deduced amino acid sequences.

Endoproteolytic processing events were identified based on the detection of non-tryptic cleavage sites in mapped peptides [19]. Proteoforms were defined here as being any protein stretch covered by two or more overlapping or juxtaposed mapped peptides, and with its amino and/or carboxy-terminal ends generated by a non-tryptic cleavage event.

### 2.5. Differential endoproteolytic processing in *M. hyopneumoniae* and *M. flocculare* ortholog PUFs and between subcellular compartments

Differential endoproteolytic processing was identified through alignments of the ortholog PUFs of *M. hyopneumoniae* (7448 and J) and *M. flocculare* in surface and cytoplasmic protein samples. For this, stretches with 10 aa that comprised the non-tryptic cleavage site in a strain/species were aligned with corresponding stretches of the ortholog PUFs using the ClustalW algorithm from the Molecular Evolutionary Genetics Analysis (MEGA) software (https://www.megasoftware.net/), with default parameters.

Differential endoproteolytic processing between *M. hyopneumoniae* (7448 and J) and *M. flocculare* was assumed when a stretch covered by semi-tryptic peptides in a strain/species was absent in the corresponding mapped stretch of the ortholog(s). Likewise, differential endoproteolytic processing between subcellular fractions was assumed when a stretch covered by semi-tryptic peptides in the surface fraction was absent in the corresponding mapped stretch of the same protein in the cytoplasmic fraction, or vice versa.

### 2.6. Epitope predictions

Epitope predictions for the PUFs in each assessed OG were carried out using the Predicting Antigenic Peptides tool (http://imed.med.ucm.es/Tools/antigenic.pl). For these predictions, the full-length deduced amino acid sequences of each PUF were used. Epitope coverage (EC) was defined as the percentage of the whole protein or derived proteoform covered by identified antigenic peptides.

### 2.7. Phylogenetic analyses

Phylogenetic analyses were performed in order to investigate the evolutionary history among *M. hyopneumoniae* 7448 PUFs and orthologs PUFs. The full-length deduced amino acid sequences of *M. hyopneumoniae* 7448 were used to retrieve ortholog sequences from the NCBI protein database (March 2022) using the BLASTp tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The initial parameters used to search for orthologs were $\geq$30% identity with and $\geq$60% coverage of the query sequence. When necessary, an additional criterion of protein size variation $\leq$30% in relation to the query was used. The ClustalW algorithm from the MEGA software was used to perform multiple sequence alignments, with default parameters. MEGA was also used to estimate the best substitution model based on Bayesian Information Criterion (BIC), and to build the phylogenetic trees. The trees were built by the Maximum Likelihood method and with 2000 bootstrap replicates. FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) was used for the customization of the built phylogenetic trees.

## 3. Results

### 3.1. Selection of the top five most represented PUFs on the *M. hyopneumoniae* 7448 cell surface

Previous LC-MS/MS data were utilized in order to rank and select the five PUFs most represented on *M. hyopneumoniae* 7448 surface, and their orthologs from *M. hyopneumoniae* J and *M. flocculare*. OGs were ordered from 1 to 5 according to the PUF abundance on the *M. hyopneumoniae* 7448 surface, as shown in Table 1. The five most represented *M. hyopneumoniae* 7448 surface PUFs were

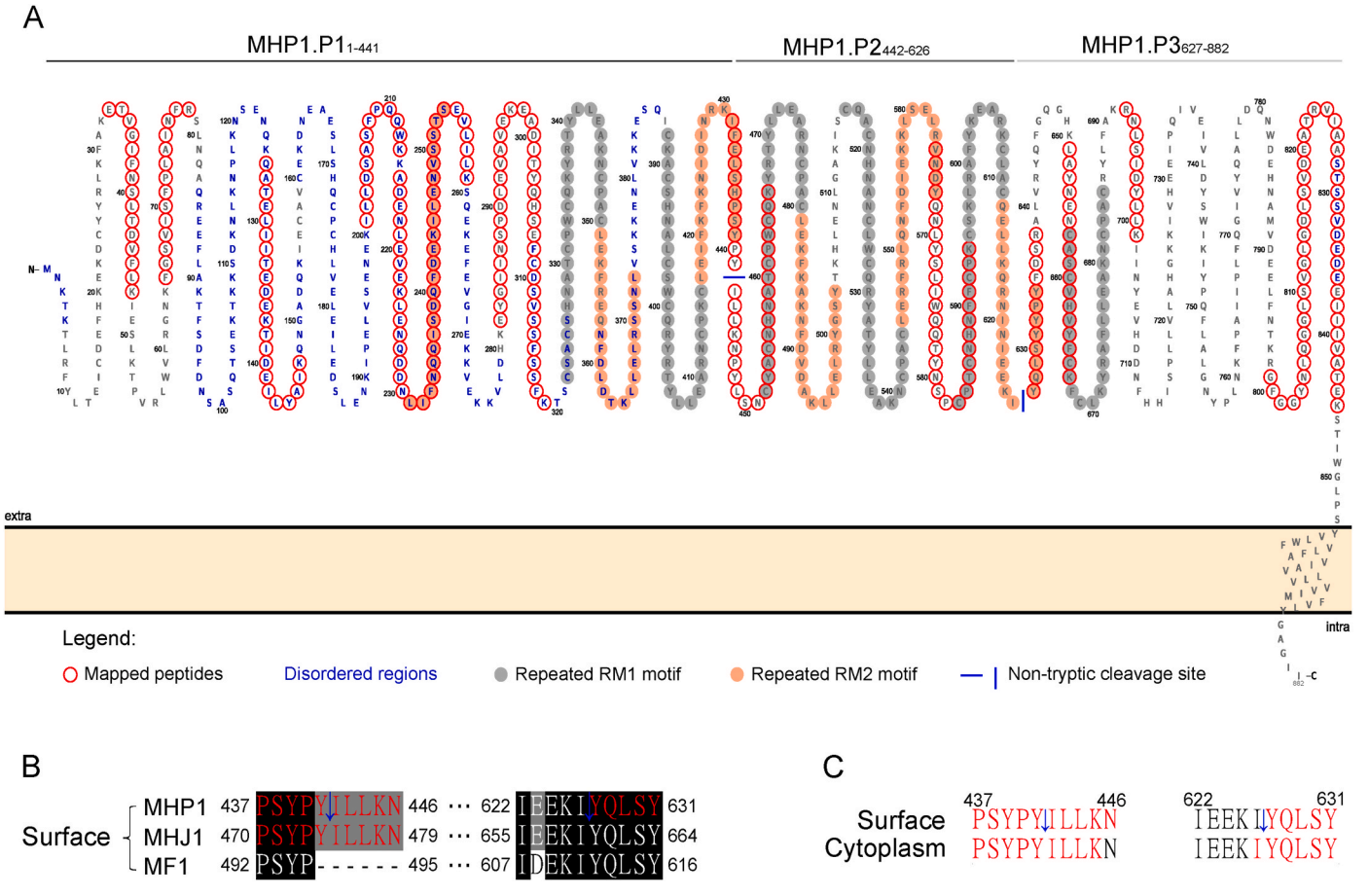


**Fig. 1.** MHP1 structural features, endoproteolytic processing, derived proteoforms and differential processing between strains/species and subcellular fractions. (A) MHP1 amino acid sequence and corresponding structural features; N- and C-terminal ends are indicated, and amino acid residue positions are numbered. MHP1 was endoproteolytically processed in two cleavage sites (blue bars) on the surface to generate three proteoforms. **(B)** Amino acid sequence alignments of the MHP1 endoproteolytic processing sites with the corresponding sites from the MHJ1 and MF1 orthologs in surface. **(C)** MHP1 differential endoproteolytic processing between surface and cytoplasm fractions. For **(B)** and **(C)**: amino acids in red are from peptides detected by LC-MS/MS; amino acids in black or white are from peptides not detected by LC-MS/MS; non-tryptic cleavage sites were indicated by blue arrows, and amino acid residue positions are numbered. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

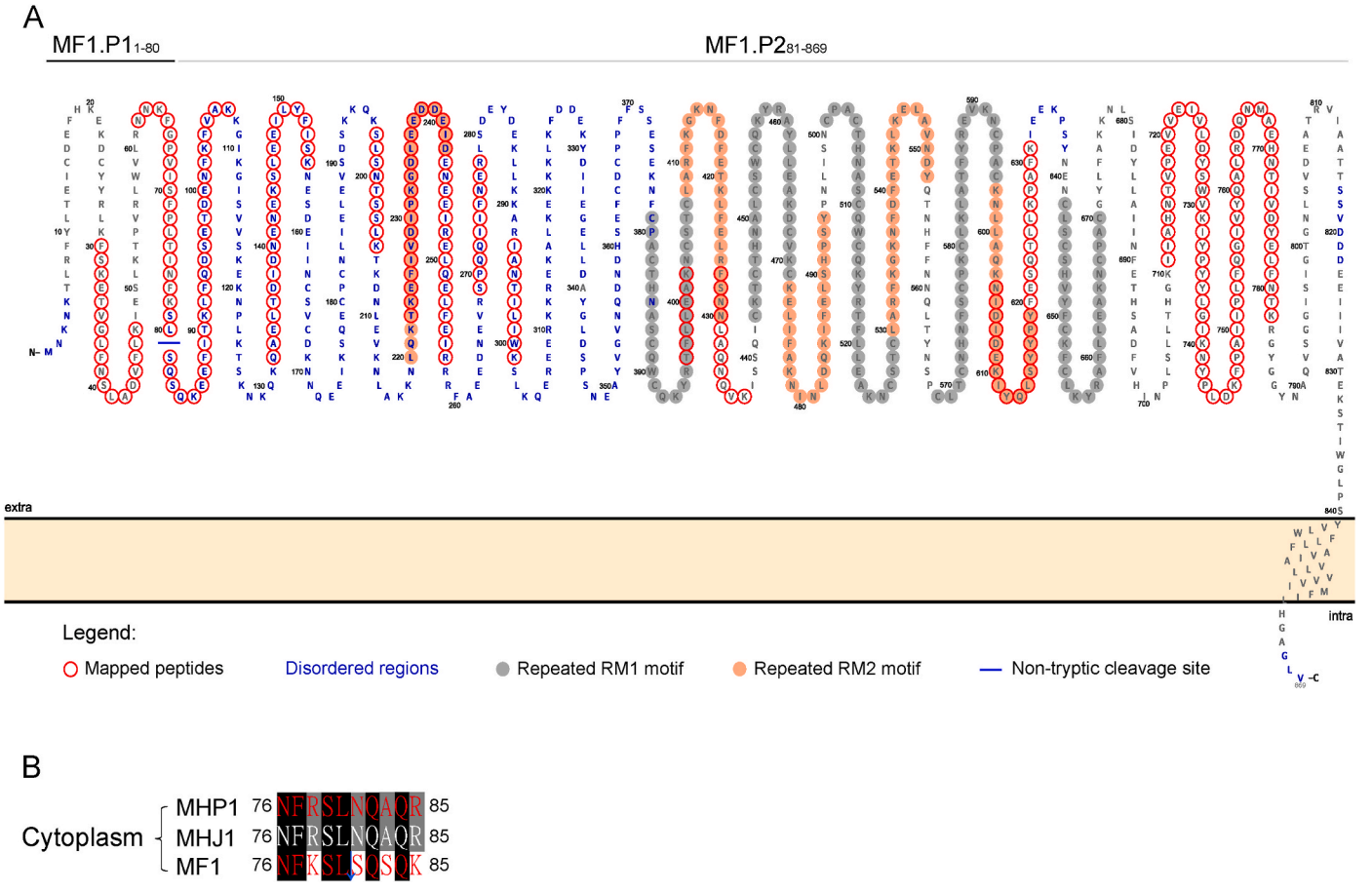


**Fig. 2.** MF1 structural features, endoproteolytic processing, derived proteoforms and differential processing between strains/species. (A) MF1 amino acid sequence and corresponding structural features; N- and C-terminal ends are indicated, and amino acid residue positions are numbered. MF1 was endoproteolytically processed in a single cleavage site (blue bar) on the cytoplasm to generate two proteoforms. **(B)** Amino acid sequence alignments of the MF1 endoproteolytic processing site with the corresponding site from the MHP1 and MHJ1 orthologs in cytoplasm. Amino acids in red are from peptides detected by LC-MS/MS; amino acids in white are from peptides not detected by LC-MS/MS; the non-tryptic cleavage site was indicated by a blue arrow, and amino acid residue positions are numbered. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

ranked from 34th to 50th in the overall rank of the total of 267 proteins detected by LC-MS/MS in surface-enriched protein samples. Most of the rank positions of the top five *M. hyopneumoniae* 7448 surface PUFs were not the same for their *M. hyopneumoniae* J and *M. flocculare* orthologs, which ranked among the 109 and 77 most abundant surface proteins, respectively.

### 3.2. Identification of differential domains among ortholog surface PUFs from M. hyopneumoniae and M. flocculare

The *M. hyopneumoniae* (7448 and J) and *M. flocculare* ortholog PUFs were assessed to search for possible differential domains (Table S2). Between *M. hyopneumoniae* 7448 and *M. hyopneumoniae* J orthologs, only orthologs from OG1 and OG4 presented differential domains (one and three, respectively) (Table S2A). More differential domains were found in the comparisons between *M. hyopneumoniae* 7448 and *M. flocculare* orthologs, with OGs1-5 exhibiting from one (in OG5) to 19 (in OG1) (Table S2B), and between *M. hyopneumoniae* J and *M. flocculare*, with OGs1-5 exhibiting from one (in OG5) to 23 (in OG1) (Table S2C).

### 3.3. Identification and comparison of disordered regions and repeated motifs among ortholog surface PUFs from M. hyopneumoniae and M. flocculare

The top five *M. hyopneumoniae* 7448 surface PUFs were assessed in the search for disordered regions and repeated motifs. All ortholog sequences from OGs1-5 exhibited numbers of disordered regions ranging from 5 (in OG5 for MHP5 and MHJ5) to 25 (in OG4 for MF4), with different numbers of amino acids in length, ranging from 1 to 261 aa (Table S3).

Repeated motifs, in turn, were found only in MHP1. In this PUF, the MEME search identified two repeated motifs, with lengths of 29 aa (RM1) and 23 aa (RM2) (Fig. S1). The search for these repeated motifs in the members of OG1 revealed that RM1 and RM2 appear in juxtaposed blocks (RM1-RM2), repeated five times in MHP1 and MHJ1, and four times in MF1 (Fig. S2; Table S3). These juxtaposed repeated blocks are followed by an additional individual copy of RM1 towards the C-terminal end in the three OG1 orthologs, 22 aa away from the block formed by the remaining repeated motifs. Moreover, an additional individual copy of RM2 was found towards the N-terminal end of these three orthologs, 68 aa away from the remaining repeated motifs for MHP1 and MJ1, and 136 aa away from the remaining repeated motifs for MF1. The repeated RM1-RM2 sets are regularly interspaced in the three orthologs, being 12–15 aa apart from each other in MHP1 and MHJ1, and 6–15 in MF1 (Fig. S2; Table S3).

### 3.4. Endoproteolytic processing and identification of differential proteoforms derived from the ortholog M. hyopneumoniae and M. flocculare PUFs

Validated peptides and associated quali-quantitative proteomic data corresponding to the *M. hyopneumoniae* (7448 and J) and *M. flocculare* PUFs in surface protein samples and cytoplasmic protein samples were recovered from previously performed LC-MS/MS studies (Tables S4A–C and S5A–C, respectively). Endoproteolytic processing sites were then inferred from the identification of semi-tryptic peptides.

For OG1 in surface samples, only MHP1 was processed, as two non-tryptic cleavage sites (PSYPY↓ILLKN and IEEKI↓YQLSY) were identified in this PUF (Fig. 1A; Table S4A). These cleavage sites do not coincide with any of the putative disordered regions predicted for MHP1. The processing potentially generates N-terminal, central, and C-terminal proteoforms, respectively called MHP1.P1$_{1\text{-}441}$, MHP1.P2$_{442\text{-}626}$, and MHP1.P3$_{627\text{-}882}$, with the last one comprehending a transmembrane domain (Fig. 1A). Each of these predicted proteoforms includes different numbers of the repeated RM1 and RM2 motifs found for MHP1. The two cleavage sites found in MHP1 (PSYPYILLKN and IEEKIYQLSY) presented 100% of conservation in the considered 10 aa-stretches with MHJ1 (Fig. 1B). However, no evidence of cleavage was found for MHJ1, as the PSYPYILLKN site was covered only by a tryptic peptide, and no peptides covering the IEEKIYQLSY were mapped. In the MF1 ortholog, the C-terminal part of the PSYPYILLKN is absent and no peptides covering the IDEKIYQLSY were mapped (Fig. 1B). Overall, these findings point out to differential endoproteolytic processing of the MHP1, MHJ1 and MF1 orthologs. Moreover, the absence of any semi-tryptic peptide mapped to the PSYPYILLKN or IEEKIYQLSY sites in cytoplasmic samples suggests differential endoproteolytic processing for this *M. hyopneumoniae* 7448 PUF between subcellular fractions (Fig. 1C).

For OG1 in cytoplasmic samples, endoproteolytic processing was found only in MF1 (Fig. 2A; Table S5C). A single non-tryptic cleavage site was identified for MF1 (NFKSL↓SQSQK), potentially generating a N-terminal MF1.P1$_{1\text{-}80}$ proteoform and a C-terminal MF1.P2$_{81\text{-}869}$ proteoform comprehending a transmembrane domain (Fig. 2A). This non-tryptic cleavage site coincides with a disordered region (79–339) predicted for MF1. MF1.P2$_{81\text{-}869}$ proteoform included all repeated RM1 and RM2 motifs found for MF1. The NFKSLSQSQK site in MF1 is only partially conserved (60% identity) in MHP1 and MHJ1 orthologs, and peptides covering this site were mapped only for MHP1 (Fig. 2B). This suggests differential endoproteolytic processing between MF1 and MHP1.

For OG2 in surface samples, endoproteolytic processing was found only in MF2 (Fig. S3A; Table S4C). A single non-tryptic cleavage site was identified for MF2 (IIKKQ↓LQLSA), resulting potentially in a N-terminal MF2.P1$_{1\text{-}562}$ proteoform and a C-terminal MF2.P2$_{563\text{-}701}$ proteoform (Fig. S3A). This cleavage site does not coincide with any of the putative disordered regions predicted for MF2. The IIKKQLQLSA site in MF2 is 100% conserved (regarding the considered 10 aa-stretch) in the MHP2 and MHJ2 orthologs and none of the mapped peptides covering this region for both was indicative of a non-tryptic cleavage event (Fig. S3B). This suggests differential endoproteolytic processing involving OG2 between *M. flocculare* and *M. hyopneumoniae*. Moreover, the absence of any semi-tryptic peptide covering the IIKKQLQLSA site for MF2 in cytoplasmic samples suggests differential endoproteolytic processing for this *M. flocculare* PUF also between subcellular fractions (Fig. S3C).

For OG3 in surface samples, endoproteolytic processing was found only in MF3 (Fig. S4A; Table S4C). A single non-tryptic cleavage site was identified for MF3 (KNQLE↓QVDFS), resulting potentially in a N-terminal MF3.P1$_{1\text{-}187}$ proteoform and a C-terminal MF3.

$P2_{188-278}$ proteoform (Fig. S4A). This non-tryptic cleavage site coincides with a disordered region (185–200) predicted for MF3. The KNQLEQVDFS site in MF3 is partially conserved (80% identity) in MHP3 and MHJ3 orthologs, and was covered by mapped peptides (Fig. S4B). This suggests differential endoproteolytic processing in the OG3, between *M. flocculare* and *M. hyopneumoniae*. Moreover, the absence of any semi-tryptic peptide mapped to the KNQLEQVDFS site for MF3 in cytoplasmic samples suggests differential endoproteolytic processing for this *M. flocculare* PUF also between subcellular fractions (Fig. S4C). The overall results providing evidence of differential endoproteolytic processing involving the PUF OGs1-5 are summarized in Table S6.

### 3.5. Epitope predictions for *M. hyopneumoniae* and *M. flocculare* PUFs and derived surface proteoforms

*M. hyopneumoniae* and *M. flocculare* PUFs and derived surface proteoforms were assessed for the prediction and mapping of epitopes, and to estimate EC. Predicted epitopes and their positions in the assessed PUFs are shown in Table S7A. For *M. hyopneumoniae* 7448, *M. hyopneumoniae* J and *M. flocculare* PUFs from OGs1-5, from 7 to 37, 7 to 34, and 8 to 37 epitopes were predicted, respectively. The lengths of these epitopes ranged from 7 to 53 aa, for *M. hyopneumoniae* 7448, 7 to 55 aa, for *M. hyopneumoniae* J, and 7 to 48 aa, for *M. flocculare* (Table S7A) and they are evenly distributed along the proteins/proteoforms (data not shown). Several predicted epitopes were conserved among the ortholog PUFs, predominantly between *M. hyopneumoniae* 7448 and J. Interestingly, some of the predicted epitopes were found within the RM1 and RM2 repeated motifs (highlighted in Table S7A).

ECs were then assessed, as at least some ortholog proteins/proteoforms differ considerably in length. Considering the full-length sequences of PUFs in OGs1-5, the ECs for *M. hyopneumoniae* (7448 and J) and *M. flocculare* PUFs were OG1 > OG3 > OG5 > OG4 > OG2 (Fig. 3). At least 60% of the MHP1, MHJ1 and MF1 sequences were predicted as antigenic, with MHP1 presenting the highest EC (68%). Within each OG, *M. hyopneumoniae* 7448 PUFs presented the highest ECs in comparison with the *M. hyopneumoniae* J and *M. flocculare* ortholog PUFs. In the OG5, MHP5 and MHJ5 presented the same ECs, as they are 100% identical. EC differences were especially evident between *M. hyopneumoniae* 7448 and *M. flocculare* in all OGs, being always lower in the nonpathogenic species.

ECs were also assessed for the derived surface proteoforms of *M. hyopneumoniae* 7448 and *M. flocculare*. As shown in Table S7B, the ECs of the MHP1, MF2, and MF3 derived proteoforms ranged from 59% to 78%, 36%–39% and 31%–58%, respectively.

### 3.6. Phylogenetic analysis of the assessed ortholog surface PUFs

Phylogenetic trees were constructed in order to evaluate the evolutionary relationship among *M. hyopneumoniae* 7448 PUFs from OG1-5 and orthologs from other bacteria. With exception of *M. hyopneumoniae*, only one representative strain of each species was considered (Table S8), selected based on the higher identity and coverage of the orthologs with the query sequence (data not shown). Orthologs of MHP1-5 were found in totals of 4, 13, 17, 10 and 32 other bacterial species, respectively (Table S8). Orthologs were found essentially among some *Mycoplasma* species, with the exceptions of MHP3 and MHP5, for which orthologs were found also in a few bacteria from other unrelated genera. Leucine-rich repeat domains (LRRs) (Pfam CL0022) were observed among MHP3 orthologs, and two conserved domains with unknown functions (DUFs), namely DUF31 (Pfam PF01732) and DUF2714 (Pfam PF10896), were identified among several MHP4 and MHP5 orthologs, respectively (Table S8).

In the phylogenetic analysis, the MHP1 and MHP5 alignments generated by our select dataset did not support dependable trees. Phylogenetic trees were then built only for MHP2, MHP3 and MHP4 orthologs, and they are presented in Figs. S5–S7, respectively. As expected, these trees demonstrated a close relationship between the MHP2-4 PUFs and their corresponding *M. hyopneumoniae* J orthologs, which grouped into monophyletic clusters in all cases. Furthermore, other clusters including *M. flocculare*, *Mycoplasma* spp., *Mycoplasma dispar*, *Mycoplasma ovipneumoniae* and *Mycoplasma* sp. Moose_RK orthologs displayed a similar and close evolutionary history, with paraphyletic relationships within the clade of species that colonize their hosts' respiratory system.

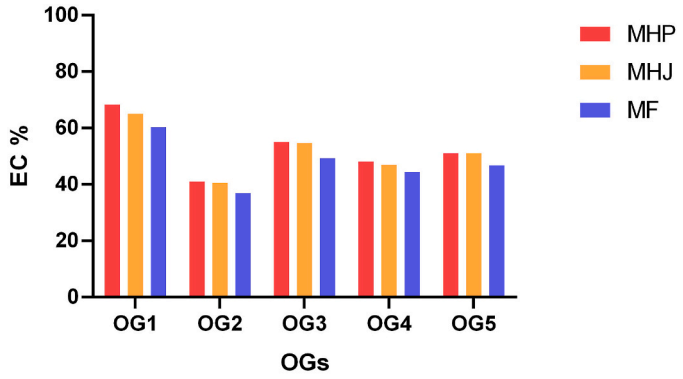


**Fig. 3.** ECs of the full-length amino acid sequences of PUFs from OGs1-5. ECs were measured for *M. hyopneumoniae* 7448 (MHP), *M. hyopneumoniae* J (MHJ) and *M. flocculare* (MF) PUFs in all OGs.

## 4. Discussion

*M. hyopneumoniae* surface-exposed proteins are protagonists in the pathogen-host relationship, due to their roles in adhesion, cytotoxicity, and immunomodulation of host response [10]. Interestingly, a large fraction of these surface proteins still remains annotated as PUFs [20], and are likely a reservoir of novel uncharacterized PEP determinants. Several studies have pointed out that PUFs may play relevant roles in the biology of many bacterial species, including *Mycoplasma* spp., and their characterization may be useful to reveal novel virulence factors, and new targets for the development of vaccines and diagnostic tests [20,22–24].

Here, we addressed the five most represented PUFs on the surface of the pathogenic *M. hyopneumoniae* 7448 and their orthologs from nonpathogenic *M. hyopneumoniae* J strain and *M. flocculare*. These surface dominant proteins were selected for the present study based on previously available proteomic quali-quantitative data generated by our group [21]. Surface-displayed molecules, specially the more abundant ones, are expected to play key roles during the infectious process of pathogenic bacteria, as they directly interact with host molecules in the extracellular milieu [25]. In line with that, PUFs overrepresented on *M. hyopneumoniae* surface may be involved in pathogen-host interactions relevant to pathogenicity, including, but not limited to, cell adhesion and invasion, antigenic variation, immunomodulation, detoxification, nutrient uptake, cell signaling, biofilm formation and so on [10,26,27]. Interestingly, our results demonstrated that the corresponding ortholog PUFs from *M. hyopneumoniae* J and *M. flocculare* are relatively less abundant on the cell surface, which is also suggestive of the importance of these proteins for the *M. hyopneumoniae* pathobiology. Moreover, endoproteolytic processing has been also investigated and revealed proteoforms derived from several surface proteins, including some PUF-derived [14,17–19,28,29]. Endoproteolytic processing is an important mechanism to modulate the protein content of the bacteria surface, with a direct impact on the interactions with the host. In our findings, the ortholog PUFs and their potential proteoforms differ in structural aspects that may influence the host-pathogen interaction, as discussed below.

Among the dominant PUFs in the *M. hyopneumoniae* surface and their *M. flocculare* orthologs, those in the assessed OGs1-5 differ in the type, the number and/or the distribution of protein domains. They occur in the form of amino acid stretches that are present/absent or are at least 70% divergent in sequence. The differences in domains were found within all OGs between the *M. hyopneumoniae* strains and *M. flocculare*, but were especially evident between these orthologs in OG1. Although these domains have not been functionally characterized so far, their differential presence or conservation among orthologs of surface full-length proteins or derived proteoforms, points out to functional differences with impact in host-pathogen interactions. Indeed, a previous study showed that differential domains between ortholog surface proteins from *M. hyopneumoniae* and *M. flocculare* induced differential immune responses in immunized mice [22].

Further structural findings came from the predictions of disordered regions and repeated motifs. Putative disordered regions were identified for all ortholog PUFs from OGs1-5. Among those endoproteolytically processed, only MF1 (NFKSL↓SQSQK) and MF3 (KNQLE↓QVDFS) were processed in predicted regions of disorder. It is likely that these regions are more susceptible to post-translational modifications, including protease processing [30]. Disordered protein stretches are assumed to be more relaxed and consequently more exposed in the three-dimensional structure of proteins. Endoproteollytic processing associated with putative disordered regions has been also described for adhesins from *Mycoplasma* species [17,31–33].

Some of the assessed ortholog surface PUFs also presented differential repeated motifs. Such motifs were found within members of OG1, in the form of two repeat units named RM1 and RM2. These repeats are highly conserved (up to 100% identical) between MHP1 and MHJ1, but are absent or are less conserved (from 13 to 93% of identity) in the RM1 or RM2 repeats of *M. flocculare* orthologs. Besides these differences in sequences, there were also differences in the number and the distribution of these repeats among the proteins within OG1 and among the MHP1 and MF1-derived proteoforms, which may determine differential molecular interactions. Repeated sequences have been reported for some characterized proteins with important functional implications in *M. hyopneumoniae*, such as P97 and other adhesion-related proteins [34,35], and the membrane-associated lipoprotein P68 [36]. Moreover, motifs have been also identified in other *Mycoplasma* species for many surface proteins related to adherence to the host and virulence [33,37,38].

We also compared the endoproteolytic processing between the strains/species and subcellular compartments for *M. hyopneumoniae* 7448 (MHP1), and *M. flocculare* (MF1, MF2 and MF3). The identified cleavage sites in these PUFs presented no conservation with processed regions previously described for other *M. hyopneumoniae* proteins [14,17–19,28,29]. It has been demonstrated that several *M. hyopneumoniae* J surface proteins may also undergo endoproteolytic processing [14,19,39], with MHJ1 among them [39]. However, no such processing was detected neither for MHJ1 nor to MHJ2-5 in our findings. Several proteases from *M. hyopneumoniae* 7448 and J strains and *M. flocculare* have been identified and a few have been functionally characterized [8,9,40]. However, it remains unclear if any of them or other so far uncharacterized ones are responsible for the observed differential endoproteolytic processing. How the differential endoproteolytic processing is regulated and the importance of this processing for virulence-related aspects deserve more investigation.

Structural differences based on repeat variation and/or post-translational proteolytic processing have also been associated with mechanisms of *M. hyopneumoniae* surface protein variation [2,35]. The ECs of the full-length assessed PUFs suggest that those from the pathogenic *M. hyopneumoniae* 7448 strain are potentially more antigenic than their orthologs from the nonpathogenic *M. hyopneumoniae* J strain and *M. flocculare*. These possible differences of antigenicity in PUFs well represented on the bacterial surface may mediate differential interactions with the host immune system. Antigenic variation was also observed in repeating units between the MHP1 and MF1 orthologs. Such differences likely have functional significance, as they may affect not only protein antigenicity/immunogenicity, but also other potential functions, including cell adhesion or transport [10,35,41]. The derived surface proteoforms of MHP1, MF2 and MF3 may also play differential roles in mycoplasma-host interactions. By being exposed on the mycoplasma surface and/or being released in the extracellular milieu, such endoproteolitically-generated proteoforms may directly interact with host molecules. For instance, the swine immune response could be modulated by these antigenic surface proteoforms.

Therefore, future experimental assays will be required to shed light on the functions carried out by these PUFs and their relationship with *M. hyopneumoniae* pathogenicity. Anyhow, as likely immunogenic, these dominant surface PUFs and their derived proteoforms deserve also future characterization as possible vaccine components.

Although MHP1-5 have not been so far functionally characterized by wet-lab experiments, and formally remain as PUFs, MHP1, MHP2 and MHP3 were recently reannotated based on *in silico* predictions as a putative DNA-binding protein, a lipoprotein and a membrane transporter, respectively [20]. DNA-binding proteins may act in the maintenance of chromosome structure, in the regulation of gene expression, or even mediate biofilm formation [27,42]. However, just a few DNA-binding proteins have been so far annotated in the *M. hyopneumoniae* genome [7–9]. Lipoproteins also have been related to pathogenicity and virulence of *Mycoplasma* species, as mediators of complex host-pathogen interactions, from adhesion to cytotoxicity or immunomodulation [36,43,44]. Membrane transporters, in turn, have also been associated with *M. hyopneumoniae* virulence and drug resistance [45,46]. Thus, it will be interesting to functionally characterize MHP1, MHP2 and MHP3, as novel potential, DNA-binding protein, lipoprotein or membrane transporter, respectively. Additionally, MHP1, MHP2 and MHP5 were *in silico* predicted as putative virulence factors [20]. In this context, the functional characterization of these PUFs will also help to investigate their relevance for bacterial virulence.

Finally, phylogenetic analyses were carried out to understand the evolutionary history of the MHP1-5 and ortholog proteins found in other bacteria. The ortholog PUFs identified for MHP1-5 were predominantly found among around 40 *Mycoplasma* species, but only four, namely *M. dispar*, *M. ovipneumoniae*, *M. flocculare* and *Mycoplasma* sp. 'Moose RK' presented orthologs for all the assessed *M. hyopneumoniae* PUFs. Interestingly, these four species are all found in the respiratory tract (of cattle, goat, pig or moose, respectively), and it will be interesting to investigate the possible function of proteins from the OGs1-5 for the specific colonization and/or infection of respiratory tissues.

No further putative functional inferences came from the found MHP1-5 orthologs. However, a couple of domains were shared among many of the identified ortholog PUFs, namely LRR and DUF domains. LRR domains have been associated with protein-protein interactions [47]. The identified DUF31 has been associated with peptidase function [48]. No function has been so far associated with the identified DUF2714 domain, but its occurrence is in the Mycoplasmastaceae family.

Phylogenetic inferences were not possible for MHP1 and MHP5, likely due to incongruences determined by biological issues involving these sequences, such as horizontal transfer or recombination [49]. However, such inferences were successfully established for MHP2, MHP3 and MHP4 orthologs, describing the evolutive history of these mycoplasma PUFs.

## 5. Conclusions

The five most abundant PUFs on *M. hyopneumoniae* 7448 surface and their orthologs from *M. hyopneumoniae* J and *M. flocculare* were compared and characterized regarding putative functional domains, disordered regions, repeated motifs, and endoproteolytic processing. Differences have been found in at least some of the investigated OGs and derived proteoforms. Such differences were discussed in terms of their association with antigenic and other functional variations, likely correlated to virulence and pathogenicity. Phylogenetic analyses showed that these *M. hyopneumoniae* PUFs and their orthologs are mostly restricted to the Mycoplasmataceae family, and that at least three of them are shared by species that colonize/infect the mammal respiratory tract. Overall, our results pointed out the relevance of these PUFs for *M. hyopneumoniae* pathobiology. Future wet lab experiments will be necessary to address these PUFs and their derived proteoforms regarding protein-protein interactions and immunological properties, with implications for their association with pathogenicity and virulence, and for their potential use as diagnostic or vaccine antigens.

## Author contribution statement

Priscila Souza dos Santos: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Jéssica Andrade Paes: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Lais Del Prá Netto Machado; Gabriela Prado Paludo: Performed the experiments; Analyzed and interpreted the data.

Arnaldo Zaha; Henrique Bunselmeyer Ferreira: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

## Data availability statement

Data included in article/supp. material/referenced in article.

## Additional information

Supplementary content related to this article has been published online at [URL].

## Funding

PSS is a recipient of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil) Ph.D. fellowship (grant number 88887.353126/2019-00). JAP is a recipient of Programa Nacional de Pós-Doutorado/CAPES (PNPD/CAPES, Brazil) post-doctoral fellowship. LDPNM was a recipient of Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil) Ph.D.

fellowship (grant number 303287/2017-9). GPP was a recipient of a Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) post-doctoral fellowship DTI-1. AZ was a recipient of a CNPq Research Productivity Fellowship (grant number 304551/2017-1). HBF is a recipient of a CNPq Research Productivity Fellowship (grant number 305083/2021-0).

## Declaration of competing interest

The authors declare no conflict of interest.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e16141.
ND, not detected in surface samples according to Paes et al. [15].

## References

[1] D. Maes, M. Sibila, P. Kuhnert, J. Segalés, F. Haesebrouck, M. Pieters, Update on Mycoplasma hyopneumoniae infections in pigs: knowledge gaps for improved disease control, Transbound. Emerg. Dis. 65 (2018) 110–124, https://doi.org/10.1111/tbed.12677.
[2] A.M. Betlach, D. Maes, L. Garza-Moreno, P. Tamiozzo, M. Sibila, F. Haesebrouck, J. Segalés, M. Pieters, Mycoplasma hyopneumoniae variability: current trends and proposed terminology for genomic classification, Transbound. Emerg. Dis. 66 (2019) 1840–1854, https://doi.org/10.1111/tbed.13233.
[3] M. Pieters, D. Maes, Mycoplasmosis, in: JJ. Zimmermann, L.A Karriker, R Ramirez, KJ Schwartz, GW Stevenson, J Zhang (Eds.), Mycoplasmosis, 11th edn, Wiley, Hoboken, NJ, 2019, pp. 863–883. https://doi.org/10.1002/9781119350927.ch56.
[4] G. Saade, C. Deblanc, J. Bougon, C. Marois-Créhan, C. Fablet, G. Auray, C. Belloc, M. Leblanc-Maridor, C.A. Gagnon, J. Zhu, M. Gottschalk, A. Summerfield, G. Simon, N. Bertho, F. Meurens, Coinfections and their molecular consequences in the porcine respiratory tract, Vet. Res. 51 (2020) 1–19, https://doi.org/10.1186/s13567-020-00807-8.
[5] D. Maes, F. Boyen, B. Devriendt, P. Kuhnert, A. Summerfield, F. Haesebrouck, Perspectives for improvement of Mycoplasma hyopneumoniae vaccines in pigs, Vet. Res. 52 (2021) 1–20, https://doi.org/10.1186/s13567-021-00941-x.
[6] Z. Li, Y. Wang, Y. Zhang, X. Tang, X. Wang, W. Liu, Y. Qian, Y. Zhu, H. Chen, C. Tan, Attenuation of mycoplasma hyopneumoniae strain ES-2 and comparative genomic analysis of ES-2 and its attenuated form ES-2L, Front. Vet. Sci. 8 (2021) 1–9, https://doi.org/10.3389/fvets.2021.696262.
[7] F.C. Minion, E.J. Lefkowitz, M.L. Madsen, B.J. Cleary, S.M. Swartzell, G.G. Mahairas, The genome sequence of Mycoplasma hyopneumoniae strain 232, the agent of swine mycoplasmosis, J. Bacteriol. 186 (2004) 7123–7133, https://doi.org/10.1128/JB.186.21.7123-7133.2004.
[8] F.M. Siqueira, C.E. Thompson, V.G. Virginio, T. Gonchoroski, L. Reolon, L.G. Almeida, M.M. da Fonsêca, R. de Souza, F. Prosdocimi, I.S. Schrank, H.B. Ferreira, A.T.R. de Vasconcelos, A. Zaha, New insights on the biology of swine respiratory tract mycoplasmas from a comparative genome analysis, BMC Genom. 14 (2013) 1–17, https://doi.org/10.1186/1471-2164-14-175.
[9] A.T.R. Vasconcelos, H.B. Ferreira, C.V. Bizarro, S.L. Bonatto, M.O. Carvalho, P.M. Pinto, D.F. Almeida, L.G.P. Almeida, R. Almeida, L. Alves-Filho, E. N. Assunção, V.A.C. Azevedo, M.R. Bogo, M.M. Brigido, M. Brocchi, H.A. Burity, A.A. Camargo, S.S. Camargo, M.S. Carepo, D.M. Carraro, J.C. De Mattos Cascardo, L.A. Castro, G. Cavalcanti, G. Chemale, R.G. Collevatti, C.W. Cunha, B. Dallagiovanna, B.P. Dambrós, O.A. Dellagostin, C. Falcão, F. Fantinatti-Garboggini, M.S.S. Felipe, L. Fiorentin, G.R. Franco, N.S.A. Freitas, D. Frías, T.B. Grangeiro, E.C. Grisard, C.T. Guimarães, M. Hungria, S.N. Jardim, M.A. Krieger, J.P. Laurino, L.F.A. Lima, M.I. Lopes, É.L.S. Loreto, H.M.F. Madeira, G.P. Manfio, A.Q. Maranhão, C.T. Martinkovics, S.R.B. Medeiros, M.A.M. Moreira, M. Neiva, C.E. Ramalho-Neto, M.F. Nicolás, S.C. Oliveira, R.F.C. Paixão, F.O. Pedrosa, S.D.J. Pena, M. Pereira, L. Pereira-Ferrari, I. Piffer, L.S. Pinto, D.P. Potrich, A.C. M. Salim, F.R. Santos, R. Schmitt, M.P.C. Schneider, A. Schrank, I.S. Schrank, A.F. Schuck, H.N. Seuanez, D.W. Silva, R. Silva, S.C. Silva, C.M.A. Soares, K.R. L. Souza, R.C. Souza, C.C. Staats, M.B.R. Steffens, S.M.R. Teixeira, T.P. Urmenyi, M.H. Vainstein, L.W. Zuccherato, A.J.G. Simpson, A. Zaha, Swine and poultry pathogens: the complete genome sequences of two strains of Mycoplasma hyopneumoniae and a strain of Mycoplasma synoviae, J. Bacteriol. 187 (2005) 5568–5577, https://doi.org/10.1128/JB.187.16.5568-5577.2005.
[10] F.M.A. Leal Zimmer, J.A. Paes, A. Zaha, H.B. Ferreira, Pathogenicity & virulence of Mycoplasma hyopneumoniae, Virulence 11 (2020) 1600–1622, https://doi.org/10.1080/21505594.2020.1842659.
[11] S.P. Djordjevic, S.J. Cordwell, A. Michael, J. Wilton, F.C. Minion, M.A. Djordjevic, Proteolytic processing of the mycoplasma hyopneumoniae cilium adhesin proteolytic processing of the mycoplasma hyopneumoniae cilium adhesin, Infect. Immun. 72 (2004) 2791–2802, https://doi.org/10.1128/IAI.72.5.2791.
[12] P.M. Pinto, G. Chemale, L.A. de Castro, A.P.M. Costa, J.D. Kich, M.H. Vainstein, A. Zaha, H.B. Ferreira, Proteomic survey of the pathogenic Mycoplasma hyopneumoniae strain 7448 and identification of novel post-translationally modified and antigenic proteins, Vet. Microbiol. 121 (2007) 83–93, https://doi.org/10.1016/j.vetmic.2006.11.018.
[13] P.M. Pinto, C.S. Klein, A. Zaha, H.B. Ferreira, Comparative proteomic analysis of pathogenic and non-pathogenic strains from the swine pathogen Mycoplasma hyopneumoniae, Proteome Sci. 7 (2009) 1–11, https://doi.org/10.1186/1477-5956-7-45.
[14] J.L. Tacchi, B.B.A. Raymond, P.A. Haynes, I.J. Berry, M. Widjaja, D.R. Bogema, L.K. Woolley, C. Jenkins, F.C. Minion, M.P. Padula, S.P. Djordjevic, Post-translational processing targets functionally diverse proteins in Mycoplasma hyopneumoniae, Open Biol 6 (2016), https://doi.org/10.1098/rsob.150210.
[15] J. Wilton, C. Jenkins, S.J. Cordwell, L. Falconer, F.C. Minion, D.C. Oneal, M.A. Djordjevic, A. Connolly, I. Barchia, M.J. Walker, S.P. Djordjevic, Mhp493 (P216) is a proteolytically processed, cilium and heparin binding protein of Mycoplasma hyopneumoniade, Mol. Microbiol. 71 (2009) 566–582, https://doi.org/10.1111/j.1365-2958.2008.06546.x.
[16] L.M. Seymour, A.T. Deutscher, C. Jenkins, T.A. Kuit, L. Falconer, F.C. Minion, B. Crossett, M. Padula, N.E. Dixon, S.P. Djordjevic, M.J. Walker, A processed multidomain Mycoplasma hyopneumoniae adhesin binds fibronectin, plasminogen, and swine respiratory cilia, J. Biol. Chem. 285 (2010) 33971–33978, https://doi.org/10.1074/jbc.M110.104463.
[17] D.R. Bogema, A.T. Deutscher, L.K. Woolley, L.M. Seymour, B.B.A. Raymond, J.L. Tacchi, M.P. Padula, N.E. Dixon, F.C. Minion, C. Jenkins, M.J. Walker, S. P. Djordjevic, Characterization of cleavage events in the multifunctional cilium adhesin Mhp684 (P146) reveals a mechanism by which Mycoplasma hyopneumoniae regulates surface topography, mBio 3 (2012) 1–11, https://doi.org/10.1128/mBio.00282-11.
[18] A.T. Deutscher, J.L. Tacchi, F.C. Minion, M.P. Padula, B. Crossett, D.R. Bogema, C. Jenkins, T.A. Kuit, M.J. Walker, S.P. Djordjevic, Mycoplasma hyopneumoniae surface proteins Mhp385 and Mhp384 bind host cilia and glycosaminoglycans and are endoproteolytically processed by proteases that recognize different cleavage motifs, J. Proteome Res. 11 (2012) 1924–1936, https://doi.org/10.1021/pr201115v.
[19] L.D.P.N. Machado, J.A. Paes, P. Souza dos Santos, H.B. Ferreira, Evidences of differential endoproteolytic processing on the surfaces of Mycoplasma hyopneumoniae and Mycoplasma flocculare, Microb. Pathog. 140 (2020), 103958, https://doi.org/10.1016/j.micpath.2019.103958.
[20] B.A. da R. Tavares, J.A. Paes, A. Zaha, H.B. Ferreira, Reannotation of Mycoplasma hyopneumoniae hypothetical proteins revealed novel potential virulence factors, Microb. Pathog. 162 (2022), 105344, https://doi.org/10.1016/j.micpath.2021.105344.

[21] J.A. Paes, L.D.P.N. Machado, F.M.D.A. Leal, S.N. De Moraes, H. Moura, J.R. Barr, H.B. Ferreira, Comparative proteomics of two mycoplasma hyopneumoniae strains and mycoplasma flocculare identified potential porcine enzootic pneumonia determinants, Virulence 9 (2018) 1230–1246, https://doi.org/10.1080/21505594.2018.1499379.

[22] F.M. dos, A. Leal, V.G. Virginio, C.L. Martello, J.A. Paes, T.J. Borges, N. Jaeger, C. Bonorino, H.B. Ferreira, Mycoplasma hyopneumoniae and Mycoplasma flocculare differential domains from orthologous surface proteins induce distinct cellular immune responses in mice, Vet. Microbiol. 190 (2016) 50–57, https://doi.org/10.1016/j.vetmic.2016.05.008.

[23] M. Shahbaaz, K. Bisetty, F. Ahmad, M.I. Hassan, Silico approaches for the identification of virulence candidates amongst hypothetical proteins of Mycoplasma pneumoniae 309, Comput. Biol. Chem. 59 (2015) 67–80, https://doi.org/10.1016/j.compbiolchem.2015.09.007.

[24] T. Cristina, V. Rodrigues, K. Jaiswal, A. De Sarom, L. De Castro, C. José, F. Oliveira, P. Ghosh, S. Tiwari, F.M. Miranda, L. De, J. Benevides, V. Ariston, D. C. Azevedo, Reverse Vaccinology and Subtractive Genomics Reveal New Therapeutic Targets against Mycoplasma Pneumoniae : a Causative Agent of Pneumonia, 2019.

[25] D. Ribet, P. Cossart, How bacterial pathogens colonize their hosts and invade deeper tissues, Microb. Infect. 17 (2015) 173–183, https://doi.org/10.1016/j.micinf.2015.01.004.

[26] P. Li, Y. Zhang, X. Li, W. Zhou, X. Li, F. Jiang, W. Wu, Mycoplasma hyopneumoniae Mhp597 is a cytotoxicity, inflammation and immunosuppression associated nuclease, Vet. Microbiol. 235 (2019) 53–62, https://doi.org/10.1016/j.vetmic.2019.05.011.

[27] B.B.A. Raymond, C. Jenkins, L. Turnbull, C.B. Whitchurch, S.P. Djordjevic, Extracellular DNA release from the genome-reduced pathogen Mycoplasma hyopneumoniae is essential for biofilm formation on abiotic surfaces, Sci. Rep. 8 (2018) 1–12, https://doi.org/10.1038/s41598-018-28678-2.

[28] T.A. Burnett, K. Dinkla, M. Rohde, G.S. Chhatwal, C. Uphoff, M. Srivastava, S.J. Cordwell, S. Geary, X. Liao, F.C. Minion, M.J. Walker, S.P. Djordjevic, P159 is a proteolytically processed, surface adhesin of Mycoplasma hyopneumoniae: defined domains of P159 bind heparin and promote adherence to eukaryote cells, Mol. Microbiol. 60 (2006) 669–686, https://doi.org/10.1111/j.1365-2958.2006.05139.x.

[29] B.B.A. Raymond, J.L. Tacchi, V.M. Jarocki, F.C. Minion, M.P. Padula, S.P. Djordjevic, P159 from Mycoplasma hyopneumoniae binds porcine cilia and heparin and is cleaved in a manner akin to ectodomain shedding, J. Proteome Res. 12 (2013) 5891–5903, https://doi.org/10.1021/pr400903s.

[30] R. Van Der Lee, M. Buljan, B. Lang, R.J. Weatheritt, G.W. Daughdrill, A.K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D.T. Jones, P.M. Kim, R.W. Kriwacki, C. J. Oldfield, R.V. Pappu, P. Tompa, V.N. Uversky, P.E. Wright, M.M. Babu, Classification of intrinsically disordered regions and proteins, Chem. Rev. 114 (2014) 6589–6631, https://doi.org/10.1021/cr400525m.

[31] D.R. Bogema, N.E. Scott, M.P. Padula, J.L. Tacchi, B.B.A. Raymond, C. Jenkins, S.J. Cordwell, F.C. Minion, M.J. Walker, S.P. Djordjevic, Sequence TTKF ↓ QE defines the site of proteolytic cleavage in Mhp683 protein, a novel glycosaminoglycan and cilium adhesin of Mycoplasma hyopneumoniae, J. Biol. Chem. 286 (2011) 41217–41229, https://doi.org/10.1074/jbc.M111.226084.

[32] J.L. Tacchi, B.B.A. Raymond, V.M. Jarocki, I.J. Berry, M.P. Padula, S.P. Djordjevic, Cilium Adhesin P216 (MHJ-0493) is a target of ectodomain shedding and aminopeptidase activity on the surface of mycoplasma hyopneumoniae, J. Proteome Res. 13 (2014) 2920–2930, https://doi.org/10.1021/pr500087c.

[33] M. Widjaja, I.J. Berry, V.M. Jarocki, M.P. Padula, R. Dumke, S.P. Djordjevic, Cell surface processing of the P1 adhesin of Mycoplasma pneumoniae identifies novel domains that bind host molecules, Sci. Rep. 10 (2020) 1–16, https://doi.org/10.1038/s41598-020-63136-y.

[34] F.C. Minion, C. Adams, T. Hsu, R1 region of P97 mediates adherence of Mycoplasma hyopneumoniae to swine cilia, Infect. Immun. 68 (2000) 3056–3060, https://doi.org/10.1128/IAI.68.5.3056-3060.2000.

[35] L.A. de Castro, T.R. Pedroso, S.S. Kuchiishi, M. Ramenzoni, J.D. Kich, A. Zaha, M.H. Vainstein, H.B. Ferreira, Variable number of tandem aminoacid repeats in adhesion-related CDS products in Mycoplasma hyopneumoniae strains, Vet. Microbiol. 116 (2006) 258–269, https://doi.org/10.1016/j.vetmic.2006.04.022.

[36] W. Liu, D. Zhou, F. Yuan, Z. Liu, Z. Duan, K. Yang, R. Guo, M. Li, S. Li, L. Fang, S. Xiao, Y. Tian, Surface proteins mhp390 (P68) contributes to cilium adherence and mediates inflammation and apoptosis in Mycoplasma hyopneumoniae, Microb. Pathog. 126 (2019) 92–100, https://doi.org/10.1016/j.micpath.2018.10.035.

[37] S. Pereyre, P. Sirand-Pugnet, L. Beven, A. Charron, H. Renaudin, A. Barré, P. Avenaud, D. Jacob, A. Couloux, V. Barbe, A. De Daruvar, A. Blanchard, C. Bébéar, Life on arginine for Mycoplasma hominis: clues from its minimal genome and comparison with other human urogenital mycoplasmas, PLoS Genet. 5 (2009), https://doi.org/10.1371/journal.pgen.1000677.

[38] K. Sachse, J.H. Helbig, I. Lysnyansky, C. Grajetzki, W. Müller, E. Jacobs, D. Yogev, Epitope mapping of immunogenic and adhesive structures in repetitive domains of Mycoplasma bovis variable surface lipoproteins, Infect. Immun. 68 (2000) 680–687, https://doi.org/10.1128/IAI.68.2.680-687.2000.

[39] I.J. Berry, V.M. Jarocki, J.L. Tacchi, B.B.A. Raymond, M. Widjaja, M.P. Padula, S.P. Djordjevic, N-terminomics identifies widespread endoproteolysis and novel methionine excision in a genome-reduced bacterial pathogen, Sci. Rep. 7 (2017) 1–17, https://doi.org/10.1038/s41598-017-11296-9.

[40] L. Moitinho-Silva, M.Y. Kondo, L.C.G. Oliveira, D.N. Okamoto, J.A. Paes, M.F.M. Machado, C.L. Veronez, G. Motta, S.S. Andrade, M.A. Juliano, H.B. Ferreira, L. Juliano, I.E. Gouvea, Mycoplasma hyopneumoniae in vitro peptidase activities: identification and cleavage of kallikrein-kinin system-like substrates, Vet. Microbiol. 163 (2013) 264–273, https://doi.org/10.1016/j.vetmic.2013.01.011.

[41] H.B. Ferreira, L.A. de Castro, A preliminary survey of M. hyopneumoniae virulence factors based on comparative genomic analysis, Genet. Mol. Biol. 30 (2007) 245–255, https://doi.org/10.1590/s1415-47572007000200012.

[42] J. Hołówka, J. Zakrzewska-Czerwińska, Nucleoid associated proteins: the small organizers that help to cope with stress, Front. Microbiol. 11 (2020) 1–7, https://doi.org/10.3389/fmicb.2020.00590.

[43] A. Christodoulides, N. Gupta, V. Yacoubian, N. Maithel, J. Parker, T. Kelesidis, The role of lipoproteins in mycoplasma-mediated immunomodulation, Front. Microbiol. 9 (2018) 1–9, https://doi.org/10.3389/fmicb.2018.01682.

[44] F. Bai, B. Ni, M. Liu, Z. Feng, Q. Xiong, G. Shao, Mycoplasma hyopneumoniae-derived lipid-associated membrane proteins induce inflammation and apoptosis in porcine peripheral blood mononuclear cells in vitro, Vet. Microbiol. 175 (2015) 58–67, https://doi.org/10.1016/j.vetmic.2014.11.013.

[45] X. Xia, L. Yang, Y. Ling, J. Yu, H. Ding, Emergence and mechanism of resistance of tulathromycin against mycoplasma hyopneumoniae in a PK/PD model and the fitness costs of 23S rRNA mutants, Front. Vet. Sci. 9 (2022) 1–14, https://doi.org/10.3389/fvets.2022.801800.

[46] M. Galvao Ferrarini, S.G. Mucha, D. Parrot, G. Meiffrein, J.F. Ruggiero Bachega, G. Comte, A. Zaha, M.F. Sagot, Hydrogen peroxide production and myo-inositol metabolism as important traits for virulence of Mycoplasma hyopneumoniae, Mol. Microbiol. 108 (2018) 683–696, https://doi.org/10.1111/mmi.13957.

[47] P. Enkhbayar, M. Kamiya, M. Osaki, T. Matsumoto, N. Matsushima, Structural principles of leucine-rich repeat (LRR) proteins, Proteins Struct. Funct. Genet. 54 (2004) 394–403, https://doi.org/10.1002/prot.10605.

[48] N.F. Goodacre, D.L. Gerloff, P. Uetz, Protein domains of unknown function are essential in bacteria, mBio 5 (2013) 1–6, https://doi.org/10.1128/mBio.00744-13.

[49] A. Som, Causes, consequences and solutions of phylogenetic incongruence, Briefings Bioinf. 16 (2014) 536–548, https://doi.org/10.1093/bib/bbu015.