



## Article

# Genomic Analysis of *Pasteurella atlantica* Provides Insight on Its Virulence Factors and Phylogeny and Highlights the Potential of Reverse Vaccinology in Aquaculture

Rebecca Marie Ellul <sup>1,†</sup> , Panos G. Kalatzis <sup>2,†</sup> , Cyril Frantzen <sup>3</sup>, Gyri Teien Haugland <sup>1</sup>, Snorre Gulla <sup>4</sup> , Duncan John Colquhoun <sup>1,4</sup>, Mathias Middelboe <sup>2</sup> , Heidrun Inger Wergeland <sup>1</sup> and Anita Rønneseth <sup>1,\*</sup>

<sup>1</sup> Department of Biological Sciences, University of Bergen, N-5006 Bergen, Norway; rebecca.ellul@uib.no (R.M.E.); gyri.haugland@uib.no (G.T.H.); Duncan.Colquhoun@vetinst.no (D.J.C.); heidrun.wergeland@uib.no (H.I.W.)

<sup>2</sup> Department of Biology, Marine Biological Section, University of Copenhagen, DK-3000 Helsingør, Denmark; panos.kalatzis@bio.ku.dk (P.G.K.); mmiddelboe@bio.ku.dk (M.M.)

<sup>3</sup> STIM/ACD Pharma AS, Karl Johans Gate 16, 0154 Oslo, Norway; cyril.frantzen@stim.no

<sup>4</sup> Norwegian Veterinary Institute, Ullevålsveien 68, P.O. Box 8146 Dep, N-0033 Oslo, Norway; snorre.gulla@vetinst.no

\* Correspondence: anita.ronneseth@uib.no

† These authors contributed equally to this work.



**Citation:** Ellul, R.M.; Kalatzis, P.G.; Frantzen, C.; Haugland, G.T.; Gulla, S.; Colquhoun, D.J.; Middelboe, M.; Wergeland, H.I.; Rønneseth, A. Genomic Analysis of *Pasteurella atlantica* Provides Insight on its Virulence Factors and Phylogeny and Highlights the Potential of Reverse Vaccinology in Aquaculture. *Microorganisms* **2021**, *9*, 1215. <https://doi.org/10.3390/microorganisms9061215>

Academic Editor: M. Javad Aman

Received: 7 May 2021

Accepted: 1 June 2021

Published: 4 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Pasteurellosis in farmed lumpstickers, *Cyclopterus lumpus*, has emerged as a serious disease in Norwegian aquaculture in recent years. Genomic characterization of the causative agent is essential in understanding the biology of the bacteria involved and in devising an efficient preventive strategy. The genomes of two clinical *Pasteurella atlantica* isolates were sequenced ( $\approx 2.3$  Mbp), and phylogenetic analysis confirmed their position as a novel species within the *Pasteurellaceae*. In silico analyses revealed 11 genomic islands and 5 prophages, highlighting the potential of mobile elements as driving forces in the evolution of this species. The previously documented pathogenicity of *P. atlantica* is strongly supported by the current study, and 17 target genes were recognized as putative primary drivers of pathogenicity. The expression level of a predicted vaccine target, an uncharacterized adhesin protein, was significantly increased in both broth culture and following the exposure of *P. atlantica* to lumpstickers head kidney leucocytes. Based on in silico and functional analyses, the strongest gene target candidates will be prioritized in future vaccine development efforts to prevent future pasteurellosis outbreaks.

**Keywords:** *Pasteurella atlantica*; lumpstickers; aquaculture; pathogenicity; phylogeny; in silico analysis; virulence factors; mobile elements; vaccine

## 1. Introduction

The prevalence of pasteurellosis in farmed lumpstickers, *Cyclopterus lumpus*, in Norway has increased in recent years since the first case was recorded in 2012 [1]. Despite a decrease in the number of affected localities [2], *Pasteurella atlantica* [3] remains a significant problem, and as the disease is non-notifiable, under-reporting of outbreaks cannot be excluded. Furthermore, while farmed lumpstickers are vaccinated against vibriosis and atypical furunculosis, there are no commercially available vaccines against pasteurellosis.

The family *Pasteurellaceae* is composed of commensals, opportunistic- and primary-pathogens and includes the genera *Pasteurella*, *Actinobacillus*, and *Hemophilus* amongst others. The genus *Pasteurella* has a broad host range, but little is known regarding the ecology of marine species. *P. atlantica* isolated from diseased Norwegian lumpstickers is related to but serologically distinct from *Pasteurella skyensis*, which causes disease in Atlantic salmon (*Salmo salar* L.) in Scotland, and a *P. atlantica* isolate first detected in 2018, which causes disease in Norwegian farmed Atlantic salmon [1,3,4]. Pasteurellosis affects all

life stages of a lump sucker, from fry to fish deployed in salmon cages [4]. As *P. atlantica* has been detected in lump sucker eggs and milt, vertical transmission may also be possible [5].

In a previous work [6], we showed that whole cell-inactivated bacterin-based vaccines do not provide adequate protection against the disease, despite the high titers of specific antibodies raised. Such a situation highlights the need for a deeper characterization of the bacterium and the identification of immunogenic and protective antigens.

Most pathogenic bacteria have several tools in their genetic arsenal to avoid host defenses and to enhance their survival. Genetic determinants including pathogenicity islands, antibiotic resistance genes, toxins, and adhesins are often shared between bacterial populations via horizontal gene transfer and are commonly associated with plasmids, prophages, and other mobile genetic elements. As adhesins are involved in the early stages of colonization, they can be utilized as targets for vaccine development via reverse vaccinology (RV).

Using *in silico* bioinformatic analyses, immunogenic antigens for vaccine development can be predicted [7,8]. RV is a rapid process and can reduce vaccine development time by up to 2 years [9]. It is also a more ethical approach due to a reduction in the number of experimental animals required for vaccine testing and results in effective vaccines. In human medicine, RV has been successful in the development of vaccines against several bacterial pathogens reviewed by Sharma et al. 2021 [10]. Pathogens of significant concern such as *Neisseria meningitidis* [11,12], *Mycobacterium tuberculosis* [13], *Chlamydia pneumoniae* [14], *Streptococcus pneumoniae* [15], *Helicobacter pylori* [16], *Porphyromonas gingivalis* [17], and *Bacillus anthracis* [18,19] have been investigated using RV.

RV technologies have improved considerably since the principles were initially developed and applied. The main criteria required for reliable prediction of vaccine targets include subcellular localization, the probability of adhesion functionality, and the number of transmembrane domains [20]. Additionally, using targets that are present only in virulent strains and having protein sequences that are dissimilar to host sequences results in more reliable vaccine targets that also induce strong immune responses [9].

Adhesins are a class of surface-bound proteins involved in facilitating bacterial attachment to host tissues. They are classed as fimbrial or non-fimbrial based on the absence or presence, respectively, of an outer membrane anchor in the protein. Among the different types of adhesins, bacterial lectins are the most common [15]. The mediation of attachment occurs through the recognition of specific carbohydrates, proteins, or lipids presented on the host cell surface.

In addition to attachment, adhesins also promote the delivery of toxins through the upregulation of virulence genes leading to invasion of the host. Furthermore, such interactions can trigger cytokine production or lectinophagocytosis by the host, compounding the severity of disease [21]. For this reason, subunit vaccines that hinder microbial attachment could be of great advantage in combating pasteurellosis in lump suckers. Examples of successful adhesin vaccines include those against enterotoxigenic *E. coli* (ETEC) strains in farm animals using the K88 fimbriae [22] as well as uropathogenic *E. coli* (UPEC) known to cause urinary tract infections (UTIs) using the FimCH complex [23]. For the aquaculture industry, vaccines have been tested by targeting major bacterial adhesins from *Aeromonas hydrophila* [24,25], *Vibrio harveyi* [26], and *Edwardsiella tarda* [27].

As research on *P. atlantica* in lump suckers is still in its infancy, no commercial vaccines are available against pasteurellosis in lump suckers. Simple phylogenetic analyses using the *rpoB* and 16S rRNA genes have shown that NVI-9100 is similar to other *Pasteurella* sp. isolated from lump suckers [4]. Whole genome analysis can, however, determine more accurate phylogenetic relationships. The purpose of this study is, therefore, to determine the taxonomic position of the species through whole genome sequencing and to predict potential vaccine antigens via *in silico* analysis. To further assess the predicted vaccine targets, their expression levels during exposure to lump sucker leucocytes were analyzed *in vitro*.

## 2. Materials and Methods

### 2.1. Bacterial Culture, Genome Isolation, and Sequencing

Two *P. atlantica* isolates described in previous studies (NVI 9100 and UiBP1-2013) [4,6,28] collected from clinically sick lumpsuckers were whole-genome sequenced. Briefly, bacteria were grown in tryptic soy broth (TSB) (Becton Dickinson, Sparks, MD, USA) supplemented with 1.5% NaCl and 10% fetal bovine serum, Australian origin, (Gibco, Waltham, MA, USA) at 20 °C with shaking (200 rpm).

Total genomic DNA was isolated using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) following manufacturer instructions. Briefly, 3 mL of an overnight culture of *P. atlantica* containing a maximum of  $2 \times 10^9$  cells was centrifuged at  $2500 \times g$  for 15 min and the pellet was resuspended in 180  $\mu$ L of ATL buffer. Twenty microliters of proteinase K were then added, and the bacteria were incubated at 56 °C in a rotating heat block for 1 h. Two hundred microliters of AL buffer and two hundred microliters of 96% ethanol were added, vortexing well between additions. The mixture was then loaded onto DNeasy Mini Spin columns and washed with the appropriate buffers (AW1 and AW2) prior to elution using buffer AE. The eluted DNA was then purified and measured for concentration before storing at  $-20$  °C until further processing.

Extracted DNA was sequenced using either an Illumina (UiBP1-2013) or PacBio (NVI 9100) platform. Illumina libraries were made using the Nextera DNA Flex Sample Prep kit (Illumina, San Diego, CA, USA) according to the manufacturer instructions and sequenced with Illumina MiSeq (Illumina, San Diego, CA, USA) using V3 chemistry. PacBio libraries were prepared with the Pacific Biosciences 20 kb library preparation protocol, and size was selected using BluePippin (Sage Science, Beverly, MA, USA), with subsequent sequencing performed on a Pacific Biosciences RS II instrument using P4-C2 chemistry and employing three SMRT cells (Pacific Biosciences, Menlo Park, CA, USA).

Raw Illumina sequences were adapter trimmed, quality filtered ( $Q > 20$ ), de novo assembled using [29,30]. Contigs shorter than 1000 bp or with  $<5$  times coverage were removed from each assembly prior to gene annotation. The genes of the Illumina sequenced genome were predicted by Glimmer 3, version 3.02b; Johns Hopkins University: USA, 1999 [31].

PacBio sequences were de novo assembled using HGAP version 3 (Pacific Biosciences, SMRT Analysis Software version 2.2.0) and circularized with the Minimus2 software version 2 (Amos package). For post-circularization correction of bases, reads were subsequently mapped to the circularized sequences using RS\_Resequencing.1 software (SMRT Analysis version version 2.3.0).

This whole genome project has been deposited at GenBank under the Accession Number PRJNA721926 (UiBP1-2013) and Accession Number CP074346 (NVI 9100).

### 2.2. Phylogenetic Analysis

Twenty-three *Pasteurellaceae* genomes (Table 1) were included for phylogenetic analyses.

In order to evaluate similarity among whole genomes, Orthologous Average Nucleotide Identity Tool using OrthoANI [43] was used to provide reliable and fast assessment of average nucleotide identity (ANI) for taxonomic classification purposes. A Similar Genome Finder with high similarity parameters (max hits: 50;  $p$ -value threshold: 1; distance: 0.5) utilizing Mash, a fast genome distance estimation tool [44], mounted on the Pathosystems Resource Integration Center (PATRIC) platform (<https://patricbrc.org/> (accessed on 3 June 2021)) [45,46] was also used to assess the genomic similarity of *P. atlantica* to the bacterial database.

PATRIC [45,46] was used to construct a whole-genome codon-based tree built on 500 single-copy genes [47] present in all genomes studied (Table 1). Both amino acid and nucleotide sequences were aligned using MUSCLE [48] and the codon align function of Biopython [49], respectively. The concatenated alignments of protein and nucleotide sequences were used in order to generate a RAxML method-based phylogenetic tree with

branch support values determined by 100 rounds of rapid bootstrapping [50,51]. The tree was visualized using iTOL online platform [52].

**Table 1.** Twenty-three *Pasteurellaceae* genomes used to evaluate the phylogenetic position of *P. atlantica*.

Bacterial Species Name (GenBank)	Animal Host	Country	Genome	Ref.
<i>Pasteurella</i> strains				
<i>Pasteurella multocida</i> strain ATCC43137	Pig	n/a	Assembly	NCBI
<i>Pasteurella dagmatis</i> strain NCTC11617	Human	n/a	Assembly	[32]
<i>Pasteurella testudinis</i> strain DSM23072	Desert tortoise	USA	WGS—69 contigs	[33]
<i>Pasteurella bettyae</i> strain CCUG2042	Human	USA	WGS—47 contigs	NCBI
<i>Pasteurella langaensis</i> strain DSM22999 Ga0215636	Chicken	Denmark	WGS—29 contigs	[32]
<i>Pasteurella caecimuris</i> strain NM44_TS2-9	Mouse	Canada	WGS—53 contigs	NCBI
<i>Pasteurella canis</i> strain NCTC11621	Human	n/a	WGS—20 contigs	[32]
<i>Pasteurella oralis</i> strain WCHPO000540	Human	China	WGS—30 contigs	[34]
<i>Pasteurella skyensis</i> strain DSM24204	Atlantic salmon	Scotland	WGS—56 contigs	[35]
[ <i>Pasteurella</i> ] <i>aerogenes</i> strain NCTC13378	n/a	Finland	Assembly	NCBI
<i>Pasteurella pneumotropica</i> strain DSM21403	Mouse	USA	WGS—22 contigs	NCBI
Non- <i>Pasteurella</i> strains				
<i>Phocoenobacter uteri</i> strain NCTC12872	Harbor porpoise	Scotland	WGS—5 contigs	[36]
<i>Hemophilus hemolyticus</i> strain 65117 B Hi-3	Human	Australia	WGS—16 contigs	NCBI
<i>Hemophilus hemolyticus</i> strain M19346	Human	USA	Assembly	[37]
<i>Hemophilus influenzae</i> strain 156_HINF	Human	USA	WGS—41 contigs	[38]
<i>Hemophilus influenzae</i> strain NCTC8143	Human	UK	Assembly	NCBI
<i>Hemophilus parainfluenzae</i> strain 155_HPAR	Human	USA	WGS—61 contigs	[38]
<i>Hemophilus parainfluenzae</i> strain T3T1	n/a	n/a	Assembly	NCBI
<i>Aggregatibacter aphrophilus</i> strain C2008001782	Human	USA	WGS—16 contigs	[39]
<i>Aggregatibacter aphrophilus</i> ATCC33389 strain NCTC5906	n/a	UK	Assembly	NCBI
<i>Glaesserella parasuis</i> strain F9	Pig	Spain	WGS—182 contigs	[40]
<i>Actinobacillus indolicus</i> strain AIFJ1607	Pig	China	Assembly	[41]
<i>Mannheimia hemolytica</i> strain M42548	Ruminants	n/a	Assembly	[42]

### 2.3. Analysis of Genomic Regions and Reverse Vaccinology Approach

The novel genomes were screened against major virulence factor databases. Specifically, the proteome of *P. atlantica* was blasted against VFDB [53] and VICTORS [54] using an e-value of  $1 \times 10^{-10}$ . A moderate e-value was applied to avoid false positive hits while identifying as many presumptive virulence factors as possible. The prediction was complemented by screening against PATRIC\_VF [55], the integrated virulence factor database inside the PATRIC platform (<https://patricbrc.org/> (accessed on 3 June 2021)) [45,46]. Subcellular localization for all presumptive virulence factors was examined by PSORTb 3.0 [56] to evaluate putative interaction with the extracellular environment (localization score threshold  $>7.5$ ). Proteins with both extracellular and outer membrane subcellular localization were imported to SignalP server [57] in order to further assess any known types of signal peptides as well as the pathways in which they are involved during their secretion or anchoring process (threshold  $>90\%$ ). The detection of possible adhesion-related functionality was tested in SPAAN [58]. A Protein Fold Recognition Server, Phyre2 [59], was applied to obtain insights on structure and folding properties of presumptive virulence factors using confidence level of above 97%.

Clusters of genes of probable horizontal origin, known as genomic islands (GIs), were identified by IslandViewer 4, version 4; Simon Fraser University: Canada, 2017 [60]. Inducible prophages were detected using the prophage finder tool, PHASTER [61], while VIRFAM [62] was used to classify intact prophages.

In line with reverse vaccinology principles, the identification of the strongest virulence candidates was deduced through Dynamic Vaxign Analysis, Vaxign [63], a software platform that has been developed and dedicated to vaccine design. The pipeline of Vaxign has integrated PSORTb [56] and SPAAN [58] so it can also be used to validate the previously identified virulence factors. Vaxign-ML [20] was also included in the reverse vaccinology analysis as a machine learning model to improve the prediction of bacterial protective antigens. Vaxijen [64] complemented the analysis to corroborate that the finally selected candidates for vaccine development also can function as protective antigens, which is a prerequisite for vaccine development.

## 2.4. Functional Studies

### 2.4.1. Processing Bacteria for qPCR Analysis

Cultures of *P. atlantica* (UiBP1-2013) used for the in vitro exposure experiment (Section 2.4.5) were cultivated as described in Section 2.1, harvested in the late exponential growth phase (18 h post inoculation) and centrifuged at  $2500 \times g$  (Beckman Coulter Allegra X-15R) for 15 min at 4 °C. For the gene expression analysis during bacterial cell proliferation in growth medium, 1 mL samples were harvested 14, 16, 18, and 20 h after inoculation. Bacterial cell counts were measured at harvest using a cell counter (CASY Model TT (Innovatis) and CASY worX version 1.26) followed by centrifugation at  $4000 \times g$  (Beckman Coulter Allegra X-15R) for 10 min at 4 °C. For both methods, the growth medium was then discarded, and the samples were stored on ice.

Total RNA was immediately extracted using the Bacterial RNA Kit (E.Z.N.A., Norcross, GA, USA) according to manufacturer instructions. The RNA was then DNase-treated (Sigma-Aldrich, Saint-Louis, MO, USA), converted to cDNA using qScript cDNA Synthesis Kit (Quantabio, Beverly, MA, USA), and stored at  $-20$  °C.

### 2.4.2. qPCR

Each qPCR reaction contained a volume of 25  $\mu$ L and consisted of 12.5  $\mu$ L  $2 \times$  SYBR Green JumpStart Taq Ready Mix (Thermo-Fisher Scientific, Waltham, MA, USA), 1  $\mu$ L each of the forward and reverse primers (10  $\mu$ M final working concentration, Table 2), 0.5  $\mu$ L of RNase and DNase free water (Sigma-Aldrich), and 10  $\mu$ L of cDNA (concentration depended on the specific analysis). A C1000 Touch thermal cycler with CFX96 Real-Time System (Bio-Rad, Oslo, Norway) was used for qPCR, with the following cycle conditions: 94 °C for 5 min followed by 40 cycles of 94 °C for 15 s and 60 °C for 1 min. Melting curve analyses were performed after each run (60 to 92 °C at a rate of 1 °C/5 s) to ensure that the specificity of the primers and the qPCR products were visualized on a 2% agarose gel. Three parallel reactions were performed for each gene, and negative controls excluding cDNA (NTC) and cDNA reactions without reverse transcriptase (NRT) were included for all master mixes. The gene expression levels were calculated by the  $\Delta\Delta C_t$  method [65].

**Table 2.** Details of primers for reference genes and the target gene used for qPCR.

Gene	Primer Name	Sequence 5'-3'	Primer Length (bp)
<i>gyrA</i>	#27-B_GYRA_F3	GTTTCATCGGGTATTGCGGTCGGTAT	25
	#28-B_GYRA_R3	TCCTGTGCGGTAAGCGTCTTCG	22
<i>rpoD</i>	#46-B_RPOD_F1	GGACGTGATGCGACACCTGAAGAAT	25
	#47-B_RPOD_R1	AGTGGCTGTGCAAGTGCAGTATCTT	25
Putative uncharacterized protein <Hia>	#70-B_HIA_F4	AGGTGTGGGTTTCATTCGCTGTGG	23
	#68-B_HIA_R2	CCGATTGCTGCCGCTTGTGTTTC	23

### 2.4.3. Primer Design and Validation

Genes considered for this work were based on the results from an analysis of genomic regions and potential vaccine targets. The genes selected as reference genes were *rpoD* and *gyrA*. The target gene selected was the putative uncharacterized protein (<Hia>) (Table 2).

qPCR assays were designed using the software Primer Premier version 6.24 (Premier Biosoft, San Francisco, CA, USA). The five highest rated assays for each target sequence were then chosen for testing. The length of the amplicons was kept between 100 and 250 bp for optimal amplification efficiency. The specificity of the primers was confirmed by qPCR (20 ng of cDNA used in each qPCR reaction), and product size was observed by electrophoresis on 2% agarose gels. All of the qPCR assays produced single amplification products. The best assay for each target gene based on Cq value, non-template controls, melting curves, and the results of electrophoresis were then chosen for further work. The resulting primers used for qPCR are listed in Table 2.

### 2.4.4. Bacterial Exponential Growth Phase Analysis

*P. atlantica* was grown as described in Section 2.1. At different time points after inoculation, the bacteria were harvested, the RNA was isolated, and cDNA synthesis was performed as described in Section 2.4.1. The synthesized cDNA was diluted across a twofold dilution series to give a range from 10 ng/ $\mu$ L to 0.625 ng/ $\mu$ L of cDNA for both the target and reference genes in qPCR (100 ng to 6.25 ng in each qPCR reaction). The relative gene expression was calculated by the  $\Delta\Delta Cq$  method using the most suitable dilution from the range tested, and comparisons of gene expression were made to the lowest stable time point of the analysis.

### 2.4.5. Head Kidney Leucocyte (HKL) Exposure Analysis

Four lumpsuckers were quickly netted and killed by a sharp blow to the head. Leucocytes were isolated from the head kidney on discontinuous Percoll gradients as described previously [66] with the following modifications. The supplemented L-15 medium did not contain gentamicin sulphate since the cells were to be exposed to viable *P. atlantica*. Additionally, resuspension of the isolated leucocytes was performed in L-15 supplemented with 5% fetal calf serum (L-15/FCS). The leucocytes were counted in a CASY-TT Cell Counter TM (Innovatis AG), and the viability (95%) and aggregation factor (1.2) of the cells were determined. The concentration of the isolated leucocytes was then adjusted to  $3.3 \times 10^6$  cells  $mL^{-1}$  in L-15/FCS, and 250  $\mu$ L volumes were added to each well of the 24-well Nunc plates (approximately  $8 \times 10^5$  cells per well) (Thermo-Fisher Scientific) and incubated overnight at 15 °C prior to exposure to *P. atlantica*.

A late exponential phase (18 h) culture of *P. atlantica* was prepared as described in Section 2.4.1, re-suspended in L-15/FCS, and adjusted to  $1.5 \times 10^8$  bacteria  $mL^{-1}$ . Volumes of 250  $\mu$ L were then added to each well (approximately  $4 \times 10^7$  bacteria per well). Sterile L-15/FCS medium was supplied to the leucocytes and used for the non-challenged controls. The cells were then incubated at 15 °C, and samples were collected at 3, 6, 9, 12, and 24 h after exposure. Samples were collected from the wells and centrifuged to remove all media prior to storage at  $-80$  °C, until the total RNA was isolated, DNase treated, and converted to cDNA as described above. For qPCR, 80 ng per reaction was used as the template for both the target and reference genes.

### 2.4.6. Statistical Analysis

The relative gene expression during the HKL exposure was calculated by the  $\Delta\Delta Cq$  method [65] and comparisons were made to the negative controls (bacteria without leucocytes). The results were analyzed using two-way ANOVA and Fisher's LSD for post hoc tests, and differences were considered significant when  $p < 0.05$ . All statistical analyses were carried out using SigmaPlot version 12 (Systat Software, San Jose, CA, USA).

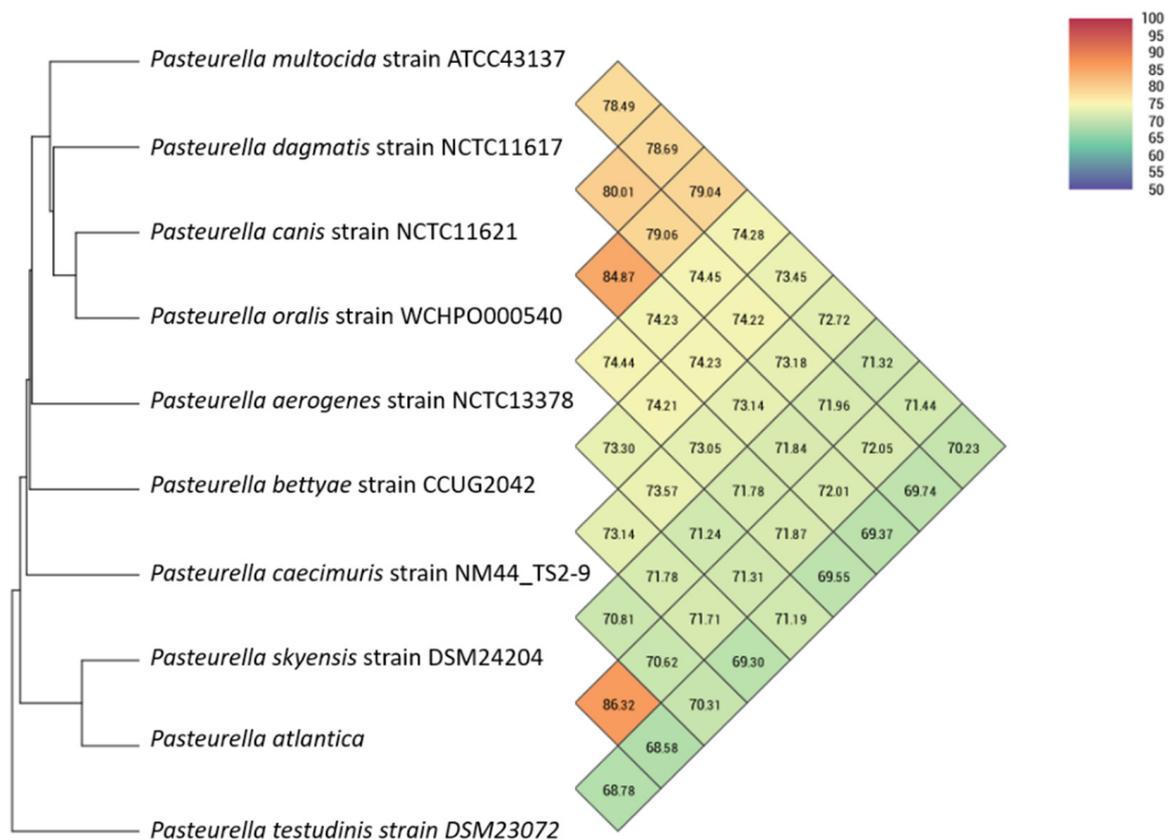
### 3. Results

#### 3.1. Novel Sequenced Genome and Accession Numbers

The length of the PacBio sequenced *P. atlantica* genome (NVI 9100) is 2,301,649 bp, with a GC content of 42.1%, and 2197 coding regions (CDS) were predicted. The length of the Illumina sequenced *P. atlantica* genome (UiBP1-2013) is 2,260,408 bp.

#### 3.2. Phylogenomics and Taxonomic Classification

The *P. atlantica* strains were isolated from *Cyclopterus lumpus* in a previous study where preliminary phylogenetic analysis was performed based on 16S rRNA and *rpoB* gene sequencing [4]. To assess pairwise similarity percentages among *Pasteurella* species, nine different publicly available *Pasteurella* species genomes (Table 1) and *P. atlantica* were analyzed using the OrthoANI tool (Figure 1).



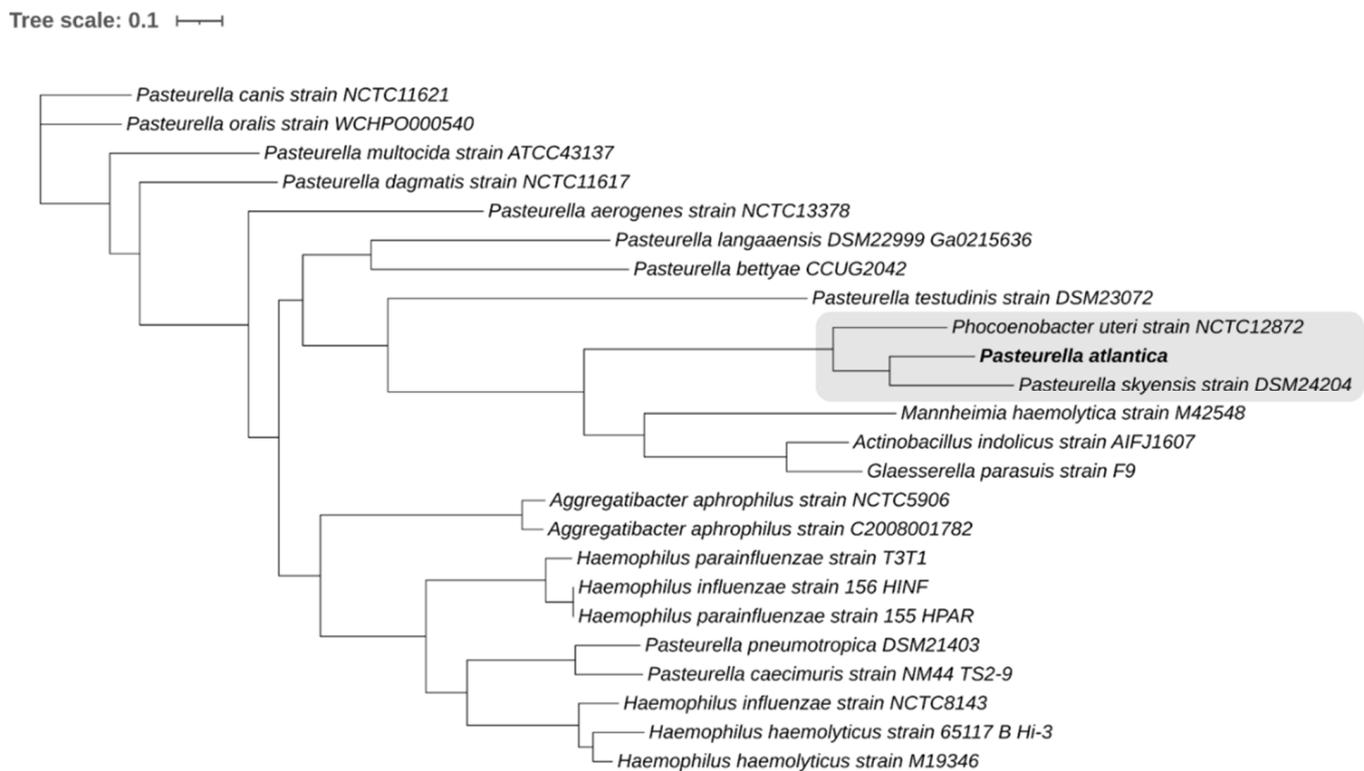
**Figure 1.** Heatmap generated with OrthoANI values calculated from the Orthologous Average Nucleotide Identity Tool (OAT) software.

When compared to our genome of interest, *P. skyensis* shows the highest OrthoANI values (86.32%) whereas the other species generated much lower values (70 to 72%).

Apart from *Pasteurella* genus, a genomic analysis of *P. atlantica* against the entire publicly available bacterial database, based on whole genome sequences, according to Similar Genome Finder, returned the *P. skyensis* strain DSM 24204 as the only similar genome at a genomic distance of 0.096 (Mash distance threshold: 0.1). The second most similar genome at a genomic distance of 0.145 was the *Ph. uteri* strain NCTC12872, and that was shown only when the threshold increased up to 0.5 (Figure 1). A bigger number of *Pasteurellaceae* family bacteria were found to be less similar, at genomic distances of >0.2, and for the sake of further analysis, representative genomes have been included in Table 1 as well. Substantially low-similarity levels compared to the bacterial database constitute an additional piece of evidence supporting *P. atlantica* as a novel species, knowing that a

pairwise Mash distance of at least  $\leq 0.05$ , approximately corresponding to an ANI of  $\geq 95\%$ , is taken for granted among conspecific bacterial genomes [44].

Multiple codons through the entire genomes of Table 1 were used as genetic markers to finalize the taxonomy of *P. atlantica* (Figure 2).



**Figure 2.** Phylogenetic tree of *P. atlantica* and Table 1 bacterial strains based on compariScheme 500 of randomly selected codons through their whole genome sequences. The monophyletic group in which the novel species belongs is highlighted, while the branch support in all nodes is  $>98\%$  and has been based on 100 rounds of rapid bootstrapping.

According to the genome-wide analysis, *P. atlantica* clusters together with *P. skyensis* and *Ph. uteri*, forming a robust phylogenetic group supported by 100% confidence level (highlighted in Figure 2). Both monophyletic taxa are phylogenetically distinct since their nodes bear 100% support values.

### 3.3. Genomic Elements

The sequence of *P. atlantica* revealed mobile elements including genomic islands and prophages that should further be analyzed due to their possible role in pathogenicity [67–69].

#### 3.3.1. Genomic Islands and Prophages

Bioinformatic analysis on the genome of *P. atlantica* illustrated eleven genomic islands (GIs) (Table 3) with various properties.

GC content of the GIs span between 29.4 and 36.6%. They encompass 207 coding sequences (CDS), approximately 10% of the *P. atlantica* total proteome, and many of them were attributed to virulence-related or broadly proteolytic activity functions. The most prominent putative virulence functions belong to CDS such as RTX-like toxin (GI\_1), putative hemolysin toxin of *HlyC* family (GI\_1), heme-binding protein (GI\_4), bacterial peptidoglycan degrading enzyme of NlpC/P60 (GI\_6), spermidine-putrescine binding (GI\_8), acetyltransferase of gnat family (GI\_9), ctx rstr-like repressor (GI\_9), two-partner secretion (TPS) system protein (GI\_10), and toxin-antitoxin system proteins (GI\_11).

The genome of *P. atlantica* (NVI 9100) harbors five different prophages that were in silico predicted to be intact, hence inducible (Table 4).

**Table 3.** Genomic features of the eleven genomic islands that were found in the genome of *P. atlantica* (NVI 9100).

#	Sequence Location	Sequence Length (bp)	GC%	CDS
GI_1	GI_64353_82229	17,877	31.3	19
GI_2	GI_322423_340166	17,744	34.8	26
GI_3	GI_685873_700034	14,162	34.4	11
GI_4	GI_922164_930860	8697	36.6	12
GI_5	GI_966979_985107	18,129	34.7	26
GI_6	GI_1040130_1045358	5229	35.2	7
GI_7	GI_1058606_1061839	3234	33	9
GI_8	GI_1287395_1293893	6499	29.4	7
GI_9	GI_1369488_1396082	26,595	31.1	34
GI_10	GI_1468334_1477249	8916	29.7	12
GI_11	GI_1534357_1558731	24,375	35	44

**Table 4.** Genomic features of the five prophages that were found in the genome of *P. atlantica* (NVI 9100).

#	Sequence Location	Sequence Length (bp)	GC%	CDS	Most Closely Related Phage	Predicted Family	NCBI Accession Number
P_1	P_322716_342002	19,287	35	26	<i>Enterobacteria</i> phage P4	Unknown	NC_001609
P_2	P_972452_1005548	33,097	37.3	40	<i>Mannheimia</i> phage vB_MhS_535AP2	Siphoviridae	NC_028853
P_3	P_1030690_1066158	35,479	35.6	49	<i>Escherichia</i> phage D108	Siphoviridae	NC_013594
P_4	P_1527862_1576157	48,296	35.2	57	<i>Mannheimia</i> phage vB_MhM_587AP1	Myoviridae	NC_028898
P_5	P_2056529_2078637	22,109	35	28	<i>Hemophilus</i> phage SuMu	Myoviridae	NC_019455

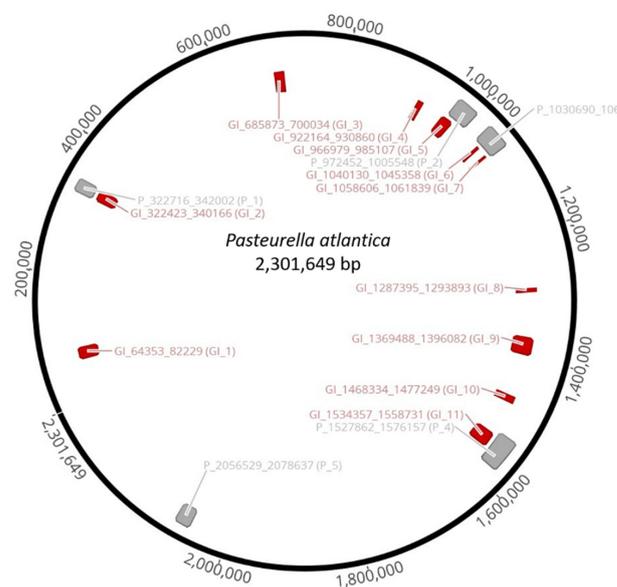
P\_1 and P\_3 were found to be somewhat related to *Enterobacteria* phages, whereas P\_2, P\_4, and P\_5 resemble prophages that have already been described in other *Pasteurellaceae* family bacteria [70,71]. According to BLAST sequence similarity, it was only possible to match 4 out of 26 CDS (7.7%) present in P\_1 with any known phage genome implying that this is a quite novel prophage, uniquely present in the genome of *P. atlantica*. Similarly, P\_3 seems to be a unique prophage as only 9 of its 49 CDS (18.4%) aligned with *Escherichia* phage D108. On the other hand, P\_2 carried 9 of 40 CDS (22.5%) of *Mannheimia* phage vB\_MhS\_535AP2 while P\_4 has homologs of 24 out of 57 CDS (42.1%) that were present in *Mannheimia* phage vB\_MhM\_587AP1. Last, P\_5 has 23 out of 28 CDS (82.1%), resembling *Hemophilus* phage SuMu more, which was also anticipated due to the close phylogenetic relationships between *P. atlantica* and *G. parasuis* [40].

Interestingly, some of the predicted genomic islands and prophages do overlap in the genome (Figure 3). Specifically, the entire GI\_2 is part of P\_1, part of GI\_5 is included in P\_2, both GI\_6 and GI\_7 belong to P\_3, while GI\_11 is a part of P\_5. Hence, there is a prominent connection between prophages and genomic islands that may be directly linked to the virulence properties of the bacterial host, with the former inciting the mobilization of the latter.

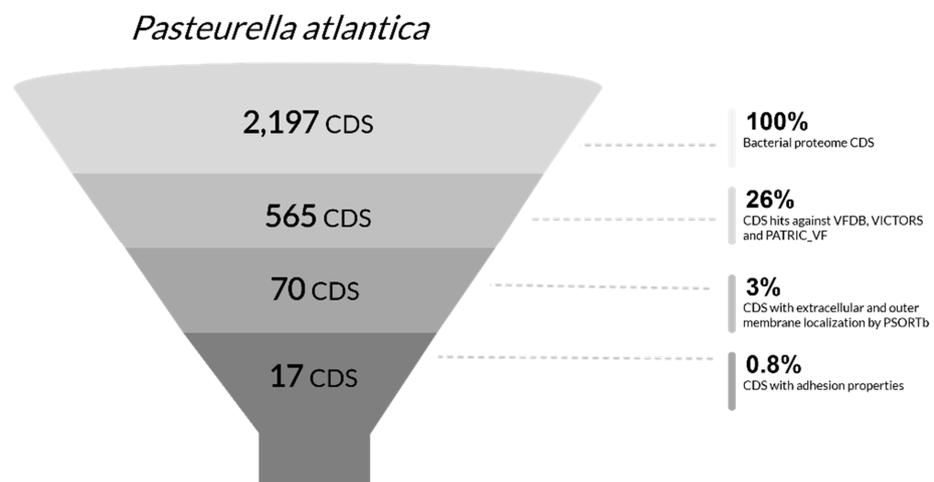
### 3.3.2. Virulence Factors

The proteome of *P. atlantica* consists of 2197 protein coding sequences (Supplementary Material S1), which were screened against an integrated matrix consisting of the VFDB, VICTORS, and PATRIC\_VF databases to generate 565 unique amino acid sequence hits after discarding overlapping hits (Supplementary Material S1). This corresponds to ap-

proximately 26% of the 2197 predicted coding sequences. However, significant uncertainty needs to be removed from the analysis to achieve a better focus on the essential virulence factors present in the species. Subcellular localization of presumptive virulence factors showed that approximately 12% (70 out of 565) are presumably either extracellularly secreted (21) or with outer membrane surface exposure (49), setting these proteins as primary candidates for the pathogenicity of *P. atlantica*. The amino acid sequences of the predicted proteins accompanied by their most similar protein sequence GI number (SCL-BLAST), subcellular localization, adhesion, and transmembrane (TM) helices probabilities are reported in Supplementary Material S1. According to their amino acid sequences, 17 out of the 70 virulence factors with extracellular and outer membrane subcellular localization were the most critical for mediating bacterial invasion, such as adhesins with probabilities higher than 75%. Adhesins correspond to 0.8% of the total proteome (Figure 4), and since their expression may be a key determinant of the success of the infection process [72], they could play an imperative role in the pathogenicity of *P. atlantica* (Table 5).



**Figure 3.** Genomic map of *P. atlantica* that highlights the eleven genomic islands (red) and five prophages (grey). Five out of eleven GIs are shown to overlap with four out of five prophages.



**Figure 4.** Filtration process of the CDS present in the genome of *P. atlantica* to determine the most likely virulence factors. According to the selection criteria applied during each step of the process, a total of 17 putative virulence factors were predicted as potentially important contributors to the pathogenicity of *P. atlantica*.

**Table 5.** Seven extracellular and ten outer membrane localized putative virulence factors of the genome of *P. atlantica* (NVI 9100). E: extracellular, OM: outer membrane, GI: Genomic island, P: Prophage, and ND: Not designated.

#	CDS	Localization (%)	Adhesin (%)	Signal Peptides	Protein Data Bank (PDB) Header (PDB Molecule)	Designated Genomic Area
1	Putative uncharacterized protein <Hia>	E (96)	92.5	n/a	Membrane protein/cell adhesion (Hia)	ND
2	Uncharacterized protein	E (96)	74.8	n/a	De novo protein (designed helical bundle)	GI_6 and P_3
3	Hemolysin-type calcium-binding region	E (96)	82	n/a	Cell adhesion (surface associated protein csha)	ND
4	Putative collagen triple helix repeat protein	E (96)	86.2	n/a	Membrane protein/cell adhesion (Hia)	GI_3
5	HbP1 protein	E (97)	81.5	n/a	Structural protein/contractile protein (collagen I alpha 1)	GI_2 and P_1
6	Uncharacterized protein	E (97)	76	Sec/SPI	Metal transport (hemophore)	ND
7	Pilus A	E (96)	82.9	n/a	Cell adhesion (fimbrial protein)	ND
8	Hep/Hag repeat protein	OM (95)	85.8	n/a	Cell adhesion (hep_hag family)	ND
9	Uncharacterized protein	OM (89)	82.6	n/a	Toxin (hemolysin)	ND
10	Protein PfhB1	OM (99)	75.4	n/a	Membrane protein (znud)	ND
11	Putative uncharacterized protein 4	OM (100)	74.7	n/a	Toxin (hemolysin)	ND
12	Putative septum site-determining protein MinC	OM (99)	82	n/a	Pili subunits (pili subunits)	ND
13	Putative uncharacterized protein 26	OM (95)	86.3	Sec/SPI	Pili subunits (pili subunits)	P_5
14	Phage-related protein tail component-like protein	OM (95)	74.6	n/a	Signaling protein receptor (interferon alpha beta receptor 1)	P_2
15	Type IV pilus biogenesis/stability protein PilW	OM (99)	80	n/a	Transferase (udp-n-acetylglucosamine-peptide n-)	ND
16	Auto transporter beta-domain protein	OM (95)	84.8	n/a	Hydrolase (esterase esta)	ND
17	Outer membrane antigenic lipoprotein B	OM (99)	79.7	n/a	Sugar binding protein (chitin elicitor-binding protein)	ND

The strongest virulence factor candidate is a novel uncharacterized protein, the function of which would have passed unnoticed without examining its protein structure. Therefore, PDB predictions based on protein fold recognition were included. All PDB predictions reported in Table 5 were above the 97% confidence level and encompassed major virulence components such as hemolysins, toxins, and pili subunits. Specific secretory signal peptides were detected in only 2 out of 17 adhesins, intended to be transported by Sec translocon followed by Signal Peptidase I cleavage (Sec/SPI) to be released to the host.

### 3.3.3. Vaccine Targets

All 70 CDS with extracellular and outer membrane localization were assessed as potential subunit vaccine candidates. However, apart from adhesion probability, additional factors such as transmembrane helices and orthologue analysis need to be considered for calculating their Vaxign-ML score, which was the metric applied to define their suitability as targets for vaccine development. All vaccine targets with Vaxign-ML scores higher than 99.5% are presented in Table 6, and they are therefore suggested as the most promising vaccine targets. All of the exact scores as well as further details about the remaining 60 extracellular and outer membrane CDS are included in Supplementary Material S1.

**Table 6.** Three extracellular and seven outer membrane localized proteins are the most promising candidates for vaccine development against *P. atlantica* (NVI 9100 and UiBP1-2013). E: extracellular, OM: outer membrane, GI: Genomic island, P: Prophage, and ND: Not designated.

#	CDS	Localization (%)	Adhesin (%)	Signal Peptides	PDB Header (PDB Molecule)
1	Putative uncharacterized protein <Hia>	E (96)	92.5	n/a	Membrane protein/cell adhesion (Hia)
2	Uncharacterized protein	E (96)	74.8	n/a	De novo protein (designed helical bundle)
3	Uncharacterized protein	OM (89)	82.6	n/a	Toxin (hemolysin)
4	Hep/Hag repeat protein	OM (95)	85.8	n/a	Cell adhesion (hep_hag family)
5	Protein PfhB1	OM (99)	75.4	n/a	Membrane protein (znud)
6	Putative uncharacterized protein 4	OM (100)	74.7	n/a	Toxin (hemolysin)
7	Serine protease sat autotransporter	E (84)	69.4	Sec/SPI	Hydrolase (hemoglobin protease)
8	PfhB1	OM (100)	69.5	n/a	Toxin (hemolysin)
9	Uncharacterized protein	OM (100)	63.3	Sec/SPI	Transport protein (translocation and assembly module tama)
10	Outer membrane protein assembly factor BamA	OM (100)	50.8	Sec/SPI	Membrane protein (outer membrane protein assembly factor BamA)

Since adhesion probability is not the only criterion for vaccine development, 6 out of the 10 CDS reported in Table 5 are potential virulence factors (Table 6). The list can further be confined by using VaxiJen as a tool to predict the protective antigenic properties of the vaccine targets. Only the putative uncharacterized protein and the Hep/Hag repeat protein had a score >80%, claiming that their antigenicity levels are high enough to trigger the host's immune response. However, the putative uncharacterized protein CDS not only was ranked first in both tables, scoring the highest Vaxign-ML (100%) and adhesin probability (92.5%), respectively, but also was characterized as a protective antigen, findings that render it the most significant candidate for virulence and potential vaccine development.

According to its amino acid sequence, BLASTP identified it as a YadA-like family protein, a collagen-binding adhesin originally characterized in *Yersinia enterocolitica* [73], that is also present in the genome of *P. skyensis*. However, although the query coverage was 100%, the identity rate was only 58.23%. Based on the predicted structure of the protein, Phyre2 [59] identified it as a membrane protein for cell adhesion similar to Hia (Tables 5 and 6), which is also a major adhesin expressed by *H. influenzae* [74]. Due to the prominent role that this uncharacterized membrane/cell adhesion protein seems to play in the virulence of *P. atlantica*, it was selected for further evaluation of its expression levels against lump sucker head kidney leucocytes.

### 3.3.4. Functional Analysis of the Major Virulence Factor

#### Expression of <hia> in Bacterial Exponential Growth Phase

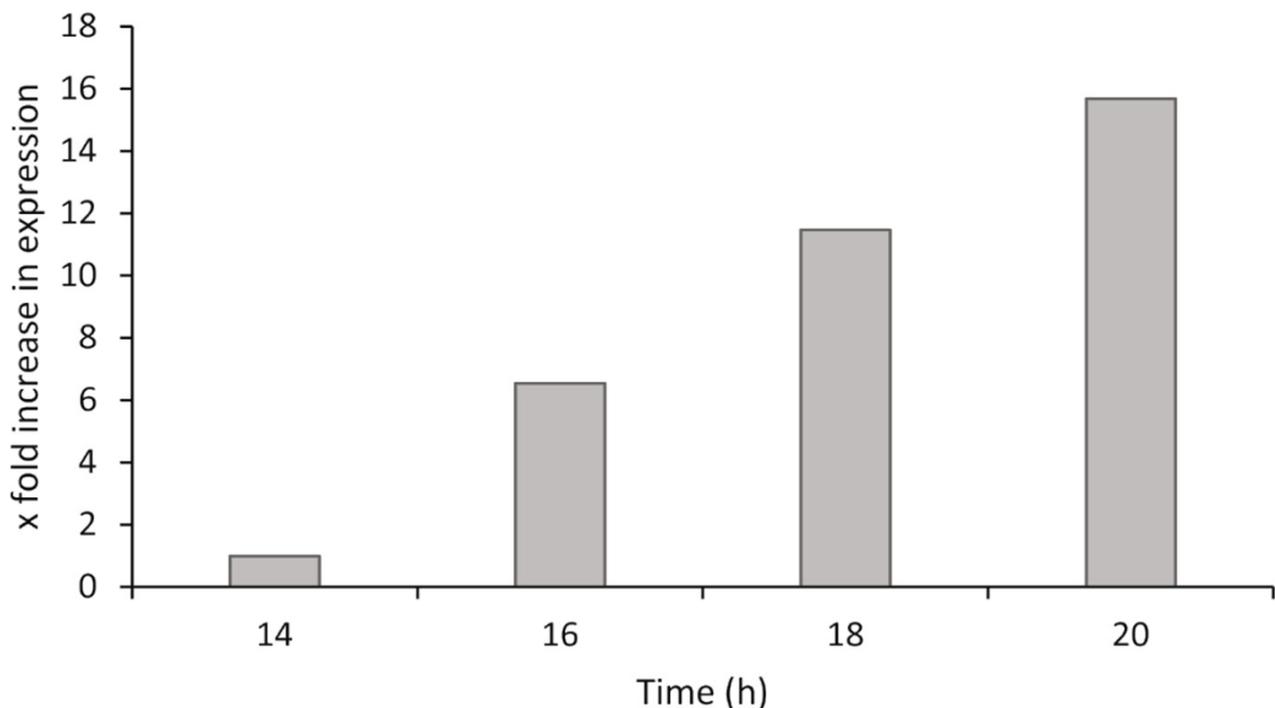
The predetermined time points for measurement were chosen based on growth curves of *P. atlantica* to coincide with the exponential growth phase of the bacteria. Furthermore, from the range of cDNA dilutions tested, 50 ng of cDNA per qPCR reaction yielded the clearest and most consistent results, and this dilution was used to calculate the <hia> expression. Out of the two reference genes tested, *gyrA* gave the most stable results compared to *rpoD* (Table 7) and was therefore used to calculate the relative expression of <hia> during cell growth and leucocyte exposure studies. This highlights the importance of

testing various reference genes for evaluation and stability for expression studies. In this study, although *rpoD* seemed to be a stable gene to use for analysis when primer assays were evaluated, experimental situations proved otherwise.

**Table 7.** Reference and target gene primers qPCR assay performance.

Gene	Amplicon Size (bp)	Assay Efficiency (%)	Correlation (R <sup>2</sup> )
<i>gyrA</i>	200	102	0.998
<i>rpoD</i>	173	104	0.987
Putative uncharacterized protein <Hia>	192	96	0.984

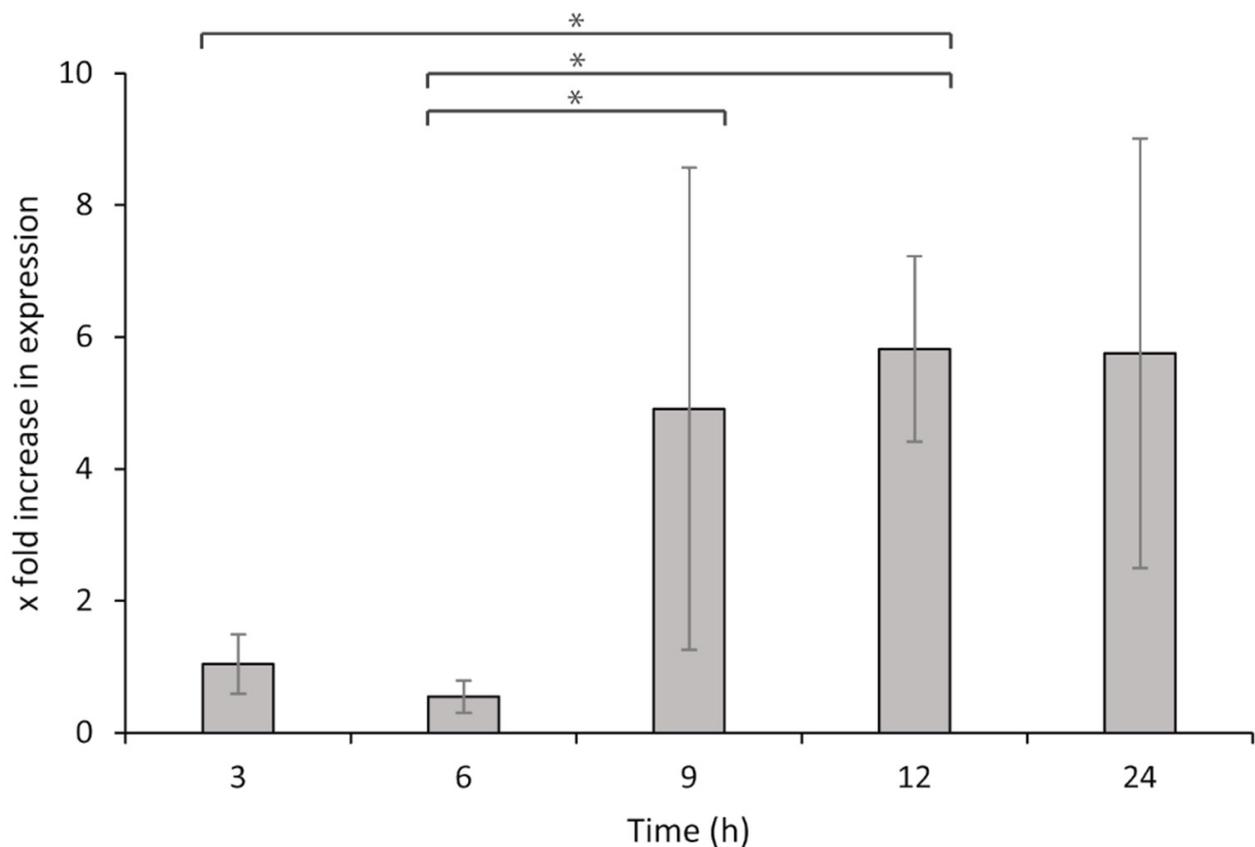
The expression level of <*hia*> changed over time, when analyzing samples collected at different time points during the exponential growth phase. The expression levels were upregulated by 6.5×, 11.5×, and 15.7× at 16, 18, and 20 h, respectively, compared to the expression level at 14 h (Figure 5). The expression levels also mirror the exponential increase seen in bacterial numbers from previous work performed on growth curves for *P. atlantica* [28].



**Figure 5.** Expression levels of <*hia*> in *P. atlantica* (UiBP1-2013) during growth in liquid medium. The time period corresponds to the exponential growth phase of the bacteria.

#### Head Kidney Leucocyte (HKL) Exposure Analysis

The expression levels of <*hia*> increased with time, with the highest levels recorded at 9 and 12 h after leucocyte exposure (Figure 6), with significantly elevated expression levels compared with those in the early stages of the exposure (3 and 6 h after exposure). At 24 h after exposure compared to 3 and 6 h after exposure, the bacteria would have entered the lag phase of growth, and the expressions of both the reference and target genes are therefore not comparable with those from earlier time points. Furthermore, statistical analysis also showed that HKL from individual fish did not have an impact on the expression of <*hia*> ( $p = 0.33$ ).



**Figure 6.** Fold increase in relative expression levels of  $\langle hia \rangle$  in *P. atlantica* (UiBP1-2013) over time in the presence of lumpsucker head kidney leucocytes, compared to bacteria not exposed to HKL. Error bars represent standard deviation ( $n = 4$ ). Asterisks represent significant differences in expression between the time points indicated by square brackets.

#### 4. Discussion

The isolates of *P. atlantica* used in this work were isolated and reported as the causative agent of confirmed outbreaks of pasteurellosis in farmed lumpsuckers from Norway.

##### 4.1. Phylogeny Reconstruction

Alarcón et al. [4] carried out preliminary phylogenetic analyses that included the isolates used in this study, using the 16S rRNA and *rpoB* genes, and it was suggested yet not concluded that the isolates might represent a novel species or subspecies. After examining the existing phylogeny of the *Pasteurellaceae* family, which is based on 16S rRNA genes, it can be deduced that using this molecular marker does not sufficiently resolve the true phylogenetic relationships, since it reflects rather incongruent phylogenetic relationships among its members. The available 16S rRNA based phylogenetic tree of the *Pasteurellaceae* family provides clear evidence that the polyphyletic nature prevails in the taxonomy of the family [75] and it has already been proven that there is room for improvement on molecular markers [76]. Despite ongoing research, the current phylogeny of the *Pasteurellaceae* family cannot be considered resolved according to the evolutionary principle of monophyletic taxa formation. Phylogenomic and molecular demarcation studies have been performed both on the major genera of the family such as *Pasteurella*, *Hemophilus*, and *Actinobacillus* [77] and on the reclassification of individual species such as the case of *Hemophilus parasuis* to *G. parasuis* [40].

Hence, the classification of *P. atlantica* in the current study was based on a dataset that includes representative species from both *Pasteurella* and other genera that belong to the *Pasteurellaceae* family to unravel a comprehensive taxonomy for *P. atlantica*. *P. multocida* is the most thoroughly studied member of *Pasteurellaceae* due to its high pathogenicity in a

broad range of livestock species [78,79]. The observed orthologous average nucleotide of 86.32% corroborates that *P. atlantica* and *P. skyensis* represent two different species. Whole genome-based phylogeny validated the novelty of the studied genome and its accurate taxonomic placement. Since *P. skyensis* and *Ph. uteri* are the only strains of marine origin out of the entire bacterial database, the habitat seems to be a strong differentiating factor for this taxon of *Pasteurellaceae*. Second, the host species of the bacteria in the marine cluster likely play a role for the speciation process. *P. atlantica* and *P. skyensis* are the closest relatives and both infect marine teleosts [35], whereas *Ph. uteri* isolated from a marine mammal is classified slightly more distant [36]. The major setback in this case is the lack of data in sequenced aquatic *Pasteurellaceae*, a fact that makes it quite precarious to formulate any solid conclusions regarding the trait of virulence. However, the highly emerging incidence and pathogenicity of *Pasteurella* in aquaculture has already brought [4,80] and will definitely bring more genomic data to the scientific community.

#### 4.2. Virulence Genes

The scarcity of aquatic *Pasteurella* genomes, particularly from the aquaculture environment, highlights the significance of these *P. atlantica* genomes in exploring genomic elements, including putative virulence factors, antimicrobial resistance genes, and their potential implications for devising precaution strategies such as vaccine development in aquaculture. The abundance of mobile elements such as GIs (11) and prophages (5) in *P. atlantica* further enforces its high pathogenicity that has already been documented in previous studies [4,6,28]. GIs and prophages are very often found in pathogenic bacteria and have consistently been reported as major contributors of virulence factors to their bacterial hosts. Five out of the seventeen strongest candidate virulence factors of *P. atlantica* are in GIs and/or prophages, which in total account for roughly 10% of the genome. This finding is in line with the current literature, which supports the potential of bacterial mobilomes to harbor significant pathogenicity traits. A recent comparative genomic analysis of *Vibrio anguillarum*, a highly virulent aquaculture pathogen, revealed that 64 GIs that belonged to the accessory genome of the species were present in six out of nine of the most virulent strains [67]. In the same study, the role of prophages was also quite prominent, while it was recently proven in another study that marine *Vibrio* harbors a large number of prophage-encoded virulence factors [81]. In the pangenome analysis of *P. multocida*, GIs were also characterized as one of the most important mobile components not only as sources of virulence factors but also as facilitators of foreign genes acquisition. Prophages were again highlighted as core elements of the mobilome, and it was speculated that they may serve as precursors for the formation of GIs [69]. Additionally, facultative intracellular pathogens, such as *P. atlantica* [6,28,82], usually demonstrate quite variable genomes with a fourfold larger genomic content of mobile DNA such as transposable elements and prophages, when compared to obligate intracellular bacteria [83].

Mobile elements are characterized by a high frequency of transfer in the environment through the mechanism of horizontal gene transfer (HGT). HGT can be phage-mediated and can facilitate the dissemination of virulence-associated GIs within and between microbial communities [84], and the presence of a pathogenicity island can potentially define virulence in bacterial pathogens [85]. Apart from sophisticated algorithms [60,61], simpler methods that are based on the distinct GC content of horizontally transferred genes compared to that of the bacterial chromosome have been suggested for the prediction of GIs and prophages [86]. For instance, the pathogenicity islands in *Salmonella typhimurium* (SPIs) have lower GC content than the rest of the bacterial genome [87]. According to the in silico analysis of our study, the average GC content for GIs and prophages were 34.5% and 35.6%, respectively, in both cases lower than 42.1%, which is the average GC content of the *P. atlantica* genome.

Recent progress in in silico analysis tools allowed us to recognize and assess potential virulence factors by exploring the novel genome of *P. atlantica*. Such methodologies constitute fundamental principles of RV, which facilitate vaccine development, while turning

extensive, time consuming, and generalized lab work into target-oriented functional analysis and in vitro/in vivo trials. It was not until 2006 when the first standalone RV program designated as NERVE [88] was published, whereas great progress as well as comparative analysis have been achieved and performed in the RV field ever since [89]. The cornerstone of virulence factor prediction has been the determination of subcellular localization of bacterial proteins. Proteins that are extracellularly secreted or exposed on the bacterial surface have good chances to play a role in bacterial virulence.

#### 4.3. Adhesins

Adhesins, for instance, which can facilitate the colonization of pathogens to mucosal or other biotic surfaces and which are essential for bacterial pathogenesis and survival may be prominent virulence factors and suitable candidates for vaccine development [72,90]. Screening of the *P. atlantica* genome for gene targets that can be further functionally evaluated as virulence factors generated the 17 strongest candidate genes that may have implications for bacterial virulence, based on subcellular localization and predicted adhesion properties.

Previously published bath, cohabitation, and intramuscular and intraperitoneal challenge trials that included the current *P. atlantica* isolate (UiBP1-2013) illustrated high and acute mortality rates, potentially indicating that extracellular proteins may have a substantial role in the infection dynamics [28], an interim conclusion that is supported by present findings. In our analysis, a putative uncharacterized protein ranked first as both a putative virulence factor and a vaccine candidate.

According to the in silico analysis, the role of this protein was directly linked to major adhesins YadA and Hia in *Yersinia* and *H. influenzae*, respectively, both of which are essential for successful infection of their hosts via human epithelia tissues [73,74,91]. Although their amino acid sequences were not identical, there was a clear connection based on query coverage and confidence level, which in both cases were 100%.

Trimeric autotransporters (TAAs) are a family of non-fimbrial, homotrimeric bacterial adhesins that are secreted by the Type 5c secretion system (T5cSS) and for which the main function is the attachment of pathogenic Gram-negative bacteria to hosts and abiotic surfaces. *Yersinia* adhesin A (YadA) is the most extensively studied member of this family, while others include *H. influenzae* adhesin (Hia) and *Neisseria* adhesin A (NadA).

The type 5c secretion system is of great importance in the development of multivalent recombinant and subunit vaccines, as it enables proteins to be exposed on the bacterial outer membrane, making them potential vaccine targets [92]. This could then also be used as the basis for further work to elucidate the structure of the Hia-like protein investigated in the current study.

YadA showcases the invasive nature of *Y. enterocolitica* and *Y. pseudotuberculosis*, which colonize the intestine by adhering to the intestinal mucosa [93]. Immunologically, YadA has been shown to be a highly immunogenic antigen. In fact, Tahir and Skurnik (2001) [93] found that immunizing mice with live bacteria raised antibodies against only epitopes located in the N-termini, whereas using inactivated bacteria raised antibodies against other epitopes as well. In addition, a study [94] found that mice immunized using a recombinant YadA vaccine had higher survival compared to those vaccinated with inactivated *Y. pseudotuberculosis*.

Hia is an adhesin found in 25% of clinical non-typable strains of *H. influenzae* (NTHi). It mediates adherence to host cells and causes various diseases in humans, including respiratory tract infections and meningitis [95]. Hia differs slightly from other TAAs and is thought to represent a new sub-family. It too is highly immunogenic, however, and antibodies were strongly induced following natural infection observed in humans. Furthermore, opsonophagocytic antibodies were generated in serum in guinea pigs and mice [96]. However, due to its prevalence in so few strains and despite its high antigenicity, it cannot be considered a vaccine target on its own but in combination with High Molecular Weight adhesin 1 (HMW1), High Molecular Weight adhesin 2 (HMW2), and outer membrane

vesicles (OMVs). In this way, a vaccine containing the three adhesins covers 95% of NTHi and confers better protection [97].

The significantly higher expression levels of *<hia>* detected in the current study following in vitro exposure to lump sucker leucocytes after 9 and 12 h together with the in silico results strengthen its candidacy as both a major virulence factor and a potential target for subunit protein antigen, mRNA, or DNA vaccine development.

The expression of *<hia>* in the absence of a host correlates with previously observed increases in bacterial growth rates [28] and indicates that the adhesin is typically expressed during bacterial growth. In the exposure experiment, bacteria were harvested from the growth medium at 18 h after inoculation when both bacterial growth and *<hia>* expression were at peak levels and combined with HKL in cell culture medium, which does not favor bacterial growth. The results from this experiment then demonstrate that the expression of *<hia>* increases significantly in the presence of the host, following an initial period of acclimation (up to 6 h, as seen in Figure 6). This implies that the presence of host leucocytes serves as a trigger for the bacteria to significantly increase the expression of adhesins (between 6 and 9 h after exposure) and shows that the increased expression is due to the presence of the host and not an increase in bacterial cell numbers. Furthermore, it indicates bacterial cell attachment to leucocytes as an initial step towards infection, as previously suggested by Ellul et al. (2019) [6].

#### 4.4. Vaccine Development

Vaccines targeting adhesins can confer protection by two major mechanisms, one of which is by inducing neutralizing antibodies, which attach to adhesins and prevent attachment and subsequent infection. The other involves inducing opsonizing antibodies, which attach to adhesins and increase recognition of the pathogen by Fc receptors on phagocytes, leading to phagocytosis, or by activation of the complement system [98], as shown by Winter and Barenkamp [96] in their experiments with Hia.

Ellul et al. [6] recently assessed the efficacy of formalin-inactivated *P. atlantica* vaccine in lump sucker, and although immune response and antibody production were prominent, the protection levels provided by the vaccine were relatively weak. Inactivated bacterin is the most common and traditional type of vaccine used in aquaculture; however, alternatives such as subunit recombinant, mRNA, or DNA vaccines may be the next step in solving the problem [99]. This is especially the case with the recent advances in human mRNA vaccines against SARS-CoV-2, which point to potential avenues for fish vaccines. These vaccines can be developed using various techniques, including harvesting and purifying the target directly from the pathogen, or using recombinant expression vectors, such as *E. coli* [99]. Pending further work, the former method could be a possible step forward in the case of *P. atlantica* as high expression rates were observed in bacteria in their exponential growth stage.

Some of the most successful recombinant protein subunit vaccines in the aquaculture industry include those against IPN (MSD Animal Health) and ISA in Chile (Centrovet). DNA vaccines have also been produced for use in Atlantic salmon against IHN in Canada and against SPD in Norway (Elanco). Subunit vaccines against the bacterial fish disease furunculosis in *Oncorhynchus mykiss* [100] have been tested. All targets for the vaccine development were identified in silico, employing the RV principles and found to be protective. Recombinant protein production has already been used in order to protect farmed fish against infections by *V. anguillarum*, *Edwardsiella tarda*, *Aeromonas hydrophila*, and *Francisella orientalis* [101] in all cases leading to the suggestion of potential vaccine candidates [102–104]. Following these lines, our study highlights the potential that in silico analysis may have in the future formulation of a novel vaccine against pasteurellosis.

Commercial vaccines are available only for a few infectious diseases in aquaculture, and given that the needs for fish and shellfish protein as well as the number of cultured aquatic species are constantly increasing, novel vaccine development will be essential to guarantee improved animal welfare and a safe, sustainable, and healthy product.

## 5. Conclusions

In this study, we report the first fully sequenced genomes of *P. atlantica*, an emerging pathogen in *Cyclopterus lumpus* in Norwegian and Scottish aquaculture. The wide phylogenetic distance from the closest bacterial relative, *P. skyensis*, and the very few available *Pasteurellaceae* members of aquatic origin highlight the need for intensification of sequencing efforts that should be focused on *Pasteurella* strains from the aquaculture environment. In silico genomic analysis revealed numerous target genes that may be responsible for the virulence of this species, and a major role was attributed to mobile elements such as prophages and GIs. Functional analyses showed that the expression levels of an uncharacterized protein increased significantly when *P. atlantica* was exposed to leucocytes from lumpsuckers, complementing the in silico analysis by designating a prospective virulence gene and by promising candidate for vaccine development. Further assessment of virulence in larger scale as well as formulation of experimental vaccines will shed more light on devising a precautionary strategy against a serious emerging disease in aquaculture.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/microorganisms9061215/s1>, Supplementary Material S1: Database with an extensive description of the sequences from the proteome-filtering process (virulence factors and adhesins)

**Author Contributions:** Conceptualization, A.R. and H.I.W.; methodology, R.M.E., C.F., P.G.K., and A.R.; software, R.M.E., C.F., S.G., G.T.H., and P.G.K.; validation, R.M.E., C.F., P.G.K., and A.R.; formal analysis, R.M.E., C.F., and P.G.K.; investigation, R.M.E. and P.G.K.; resources, H.I.W.; data curation, R.M.E., C.F., and P.G.K.; writing—original draft preparation, R.M.E. and P.G.K.; writing—review and editing, D.J.C., R.M.E., C.F., S.G., G.T.H., P.G.K., M.M., A.R., and H.I.W.; visualization, R.M.E. and P.G.K.; supervision, P.G.K., A.R., and H.I.W.; project administration, A.R. and H.I.W.; funding acquisition, H.I.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the University of Bergen and supported by ACD Pharmaceuticals and the Norwegian Veterinary Institute.

**Institutional Review Board Statement:** This work was conducted according to the regulations of the Norwegian Food Safety Authority, National Assignment department, based on the EU directive (2010/63/EU) and defined as a non-regulated procedure.

**Data Availability Statement:** This whole genome project has been deposited at GenBank under the Accession Number PRJNA721926 for isolate UiBP1-2013 and accession number CP074346 for isolate NVI 9100.

**Acknowledgments:** R.M.E. thanks Harald Sæbø Lunde for his technical support during the design of the primers for the functional analyses in this work. PacBio sequencing was performed at the Norwegian Sequencing Centre (NSC) at the Department of Medical Genetics (AMG), Oslo University Hospital.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Johansen, R. *Fiskehelse rapporten 2012*; Veterinærinstituttet: Oslo, Norway, 2013.
2. Erkinharju, T.; Bornø, G.; Grønbech, S.; Gulla, S.; Hansen, H. Helsestatusjonen hos renseskild. In *Fiskehelse rapporten 2019*; Sommerset, I., Walde, C., Bang Jensen, B., Bornø, G., Haukaas, A., Brun, E., Eds.; Veterinærinstituttet: Oslo, Norway, 2020; pp. 134–143.
3. Nilsen, H.; Colquhoun, D.; Gulla, S. *Pasteurella* infeksjon hos laks. In *Fiskehelse rapporten 2020*; Sommerset, I., Bang Jensen, B., Bornø, G., Haukaas, A., Brun, E., Eds.; Veterinærinstituttet: Oslo, Norway, 2021; pp. 98–99.
4. Alarcón, M.; Gulla, S.; Røsæg, M.V.; Rønneseth, A.; Wergeland, H.; Poppe, T.T.; Nilsen, H.; Colquhoun, D.J. Pasteurellosis in lumpsucker *Cyclopterus lumpus*, farmed in Norway. *J. Fish Dis.* **2016**, *39*, 489–495. [[CrossRef](#)]
5. Kui, S. Stamfisk screening. In *Proceedings of the Renseskildkonferansen*; PatoGen Analyse: Trondheim, Norway, 2017.
6. Ellul, R.M.; Bulla, J.; Brudal, E.; Colquhoun, D.; Wergeland, H.; Rønneseth, A. Protection and antibody reactivity in lumpsucker (*Cyclopterus lumpus* L.) following vaccination against *Pasteurella* sp. *Fish Shellfish Immunol.* **2019**, *95*, 650–658. [[CrossRef](#)]

7. Bidmos, F.A.; Siris, S.; Gladstone, C.A.; Langford, P.R. Bacterial vaccine antigen discovery in the reverse vaccinology 2.0 Era: Progress and challenges. *Front. Immunol.* **2018**, *9*, 2315. [[CrossRef](#)]
8. Ong, E.; Cooke, M.F.; Huffman, A.; Xiang, Z.; Wong, M.U.; Wang, H.; Seetharaman, M.; Valdez, N.; He, Y. Vaxign2: The second generation of the first Web-based vaccine design program using reverse vaccinology and machine learning. *Nucleic Acids Res.* **2021**, gkab279. [[CrossRef](#)] [[PubMed](#)]
9. He, Y.; Xiang, Z.; Mobley, H.L.T. Vaxign: The first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J. Biomed. Biotechnol.* **2010**, *2010*, 297505. [[CrossRef](#)] [[PubMed](#)]
10. Sharma, A.; Sanduja, P.; Anand, A.; Mahajan, P.; Guzman, C.A.; Yadav, P.; Awasthi, A.; Hanski, E.; Dua, M.; Johri, A.K. Advanced strategies for development of vaccines against human bacterial pathogens. *World J. Microbiol. Biotechnol.* **2021**, *37*, 67. [[CrossRef](#)] [[PubMed](#)]
11. Pizza, M.; Scarlato, V.; Maignani, V.; Giuliani, M.M.; Aricò, B.; Comanducci, M.; Jennings, G.T.; Baldi, L.; Bartolini, E.; Capocchi, B.; et al. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science.* **2000**, *287*, 1816–1820. [[CrossRef](#)] [[PubMed](#)]
12. Maignani, V.; Pizza, M.; Moxon, E.R. The development of a vaccine against Meningococcus B using reverse vaccinology. *Front. Immunol.* **2019**, *10*, 751. [[CrossRef](#)] [[PubMed](#)]
13. Betts, J.C. Transcriptomics and Proteomics: Tools for the Identification of Novel Drug Targets and Vaccine Candidates for Tuberculosis. *IUBMB Life* **2002**, *53*, 239–242. [[CrossRef](#)]
14. Montigiani, S.; Falugi, F.; Scarselli, M.; Finco, O.; Petracca, R.; Galli, G.; Mariani, M.; Manetti, R.; Agnusdei, M.; Cevenini, R.; et al. Genomic approach for analysis of surface proteins in *Chlamydia pneumoniae*. *Infect. Immun.* **2002**, *70*, 368–379. [[CrossRef](#)]
15. Wizemann, T.M.; Heinrichs, J.H.; Adamou, J.E.; Erwin, A.L.; Kunsch, C.; Choi, G.H.; Barash, S.C.; Rosen, C.A.; Masure, H.R.; Tuomanen, E.; et al. Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. *Infect. Immun.* **2001**, *69*, 1593–1598. [[CrossRef](#)]
16. Chakravarti, D.N.; Fiske, M.J.; Fletcher, L.D.; Zagursky, R.J. Application of genomics and proteomics for identification of bacterial gene products as potential vaccine candidates. *Vaccine* **2000**, *19*, 601–612. [[CrossRef](#)]
17. Ross, B.C.; Czajkowski, L.; Hocking, D.; Margetts, M.; Webb, E.; Rothel, L.; Patterson, M.; Agius, C.; Camuglia, S.; Reynolds, E.; et al. Identification of vaccine candidate antigens from a genomic analysis of *Porphyromonas gingivalis*. *Vaccine* **2001**, *19*, 4135–4142. [[CrossRef](#)]
18. Ariel, N.; Zvi, A.; Grosfeld, H.; Gat, O.; Inbar, Y.; Velan, B.; Cohen, S.; Shafferman, A. Search for potential vaccine candidate open reading frames in the *Bacillus anthracis* virulence plasmid pXO1: *In silico* and *in vitro* screening. *Infect. Immun.* **2002**, *70*, 6817–6827. [[CrossRef](#)]
19. Kempell, K.E.; Kidd, S.P.; Lewandowski, K.; Elmore, M.J.; Charlton, S.; Yeates, A.; Cuthbertson, H.; Hallis, B.; Altmann, D.M.; Rogers, M.; et al. Whole genome protein microarrays for serum profiling of immunodominant antigens of *Bacillus anthracis*. *Front. Microbiol.* **2015**, *6*, 747. [[CrossRef](#)] [[PubMed](#)]
20. Ong, E.; Wang, H.; Wong, M.U.; Seetharaman, M.; Valdez, N.; He, Y. Vaxign-ML: Supervised machine learning reverse vaccinology model for improved prediction of bacterial protective antigens. *Bioinformatics* **2020**, *36*, 3185–3191. [[CrossRef](#)] [[PubMed](#)]
21. Abraham, S.; Sharon, N.; Ofek, I.; Schwartzman, J. Adhesion and colonisation. In *Molecular Medical Microbiology*; Sails, A., Poxton, I., Liu, D., Schwartzman, J., Eds.; Elsevier Science: Brussels, Belgium, 2014; pp. 409–421.
22. Moon, H.W.; Bunn, T.O. Vaccines for preventing enterotoxigenic *Escherichia coli* infections in farm animals. *Vaccine* **1993**, *11*, 213–220. [[CrossRef](#)]
23. Langermann, S.; Möllby, R.; Burlein, J.E.; Palaszynski, S.R.; Auguste, C.G.; DeFusco, A.; Strouse, R.; Schenerman, M.A.; Hultgren, S.J.; Pinkner, J.S.; et al. Vaccination with FimH Adhesin Protects Cynomolgus Monkeys from Colonization and Infection by Uropathogenic *Escherichia coli*. *J. Infect. Dis.* **2000**, *181*, 774–778. [[CrossRef](#)] [[PubMed](#)]
24. Fang, H.M.; Ge, R.; Sin, Y.M. Cloning, characterisation and expression of *Aeromonas hydrophila* major adhesin. *Fish Shellfish Immunol.* **2004**, *16*, 645–658. [[CrossRef](#)]
25. Maiti, B.; Shetty, M.; Shekar, M.; Karunasagar, I.; Karunasagar, I. Evaluation of two outer membrane proteins, Aha1 and OmpW of *Aeromonas hydrophila* as vaccine candidate for common carp. *Vet. Immunol. Immunopathol.* **2012**, *149*, 298–301. [[CrossRef](#)]
26. Zhu, Z.; Dong, C.; Weng, S.; He, J. Identification of outer membrane protein TolC as the major adhesin and potential vaccine candidate for *Vibrio harveyi* in hybrid grouper, *Epinephelus fuscoguttatus* (♀) × *E. lanceolatus* (♂). *Fish Shellfish Immunol.* **2019**, *86*, 143–151. [[CrossRef](#)] [[PubMed](#)]
27. Jin, C.-D.; Li, M.-F. *Edwardsiella tarda* outer membrane protein 1 is a virulence factor and a protective immunogen. *Aquaculture* **2021**, *539*, 736596. [[CrossRef](#)]
28. Ellul, R.M.; Walde, C.; Haugland, G.T.; Wergeland, H.; Rønneseth, A. Pathogenicity of *Pasteurella* sp. in lumpsuckers (*Cyclopterus lumpus* L.). *J. Fish Dis.* **2018**, *42*, 35–46. [[CrossRef](#)] [[PubMed](#)]
29. Nurk, S.; Bankevich, A.; Antipov, D.; Gurevich, A.A.; Korobeynikov, A.; Lapidus, A.; Prjibelski, A.D.; Pyshkin, A.; Sirotkin, A.; Sirotkin, Y.; et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J. Comput. Biol.* **2013**, *20*, 714–737. [[CrossRef](#)]
30. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [[CrossRef](#)]
31. Delcher, A.L.; Bratke, K.A.; Powers, E.C.; Salzberg, S.L. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **2007**, *23*, 673–679. [[CrossRef](#)] [[PubMed](#)]

32. Mutters, R.; Ihm, P.; Pohl, S.; Frederiksen, W.; Mannheim, W. Reclassification of the genus *Pasteurella* Trevisan 1887 on the basis of deoxyribonucleic acid homology, with proposals for the new species *Pasteurella dagmatis*, *Pasteurella canis*, *Pasteurella stomatis*, *Pasteurella anatis*, and *Pasteurella langaa*. *Int. J. Syst. Bacteriol.* **1985**, *35*, 309–322. [[CrossRef](#)]
33. Snipes, K.P.; Biberstein, E.L. *Pasteurella testudinis* sp. nov.: A Parasite of Desert Tortoises (*Gopherus agassizi*). *Int. J. Syst. Bacteriol.* **1982**, *32*, 201–210. [[CrossRef](#)]
34. Christensen, H.; Bertelsen, M.F.; Bojesen, A.M.; Bisgaard, M. Classification of *Pasteurella* species B as *Pasteurella oralis* sp. nov. *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 1396–1401. [[CrossRef](#)]
35. Birkbeck, T.H.; Laidler, L.A.; Grant, A.N.; Cox, D.I. *Pasteurella skyensis* sp. nov., isolated from Atlantic salmon (*Salmo salar* L.). *Int. J. Syst. Evol. Microbiol.* **2002**, *52*, 699–704. [[CrossRef](#)]
36. Foster, G.; Ross, H.M.; Malnick, H.; Willems, A.; Hutson, R.A.; Reid, R.J.; Collins, M.D. *Phocoenobacter uteri* gen. nov., sp. nov., a new member of the family Pasteurellaceae Pohl (1979) 1981 isolated from a harbour porpoise (*Phocoena phocoena*). *Int. J. Syst. Evol. Microbiol.* **2000**, *50*, 135–139. [[CrossRef](#)]
37. Topaz, N.; Boxrud, D.; Retchless, A.C.; Nichols, M.; Chang, H.Y.; Hu, F.; Wang, X. BMScan: Using whole genome similarity to rapidly and accurately identify bacterial meningitis causing species. *BMC Infect. Dis.* **2018**, *18*, 1–9. [[CrossRef](#)] [[PubMed](#)]
38. Roach, D.J.; Burton, J.N.; Lee, C.; Stackhouse, B.; Butler-Wu, S.M.; Cookson, B.T.; Shendure, J.; Salipante, S.J. A Year of Infection in the Intensive Care Unit: Prospective Whole Genome Sequencing of Bacterial Clinical Isolates Reveals Cryptic Transmissions and Novel Microbiota. *PLoS Genet.* **2015**, *11*, e1005413. [[CrossRef](#)] [[PubMed](#)]
39. Nichols, M.; Topaz, N.; Wang, X.; Wang, X.; Boxrud, D. Draft genome sequences for a diverse set of isolates from 10 *Neisseria* species. *Genome Announc.* **2018**, *6*. [[CrossRef](#)]
40. Dickerman, A.; Bandara, A.B.; Inzana, T.J. Phylogenomic analysis of *Haemophilus parasuis* and proposed reclassification to *Glaesserella parasuis*, gen. Nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* **2020**, *70*, 180–186. [[CrossRef](#)]
41. Chiers, K.; Haesebrouck, F.; Mateusen, B.; Van Overbeke, I.; Ducatelle, R. Pathogenicity of *Actinobacillus minor*, *Actinobacillus indolicus* and *Actinobacillus porcicus* Strains for Gnotobiotic Piglets. *J. Vet. Med. Sci.* **2001**, *131*, 127–131. [[CrossRef](#)]
42. Eidam, C.; Poehlein, A.; Michael, G.B.; Kadlec, K.; Liesegang, H.; Brzuszkiewicz, E.; Daniel, R.; Sweeney, M.T.; Murray, R.W.; Watts, J.L.; et al. Complete genome sequence of *Mannheimia haemolytica* strain 42548 from a case of bovine respiratory disease. *Genome Announc.* **2013**, *1*, 2010–2011. [[CrossRef](#)]
43. Lee, I.; Kim, Y.O.; Park, S.C.; Chun, J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int. J. Syst. Evol. Microbiol.* **2016**, *66*, 1100–1103. [[CrossRef](#)]
44. Ondov, B.D.; Treangen, T.J.; Melsted, P.; Mallonee, A.B.; Bergman, N.H.; Koren, S.; Phillippy, A.M. Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biol.* **2016**, *17*, 132. [[CrossRef](#)]
45. Wattam, A.R.; Abraham, D.; Dalay, O.; Disz, T.L.; Driscoll, T.; Gabbard, J.L.; Gillespie, J.J.; Gough, R.; Hix, D.; Kenyon, R.; et al. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res.* **2014**, *42*, D581–D591. [[CrossRef](#)]
46. Wattam, A.R.; Davis, J.J.; Assaf, R.; Boisvert, S.; Brettin, T.; Bun, C.; Conrad, N.; Dietrich, E.M.; Disz, T.; Gabbard, J.L.; et al. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Res.* **2017**, *45*, D535–D542. [[CrossRef](#)] [[PubMed](#)]
47. Davis, J.J.; Gerdes, S.; Olsen, G.J.; Olson, R.; Pusch, G.D.; Shukla, M.; Vonstein, V.; Wattam, A.R.; Yoo, H. PATtyFams: Protein families for the microbial genomes in the PATRIC database. *Front. Microbiol.* **2016**, *7*, 1–12. [[CrossRef](#)] [[PubMed](#)]
48. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
49. Cock, P.J.A.; Antao, T.; Chang, J.T.; Chapman, B.A.; Cox, C.J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; et al. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **2009**, *25*, 1422–1423. [[CrossRef](#)]
50. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]
51. Stamatakis, A.; Hoover, P.; Rougemont, J. A rapid bootstrap algorithm for the RAxML web-servers. *Syst. Biol.* **2008**, *57*, 758–771. [[CrossRef](#)]
52. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)]
53. Liu, B.; Zheng, D.; Jin, Q.; Chen, L.; Yang, J. VFDB 2019: A comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* **2019**, *47*, D687–D692. [[CrossRef](#)]
54. Sayers, S.; Li, L.; Ong, E.; Deng, S.; Fu, G.; Lin, Y.; Yang, B.; Zhang, S.; Fa, Z.; Zhao, B.; et al. Victors: A web-based knowledge base of virulence factors in human and animal pathogens. *Nucleic Acids Res.* **2019**, *47*, D693–D700. [[CrossRef](#)]
55. Mao, C.; Abraham, D.; Wattam, A.R.; Wilson, M.J.C.; Shukla, M.; Yoo, H.S.; Sobral, B.W. Curation, integration and visualization of bacterial virulence factors in PATRIC. *Bioinformatics* **2015**, *31*, 252–258. [[CrossRef](#)]
56. Yu, N.Y.; Wagner, J.R.; Laird, M.R.; Melli, G.; Rey, S.; Lo, R.; Dao, P.; Sahinalp, S.C.; Ester, M.; Foster, L.J.; et al. PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **2010**, *26*, 1608–1615. [[CrossRef](#)]
57. Armenteros, J.J.A.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; Von Heijne, G.; Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **2019**, *37*, 420–423. [[CrossRef](#)]

58. Sachdeva, G.; Kumar, K.; Jain, P.; Ramachandran, S. SPAAN: A software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics* **2005**, *21*, 483–491. [[CrossRef](#)]
59. Kelly, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)]
60. Bertelli, C.; Laird, M.R.; Williams, K.P.; Lau, B.Y.; Hoad, G.; Winsor, G.L.; Brinkman, F.S.L.; Simon Fraser University Research Computing Group. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* **2017**, *45*, W30–W35. [[CrossRef](#)]
61. Arndt, D.; Grant, J.R.; Marcu, A.; Sajed, T.; Pon, A.; Liang, Y.; Wishart, D.S. PHASTER: A better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* **2016**, *44*, W16–W21. [[CrossRef](#)] [[PubMed](#)]
62. Lopes, A.; Tavares, P.; Petit, M.-A.; Guérois, R.; Zinn-Justin, S. Automated classification of tailed bacteriophages according to their neck organization. *BMC Genomics* **2014**, *15*, 1027. [[CrossRef](#)] [[PubMed](#)]
63. Løvoll, M.; Wiik-Nielsen, C.R.; Tunsjø, H.S.; Colquhoun, D.; Lunder, T.; Sørum, H.; Grove, S. Atlantic salmon bath challenged with *Moritella viscosa*—Pathogen invasion and host response. *Fish Shellfish Immunol.* **2009**, *26*, 877–884. [[CrossRef](#)] [[PubMed](#)]
64. Doytchinova, I.A.; Flower, D.R. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* **2007**, *8*, 4. [[CrossRef](#)] [[PubMed](#)]
65. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)]
66. Haugland, G.T.; Jakobsen, R.A.; Vestvik, N.; Ulven, K.; Stokka, L.; Wergeland, H.I. Phagocytosis and Respiratory Burst Activity in Lumpfish (*Cyclopterus lumpus* L.) Leucocytes Analysed by Flow Cytometry. *PLoS ONE* **2012**, *7*, e47909. [[CrossRef](#)]
67. Castillo, D.; Alvise, P.D.; Xu, R.; Zhang, F.; Middelboe, M.; Gram, L. Comparative Genome Analyses of *Vibrio anguillarum* Strains Reveal a Link with Pathogenicity Traits. *mSystems* **2017**, *2*, e00001-17. [[CrossRef](#)] [[PubMed](#)]
68. Hurtado, R.; Carhuaricra, D.; Soares, S.; Viana, M.V.C.; Azevedo, V.; Maturrano, L.; Aburjaile, F. Pan-genomic approach shows insight of genetic divergence and pathogenic-adaptation of *Pasteurella multocida*. *Gene* **2018**, *670*, 193–206. [[CrossRef](#)]
69. Zhu, D.; He, J.; Yang, Z.; Wang, M.; Jia, R.; Chen, S.; Liu, M.; Zhao, X.; Yang, Q.; Wu, Y.; et al. Comparative analysis reveals the Genomic Islands in *Pasteurella multocida* population genetics: On Symbiosis and adaptability. *BMC Genom.* **2019**, *20*, 63. [[CrossRef](#)]
70. Niu, Y.D.; Cook, S.R.; Wang, J.; Klima, C.L.; Hsu, Y.; Kropinski, A.M.; Turner, D.; Mcallister, T.A. Comparative analysis of multiple inducible phages from *Mannheimia haemolytica*. *BMC Microbiol.* **2015**, *15*, 175. [[CrossRef](#)]
71. Zehr, E.S.; Tabatabai, L.B. Detection of a bacteriophage gene encoding a Mu-like portal protein in *Haemophilus parasuis* reference strains and field isolates by nested polymerase chain reaction. *J. Vet. Diagnostic Investig.* **2011**, *23*, 538–542. [[CrossRef](#)]
72. Klemm, P.; Schembri, M.A. Bacterial adhesins: Function and structure. *Int. J. Med. Microbiol.* **2000**, *290*, 27–35. [[CrossRef](#)]
73. Nummelin, H.; Merckel, M.C.; Leo, J.C.; Lankinen, H.; Skurnik, M.; Goldman, A. The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel  $\beta$ -roll. *EMBO J.* **2004**, *23*, 701–711. [[CrossRef](#)]
74. St. Geme, J.W.; Cutter, D. The *Haemophilus influenzae* Hia adhesin is an autotransporter protein that remains uncleaved at the C terminus and fully cell associated. *J. Bacteriol.* **2000**, *182*, 6005–6013. [[CrossRef](#)]
75. Rosenberg, E.; DeLong, E.F.; Lory, S.; Stackebrandt, E.; Thompson, F. (Eds.) *The Prokaryotes: Gammaproteobacteria*; Springer: Berlin/Heidelberg, Germany, 2013; ISBN 9783642389221.
76. Machado, H.; Gram, L. The *fur* gene as a new phylogenetic marker for Vibrionaceae species identification. *Appl. Environ. Microbiol.* **2015**, *81*, 2745–2752. [[CrossRef](#)] [[PubMed](#)]
77. Naushad, S.; Adeolu, M.; Goel, N.; Khadka, B.; Al-Dahwi, A.; Gupta, R.S. Phylogenomic and molecular demarcation of the core members of the polyphyletic Pasteurellaceae genera *Actinobacillus*, *Haemophilus*, and *Pasteurella*. *Int. J. Genomics* **2015**, *2015*, 19856. [[CrossRef](#)]
78. Wilkie, I.W.; Harper, M.; Boyce, J.D.; Adler, B. *Pasteurella multocida*: Diseases and Pathogenesis. In *Pasteurella multocida: Molecular Biology, Toxins and Infection*; Aktories, K., Orth, J.H.C., Adler, B., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 1–22. ISBN 978-3-642-31017-1.
79. García-Alvarez, A.; Vela, A.I.; San Martín, E.; Chaves, F.; Fernández-garayzabal, J.F.; Lucas, D.; Cid, D. Characterization of *Pasteurella multocida* associated with ovine pneumonia using multi-locus sequence typing (MLST) and virulence-associated gene profile analysis and comparison with porcine isolates. *Vet. Microbiol.* **2017**, *204*, 180–187. [[CrossRef](#)]
80. Legård, B.K.; Strøm, S.B. Pasteurellosis in Atlantic salmon (*Salmo salar*) in Western Norway. *Bull. Eur. Assoc. Fish Pathol.* **2020**, *40*, 148–155.
81. Castillo, D.; Kauffman, K.; Hussain, F.; Kalatzis, P.; Rørbo, N.; Polz, M.F.; Middelboe, M. Widespread distribution of prophage-encoded virulence factors in marine *Vibrio* communities. *Sci. Rep.* **2018**, *8*. [[CrossRef](#)] [[PubMed](#)]
82. Hundt, M.J.; Ruffolo, C.G. Interaction of *Pasteurella multocida* with free-living amoebae. *Appl. Environ. Microbiol.* **2005**, *71*, 5458–5464. [[CrossRef](#)]
83. Bordenstein, S.R.; Reznikoff, W.S. Mobile DNA in obligate intracellular bacteria. *Nat. Rev. Microbiol.* **2005**, *3*, 688–699. [[CrossRef](#)] [[PubMed](#)]
84. Moon, B.Y.; Park, J.Y.; Robinson, D.A.; Thomas, J.C.; Park, Y.H.; Thornton, J.A.; Seo, K.S. Mobilization of genomic islands of *Staphylococcus aureus* by temperate bacteriophage. *PLoS ONE* **2016**, *11*, 1–16. [[CrossRef](#)] [[PubMed](#)]

85. Ingle, D.J.; Tauschek, M.; Edwards, D.J.; Hocking, D.M.; Pickard, D.J.; Azzopardi, K.I.; Amarasena, T.; Bennett-Wood, V.; Pearson, J.S.; Tamboura, B.; et al. Evolution of atypical enteropathogenic *E. coli* by repeated acquisition of LEE pathogenicity island variants. *Nat. Microbiol.* **2016**, *1*, 15010. [[CrossRef](#)] [[PubMed](#)]
86. Zhang, R.; Ou, H.; Gao, F.; Luo, H. Identification of Horizontally-transferred Genomic Islands and Genome Segmentation Points by Using the GC Profile Method. *Curr. Genom.* **2014**, *15*, 113–121. [[CrossRef](#)]
87. Marcus, S.L.; Brumell, J.H.; Pfeifer, C.G.; Finlay, B.B. *Salmonella* pathogenicity islands: Big virulence in small packages. *Microbes Infect.* **2000**, *2*, 145–156. [[CrossRef](#)]
88. Vivona, S.; Bernante, F.; Filippini, F. NERVE: New Enhanced Reverse Vaccinology Environment. *BMC Biotechnol.* **2006**, *6*, 35. [[CrossRef](#)]
89. Dalsass, M.; Brozzi, A.; Medini, D.; Rappuoli, R. Comparison of open-source reverse vaccinology programs for bacterial vaccine antigen discovery. *Front. Immunol.* **2019**, *10*, 1–12. [[CrossRef](#)]
90. He, Y. Bacterial Whole-Genome Determination and Applications. In *Molecular Medical Microbiology*; Academic Press: Amsterdam, The Netherlands, 2015; ISBN 9780123971692.
91. Spahich, N.A.; St Geme, J.W. Structure and function of the *Haemophilus influenzae* autotransporters. *Front. Cell. Infect. Microbiol.* **2011**, *1*, 5. [[CrossRef](#)]
92. Thibau, A.; Dichter, A.A.; Vaca, D.J.; Linke, D.; Goldman, A.; Kempf, V.A.J. Immunogenicity of trimeric autotransporter adhesins and their potential as vaccine targets. *Med. Microbiol. Immunol.* **2020**, *209*, 243–263. [[CrossRef](#)]
93. El Tahir, Y.; Skurnik, M. YadA, the multifaceted *Yersinia* adhesin. *Int. J. Med. Microbiol.* **2001**, *291*, 209–218. [[CrossRef](#)]
94. Tsugo, K.; Nakamura, S.; Yamanaka, H.; Une, Y. A study on the efficacy of the recombinant *Yersinia* adhesin A vaccine against yersiniosis in the early phase. *J. Vet. Med. Sci.* **2017**, *79*, 855–863. [[CrossRef](#)]
95. Linke, D.; Riess, T.; Autenrieth, I.B.; Lupas, A.; Kempf, V.A.J. Trimeric autotransporter adhesins: Variable structure, common function. *Trends Microbiol.* **2006**, *14*, 264–270. [[CrossRef](#)]
96. Winter, L.E.; Barenkamp, S.J. Antibodies specific for the Hia adhesion proteins of nontypeable *Haemophilus influenzae* mediate opsonophagocytic activity. *Clin. Vaccine Immunol.* **2009**, *16*, 1040–1046. [[CrossRef](#)]
97. Winter, L.E.; Barenkamp, S.J. Immunogenicity of Nontypeable *Haemophilus influenzae* Outer Membrane Vesicles and Protective Ability in the Chinchilla Model of Otitis Media. *Clin. Vaccine Immunol.* **2017**, *24*. [[CrossRef](#)]
98. Raynes, J.M.; Young, P.G.; Proft, T.; Williamson, D.A.; Baker, E.N.; Moreland, N.J. Protein adhesins as vaccine antigens for Group A *Streptococcus*. *Pathog. Dis.* **2018**, *76*, 16. [[CrossRef](#)]
99. Ma, J.; Bruce, T.J.; Jones, E.M.; Cain, K.D. A Review of Fish Vaccine Development Strategies: Conventional Methods and Modern Biotechnological Approaches. *Microorganisms* **2019**, *7*, 569. [[CrossRef](#)] [[PubMed](#)]
100. Marana, M.H.; Von Gersdorff Jørgensen, L.; Skov, J.; Chettri, J.K.; Mattsson, A.H.; Dalsgaard, I.; Kania, P.W.; Buchmann, K. Subunit vaccine candidates against *Aeromonas salmonicida* in rainbow trout *Oncorhynchus mykiss*. *PLoS ONE* **2017**, *12*, 1–15. [[CrossRef](#)]
101. Shahin, K.; Pirezan, F.; Rogge, M.; LaFrentz, B.R.; Shrestha, R.P.; Hildebrand, M.; Lu, F.; HogenEsch, H.; Soto, E. Development of IglC and GroEL recombinant vaccines for francisellosis in Nile tilapia, *Oreochromis niloticus*. *Fish Shellfish Immunol.* **2020**, *105*, 341–349. [[CrossRef](#)] [[PubMed](#)]
102. Khushiramani, R.; Girisha, S.K.; Karunasagar, I.; Karunasagar, I. Protective efficacy of recombinant OmpTS protein of *Aeromonas hydrophila* in Indian major carp. *Vaccine* **2007**, *25*, 1157–1158. [[CrossRef](#)] [[PubMed](#)]
103. Hamod, M.A.; Nithin, M.S.; Shukur, Y.N.; Karunasagar, I.; Karunasagar, I. Outer membrane protein K as a subunit vaccine against *V. anguillarum*. *Aquaculture* **2012**, *354–355*, 107–110. [[CrossRef](#)]
104. Khushiramani, R.M.; Maiti, B.; Shekar, M.; Girisha, S.K.; Akash, N.; Deepanjali, A.; Karunasagar, I.; Karunasagar, I. Recombinant *Aeromonas hydrophila* outer membrane protein 48 (Omp48) induces a protective immune response against *Aeromonas hydrophila* and *Edwardsiella tarda*. *Res. Microbiol.* **2012**, *163*, 286–291. [[CrossRef](#)]