Annals of Bangladesh Agriculture
Journal homepage: bsmrau.edu.bd/aba

ORIGINAL ARTICLES

Comparative genomics analyses of fish pathogenic *Streptococcus* spp. isolated from tilapia and flounder

Tasmina Akter[1*]

[1]Department of Fisheries Management, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur 1706, Bangladesh

ARTICLE INFO

Keywords:
Streptococcus, Streptococcosis, virulence factors, antibiotic resistance, pathogenicity.

Received : 09 November 2023
Revised : 24 December 2023
Accepted : 28 December 2023
Published : 30 December 2023

Citation:
Akter, T. 2023. Comparative genomics analyses of fish pathogenic *Streptococcus* spp. isolated from tilapia and flounder. *Ann. Bangladesh Agric*. 27(2): 1-18.

ABSTRACT

This comparative study was performed to analyze the whole genome sequences of three species of *Streptococcus* (*S. agalactiae, S. iniae* and *S. parauberis*) isolated from Tilapia and Flounder from six different countries (Brazil, China, Israel, South Korea, Taiwan and USA). The objectives were to compare the genomic features, virulence, and antibiotic resistance genes in the genome sequences of 11 isolates of *Streptococcus* spp. A total of 44 virulence genes were identified in the genomes of 11 strains. These genes are responsible for adherence, various enzyme production, immune evasion, immunoreactive antigen, and toxin production. Eight antibiotic resistance genes were identified in the eleven genome sequences of the *Streptococcus* sp. strains. All strains of *S. agalactiae* and *S. iniae* harbor macrolide resistance gene *mreA*. Although five secondary metabolites such as Arylpolyene (*ary*), Type III Polyketide synthases (*T3PKS*), RiPP-like peptide, linear azol(in)e-containing peptides (*LAPs*), *RaS-RiPP* antimicrobial compound, and *T3PKS* were detected in all 11 genomes; only *T3PKS* was common in all strains. Additionally, *Cas* cluster CAS-TypeIC and CAS-TypeIIA were identified among the ten strains of *S. agalactiae* and *S. iniae*. The findings indicated that the degree of pathogenicity of *Streptococcus* sp. remained closer regardless of origin, distribution and host. The results would be useful to understand the virulence factors of the *Streptococcus* sp. and the antibiotic resistance genes associated with their virulence in fish.

Introduction

Rapid growth and intensification of aquaculture face great challenges of disease infestation and causes economic loss in fish farming. Bacteria are the leading pathogens both in fresh and marine water fishes all over the world. Streptococcosis is an infectious septicemia disease caused by some Gram-positive cocci bacteria belonging to the genera *Streptococcus, Lactococcus, Vagococcus,* and *Enterococcus* that affects both wild and culture

[*]**Corresponding Author**: Department of Fisheries Management, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur 1706, Bangladesh. Email: tasmina.fmg@bsmrau.edu.bd

https://doi.org/10.3329/aba.v27i2.72514
ISSN 1025-482X (Print)/2521-5477 (Online)

species of fish (Rahman *et al.*, 2017; Toranzo *et al.*, 2005). It was found that streptococcosis infection caused in a variety of economically important fish species (i.e., tilapia, flounder, sea bass, rainbow trout etc.) in the different parts of the world (Toranzo *et al.*, 2005). These bacteria produce almost similar disease symptoms in fish. General pathological symptoms of streptococcosis in fishes leads to various clinical signs which include corneal opacity, hemorrhages of the eye and gill plate, loss of appetite, spine displacement, hemorrhages at the base of fins and in the opercula (Akter *et al.*, 2021). The most prominent signs are uni- or bi-lateral exophthalmia, also known as pop-eye. The bacteria can also attack the central nervous system of fish and lead to all sorts of erratic behavior (Akter *et al.*, 2021; Legario *et al.*, 2020). As multiple complex pathogens are responsible for fish streptococcosis, it is difficult to identify the bacteria accurately and thereby take control measures. A significant economic loss has been reported globally due to streptococcal infection in the fish farm (Toranzo *et al.*, 2005).

The pathogens were identified by using various molecular methods such as 16S rRNA, whole genome sequencing (Akter *et al.*, 2021; Facimoto *et al.*, 2017; Nho *et al.*, 2011). Recently, several researchers isolated and identified a numbers of *Streptococcus* sp. such as *S. agalactiae*, *S. iniae* and *S. parauberis* from different types of diseased fish in the different part of the world (Facimoto *et al.*, 2017; Nho *et al.*, 2011). Genome sequencing of fish pathogenic *Streptococcus* sp. has increased our knowledge on host adaptation and virulence factors. Genomic study of different functional as well as virulence genes is crucial to know about the host specificity, biology, and adaptation mechanism of the pathogen (Akter *et al.*, 2023). There is a large variation in the size and content of bacterial genomes among different genera, species and strains of the same species. Closely related bacteria generally have very similar genomes. Although three species of fish pathogenic *Streptococcus* such as *S. agalactiae*, *S. iniae*, and *S. parauberis* were studied by whole genome sequences, no comparative studies were performed as they cause similar diseases in fish. The objectives of the present piece of researher wre to *in-silico* analyses of the whole genome sequence of a number of strains of *S. agalactiae*, *S. iniae* and *S. parauberis,* and to compare their genomic features especially their virulence genes, antibiotic resistance genes, and secondary metabolites.

Materials and Methods

Assembly of whole genome sequence data

The whole genome sequences of 11 isolates of fish pathogenic *Streptococcus* sp. belonging to three species (*S. agalactiae*, *S. iniae, S. parauberis*) were obtained from the National Center for Biotechnology Information (NCBI) genome repository in the Fasta format (Table 1). The isolates were obtained from two fish (Tilapia and Flounder) species from six different countries such as Brazil, China, Israel, South Korea, Taiwan, and USA.

Bioinformatic in-silico analysis

Genomic comparison

An online server RAST (Rapid Annotation using Subsystem Technology (http://rast.nmpdr.org/rast.cgi, (Overbeek *et al.*, 2014) was used to identify the coding regions (CDs) among 11 isolates of *Streptococcus* sp. The SEED Viewer (http://rast.nmpdr.org/seedviewer.cgi) (Overbeek *et al.*, 2014) was used for comparison of the unique and shared genes of these strains. To detect the plasmids, web tool PlasmidFinder (https://cge.cbs.dtu.dk/services/PlasmidFinder/) was used with the setting of the threshold for a minimum 95% identity over 60% coverage of length50 (Carattoli *et al.*, 2014).

Table1: The list of fish pathogenic *Streptococcus* spp. strains used for the present study. The whole genome sequence data of these strains were obtained from the NCBI database

Sl. No.	Species	Strains	Isolated Fish	Origin	NCBI Accession No.
1	*Streptococcus agalactiae*	GD201008-001	Tilapia	China	CP003810
2	*S. agalactiae*	S25	Tilapia	Brazil	CP015976
3	*S. agalactiae*	SA20	Tilapia	Brazil	CP003919
4	*S. agalactiae*	SA623	Tilapia	Brazil	CP019836
5	*S. agalactiae*	S13	Tilapia	Brazil	CP018623
6	*S. iniae*	ISET0901	Tilapia	Israel	CP007586
7	*S. iniae*	89353	Tilapia	Taiwan	CP017952
8	*S. iniae*	SF1	Flounder	China	CP005941
9	*S. iniae*	ISNO	Tilapia	USA	CP007587
10	*S. iniae*	YSFST01-82	Flounder	South Korea	CP010783
11	*S. parauberis*	KCTC11537	Flounder	South Korea	CP002471

The BRIG (BLAST Ring Image Generator) was applied to generate a circular genomic plot to compare the similarities of the genome among same species as well as with other two species of *Streptococcus* (Alikhan *et al.*, 2011) . As a reference sequence *S. agalactiae* strain 2603V/R (NCBI accession No. AE009948) was used to generate the ring image. The prophage hunter web tool PHASTER server (http://phaster.ca/) (Arndt *et al.*, 2016) was used to identify the prophage-associated Gene clusters in the genome sequences of the studied isolates. Based on the known genes/proteins in the predicted phage-associated regions, this tool matches three scenarios such as intact (≥90%), questionable (90–60%), and incomplete (≤60%). The position of the prophase regions identified in the genome sequences were visualized on the circular map obtained from BRIG software (Alikhan *et al.*, 2011).

Assembly and identification of interested virulence gene

The virulence genes were identified in the *de novo* assembled contigs of 11 *Streptococcus* sp. strains using the online web-service VFanalyzer of virulence factor database (VFDB, http://www.mgc.ac.cn/VFs/) (Liu *et al.*, 2018) . The common genes to all of the strains were identified using default settings of the server database. The presence and absence of the virulence genes among the strains were visualized using R statistical computing program.

Identification of antibiotic resistance gene (ARG)

Three online databases were used to predict the antibiotic resistance genes in the whole genome sequence of 11 strains namely Comprehensive Antibiotic-resistance Database CARD (CARD, https://card.mcmaster.ca/analyze/rgi) (McArthur *et al.*, 2013) , Antibiotic-resistance Gene-ANNOTation V6 (ARG-ANNOT, https://ifr48.timone.univ-mrs.fr/blast/arg-annot_v6.html) (Gupta *et al.*, 2014) and ResFinder 3.1 (https://cge.cbs.dtu.dk/services/ResFinder/) (Zankari *et al.*, 2012) . Default setting for the web databases was used to identify the ARG.

Analysis of prophage region, Secondary metabolites and CRISPR-Cas

Gene cluster of Secondary metabolites synthesis were identified by another online database namely Antibiotics and Secondary Metabolite Analysis Shell V 5.1.2 (antiSMASH, https://antismash.secondarymetabolites.org/#!/start)51) (Medema *et al.*, 2011). The Clustered Regularly Interspaced Short Palindromic Repeats, (CRISPR) and CRISPR-associated genes (cas) were predicted by using an online program namely the CRISPRCasFinder v.1.1.2.(https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index) (Couvin *et al.*, 2018). The statistics obtained from secondary metabolites and CRISPR-Cas analyses were visualized using the R statistical computing program.

Analysis of phylogenetic tree

To study the similarities among the bacterial strains of *Streptococcus* sp., the online web-tool CSI Phylogeny v1.2 (https://cge.cbs.dtu.dk/services/CSIPhylogeny/) (Kaas *et al.*, 2014) was used to construct a Phylogenetic tree. The analysis of phylogenetic tree was carried out on the basis of single nucleotide polymorphism (SNP). The obtained figure of phylogenetic tree was modified using FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). As an output from the phylogenetic analysis, a matrix was obtained which contained the counts of nucleotides difference for all sequences. Default values were used in the server during the analysis of the SNP tree (Kaas *et al.*, 2014).

Results and Discussion

General features of the genomes

An overview of the genome features of the 11 streptococcal strains and their subsystem statistics were shown in Table 2. According to the subsystem analysis of seed viewer, the genome size of the 5 strains of *S. agalactiae* and 5 strains of *S. iniae* varied from 1.84 to 2.06 and 2.07 to 2.15 Mb, respectively. The genomic size of *S. parauberis* KCTC3651 was 2.14Mb. The GC (%) contents among the 11 isolates were very identical according to species. The predicted coding sequence (CDSs) of the *S. agalactiae* and *S. iniae* strains were ranged from 1919 to 2035 and 2022 to 2177, respectively. *Streptococcus parauberis* KCTC3651strain yielded 2423 CDs. Only 30 to 46% (less than 50%) CDSs in each isolate could be functionally categorized into 222 to 569 subsystems (Table 2).

The basic genomic characteristics and subsystem features of different strains in the current study were almost similar according to species of the pathogens. Two strains of *S. agalactiae* (FNA07, FPrA02) from the Nile tilapia were isolated in Thailand with genome sizes 2.1 and 2.05 Mb with GC content 35.5 and 35.4 %, respectively (Kayansamruaj *et al.*, 2015). The genomic features of *S. agalactiae* in the current study were found to be very similar to the strains GBS85147 and HU-GS5823 isolated from humans (de Aguiar *et al.*, 2016; Nagaoka *et al.*, 2018). While reduced genomic size of *S. agalactae* were identified from four Brazilian isolates compared with that reported in other study (Liu *et al.*, 2013) . Almost similar genomic size of *S. iniae* UEL-Si1 (2,395,193 bp) was identified with an average GC content 36.3% from diseased Nile Tilapia in Sothern Brazil (Vilas-Boas *et al.*, 2017) .

According to RAST and Seed viewer analysis, less than 50% subsystem coverage were identified in all strains of *Streptococcus* sp. and a large number of genes were found to be responsible for carbohydrates, amino acids, and derivatives of protein metabolism. No plasmids were found in the genome sequences of all strains of *Streptococcus* using the PlasmidFinder web-tool. It was found that the genome sequences were identical according to the species of pathogens (Fig. 1). Similar results were found in the other pathogenic bacteria responsible for streptococcosis in fish (Akter *et al.*, 2023; Kayansamruaj *et al.*, 2015).

Prophages are transposable elements that can take part in different cellular approaches i.e., they can develop virulence characteristics or acquire antibiotic resistance genes or develop harmful metabolic pathways to adapt to new environment. Prophage analysis with PHASTER web tool detected a number of prophage regions with variable lengths in all the strains of *Streptococcus* sp (Fig. 1). More specifically, PHASTER analysis showed that all strains of *S. agalactiae* conserved one incomplete prophage region with different length at a completeness score ranging from 20 to 50. A total of 42 prophage were detected from the five strains of *S. iniae* under three categories of PHASTER service tool. Interestingly, three East Asian *S. iniae* strains 89353, SF1, and YSFST01-82 conserved the highest (9) phages region, where each of them have 7 incomplete regions (Fig. 1). On the other hand, eight incomplete prophages were detected in the genome sequence of *S. iniae* strain YSFST01-82. Furthermore, streptococcal strains ISET0901 and ISNO conserved 8 and 7 prophage regions, respectively (Fig.1). Moreover, four prophage regions were detected from *S. parauberis* KCTC3651 that matched with 2 intact, 1 incomplete, and 1 questionable region. Likewise, the identical genome sequences of species, and the prophage region of the same species of *Streptococcus* were

also clustered in the same area on the genome (Fig. 1). Similar to our results, incomplete bacteriophage sequences were identified in two other Brazilian strains (LGMAI_St_11, and LGMAI_St_14) of *S. agalactiae* (Vidal Amaral *et al.*, 2022). Furthermore, incomplete prophage was identified from two Thai

Table 2: General features of the genomes and their subsystems of 11 isolates of fish pathogenic *Streptococcus* spp. responsible for streptococcosis

	Species										
	S. agalactiae						*S. iniae*				*S. parauberis*
Strains	GD201 008-001	S13	S25	SA20	SA623	ISET 0901	89353	SF1	ISNO	YSFST 01- 82	KCTC 11537
Size (Mb)	2.06	1.84	1.84	1.84	1.84	2.07	2.10	2.15	2.07	2.09	2.14
GC Content (%)	35.6	35.4	35.5	35.5	35.5	36.8	36.8	36.7	36.8	36.8	35.5
Subsystem coverage (%)	44	34	46	46	46	42	42	41	42	30	41
Number of Subsystems	531	224	520	520	520	553	554	557	551	222	569
Number of Coding Sequences (CDs)	2035	1919	1927	1938	1939	2057	2033	2177	2058	2022	2423
Number of RNAs	98	74	97	81	83	57	86	58	57	73	75
Subsystem feature counts											
Cofactors, Vitamins, Prosthetic Groups, Pigments	681	76	639	638	638	905	914	935	905	64	815
Cell Wall and Capsule	709	61	650	649	652	823	831	845	821	51	949
Virulence, Disease and Defense	510	27	465	463	463	536	540	549	535	35	532
Potassium metabolism	105	3	92	92	92	110	114	114	110	3	113
Photosynthesis	1	0	1	1	1	2	2	2	3	0	0
Miscellaneous	134	11	99	98	98	152	154	155	152	11	169
Phages, Prophages, Transposable elements, Plasmids)	132	0	64	64	64	90	91	173	90	0	234
Membrane Transport	476	29	442	442	442	488	488	501	487	25	515
Iron acquisition and metabolism	77	15	77	77	77	135	138	142	135	21	43
RNA Metabolism	606	30	529	529	528	659	662	659	656	30	667
Nucleosides and Nucleotides	461	91	428	427	427	493	500	513	495	58	512
Protein Metabolism	1092	110	1060	1062	1061	1214	1222	1239	1212	115	1117
Cell Division and Cell Cycle	206	4	204	204	204	226	228	231	225	4	250
Motility and Chemotaxis	56	0	48	48	48	86	88	92	86	0	118
Regulation and Cell signaling	265	18	244	242	242	190	191	191	188	19	307
Secondary Metabolism	13	0	14	14	14	45	47	49	45	8	10
DNA Metabolism	772	58	748	742	742	858	865	908	858	45	975
Fatty Acids, Lipids, and Isoprenoids	458	41	420	420	420	553	578	564	553	22	648
Nitrogen Metabolism	85	3	78	78	78	79	79	81	79	0	100
Dormancy and Sporulation	44	1	40	39	39	7	7	7	7	1	6
Respiration	283	16	248	246	246	325	327	333	324	17	360
Stress Response	361	22	348	346	345	413	424	430	414	16	461
Metabolism of Aromatic Compounds	42	2	35	35	35	83	82	81	83	2	66
Amino Acids and Derivatives	1094	109	1123	1122	1124	1480	1487	1511	1477	111	1482
Sulfur Metabolism	107	5	75	75	74	135	139	139	135	4	143
Phosphorus Metabolism	243	5	232	232	232	249	250	258	249	4	278
Carbohydrates	1701	165	1726	1717	1716	2088	2121	2138	2085	154	2283

fish strains of *S. agalactiae* (ST7 and CF01173) (Kayansamruaj *et al.*, 2015). Compared with the Brazilian strains (S25, SA20, SA623, and S13), the few distinct evolutionary features.

Chinese isolate (GD201008-001) conserved a very few distinct evolutionary features.

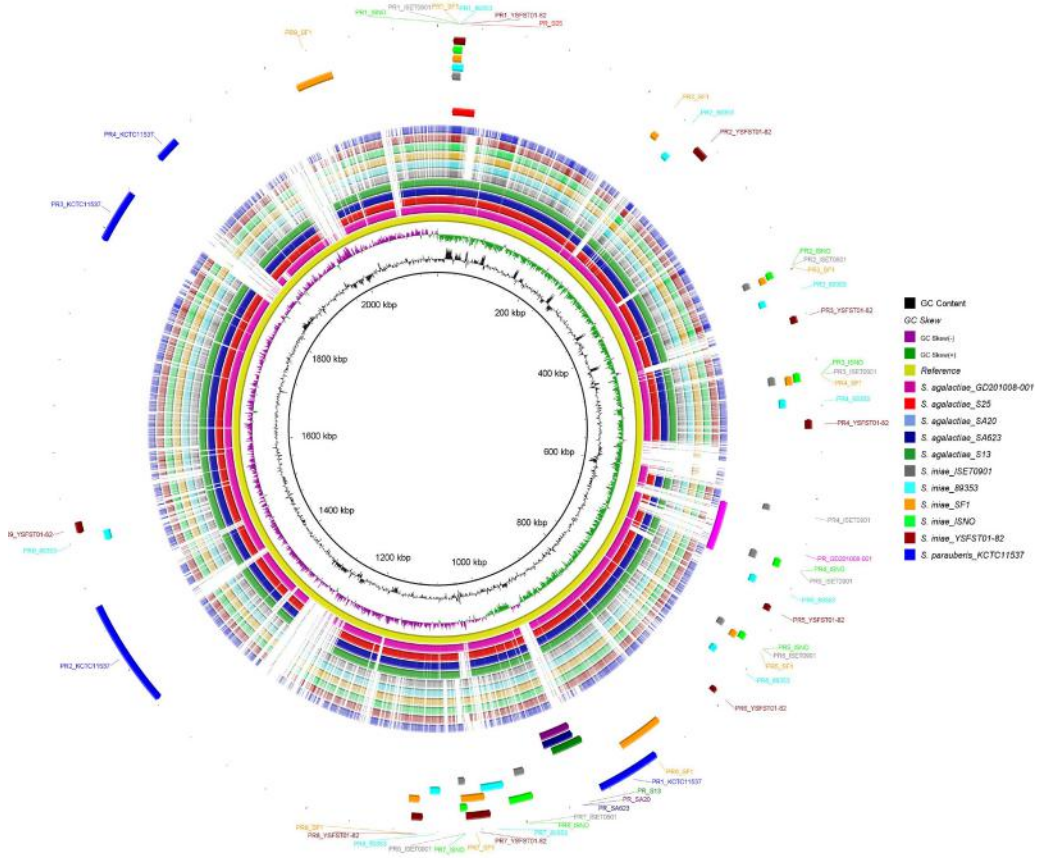

Fig.1. The BLAST genome circular map visualization of the studied 11 strains of *Streptococcus* spp. *S. agalactiae* strain 2603V/R (NCBI accession No. AE009948) was used as a reference. The innermost two circles represent the GC content (black) and the third ring shows GC skew (purple/green) of the reference strain. The remaining circles (4 to 15) represent the BLASTn search of the complete genome sequence of 11 *Streptococcus* spp. against the complete genome of the reference strain. The different colours arranged from the inner to the outer ring were as follows: reference strain *S. agalactiae* 2603V/R (yellow); *S. agalactiae* strains GD201008-001 (pink), S25 (red), SA20 (purple), SA623 (Navy blue) and S13 (green); *S. iniae* strains ISET0901 (grey), 89353 (aqua), SF1 (orange), ISNO (lime) and YSFST01-82 (maroon) and *S. parauberis* KCTC11537 (blue). The BLASTn identities were performed on the basis of colour intensity, where high nucleotide identities were with dark regions and light regions were as little identity or no nucleotide identity. The outermost 11 rings show the prophage regions (PR) of 11 strains. The prophage ring colours of each strain were similar to the respective draft genome sequence ring colour. Five prophage regions of five strains of *S. agalactiae* were denoted as PR_ GD201008- 001, PR_S25, PR_SA20, PR_SA623, and PR_S13. The prophages of five strains of *S. iniae* were as follows: 89353: PR1_89353 to PR9_89353, ISET0901: PR1_ ISET0901 to PR8_ ISET0901, SF1: PR1_ SF1 to PR9_ SF1, ISNO: PR1_ ISNO to PR7_ ISNO, and YSFST01-82: PR1_ YSFST01-82 to PR9_ YSFST01-82. Finally, the outer blue coloured ring shows four PRs of *S. parauberis* KCTC11537: PR1_ KCTC11537 to PR4_KCTC11537.

Among the five strains of *S. iniae* three isolates obtained from East Asian countries conserved 9 prophage regions in their DNA sequences; whereas, South Korean strain *S. parauberis* KCTC11537 harbors four phage areas. However, it appeared that there might be a species and some extent of host specificity in case of conserve phage region as all isolates of *S. iniae* conserved multiple prophages and all five isolates of *S. agalactiae* had single prophage region. In the case of the streptococcal strains isolated from flounder fish, all *Streptococcus* sp. conserved multiple prophage regions in the present study. Interestingly, *S. iniae* strain SF1 and *S. parauberis* strain KCTC11537 were isolated from the host flounder and contained intact phage regions.

Virulence associated gene profiles

Among the 11 *Streptococcus* sp. strains, 44 virulence genes were identified using VFDB database (Fig. 2A). They had different virulence functions i. e. adherence, enzyme production (i.e., protease), immune evasion, immunoreactive antigen, and toxin. Among the identified genes, 7 virulence genes such as Fibronectin-binding proteins (*fbp54*), Streptococcal plasmin receptor (*plr/gapA*), Hyaluronidase *(hylB)*, Streptococcal enolase *(eno)*, C3-degrading protease (*cppA*), Serine protease (*htrA/degP*) and Trigger factor (*tig*/*ropA*) were identical in all 11 strains (Fig. 2A). Parallel to our study, fibrinogen-binding protein *fbp54* was identified from human pathogen *S. pyogenes (Courtney et al., 1996) and S. intermedius* (Issa *et al.*, 2019) *and* pigs pathogen *S. suis* domestic pigs *(Park et al., 2021)*. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) encoded gene (*plr*/*gap*) responsible for plasma receptor was also found in *S. agalactiae* (Sharma *et al.*, 2013), *S. sanguinis* and *S. gordonii* (Iversen *et al.*, 2020) and *S. suis* INT-01 *(Park et al., 2021).* Furthermore, enolase is a putative virulence protein that binds to plasminogen and contributes to infectious processes. Temperature and oxidative stresses tolerance gene *HtrA* (*DegP*) is an important virulence gene for the human pathogen *S. pyogenes* that helps in the maturation of cysteine protease *SpeB* (Cole *et al.*, 2007). Virulence genes encoded for extracellular enzyme (*hylB* and *eno*) and protease synthesis (*cppA*, *htrA*/*degP*) were found in the genome sequences of fish-derived strains along with human pathogens of *S. agalactiae* (Kayansamruaj *et al.*, 2015). Furthermore, C3 protease (*cppA*) was found in the pathogenic *S. iniae* (Nguyen, 2015). *Trigger factor (tig/ropA) act as a* chaperone and is involved in protein exportation and *ropA* is an essential factor for the secretion and maturation of the cysteine protease (Lyon & Caparon, 2003). The stress tolerance trigger factor (*tig*) and *ropA* were found in the genome of *S. suis* gene (Wu *et al.*, 2011) and *S. pyogenes* (Lyon & Caparon, 2003), respectively which was in the same line of the present study.

Similarly, the capsular polysaccharide (*CPS*) biosynthesis locus conserved 12 genes (*cps4A* to *cps4L*) in all 11 strains. Consistent with our results, the *cps* were also found to be conserved by other bacteria such as *S. iniae* (Lowe *et al.*, 2007), *S. agalactiae* (Toniolo *et al.*, 2015), *S. pneumoniae* (Shainheit *et al.*, 2014) and *S. suis* (Smith *et al.*, 1999).

The pneumococcal surface antigen-A encoding gene *psaA* and CAMP factor-like gene *cfa*/*cfb* were identical in ten strains of *S. agalactiae* and *S. iniae.* Similar to this study, Kayansamruaj *et al.*, (2015) reported the metal transportation (*psaA*) and toxin synthesis (*cfb*) genes were isolated from fish pathogenic *S. agalactiae*. The same results were also reported in case of the putative virulence gene *psaA* isolated from human pathogen *S. pneumoniae* (Gor *et al.*, 2005; Johnston *et al.*, 2004). On the other hand, the potential streptococcal receptor Laminin-binding protein (*lmb*) and Endoglycosidase (*endoS*) were common in six strains of *S. iniae* and *S. parauberis*. Interestingly, several studies have identified *lmb* gene from *S. agalactaie* (Kannika *et al.*, 2017; Zhang *et al.*, 2018). Moreover, *lmb* gene was also identified from *S. oriscaviae* (Teng *et al.*, 2022) and *S. uberis* (Vezina *et al.*, 2021), and *S. parauberis* (Lee *et al.*, 2021). However, VFDB analyzer in the current study did not identify any *lmb* gene from *S. agalactiae*. The use of more web-based tools could have brought different results in this case.

Furthermore, Mitogenic factor 2 (*mf2*), Neuraminidase A (*nanA*), and C5a peptidase (*scpA* and *scpB*) were found in the five strains of *S. iniae*. Close to our results, the Mitogenic factor 2 (*mf2*) was also found in *S. pyogenes* (Ferretti *et al.*, 2001;

Hasegawa *et al.*, 2002). The C5a peptidase *Scp* (Streptococcal C5a peptidase) is a serine proteases and an important virulence protein in Group A streptococci that encoded the *Scp*-like gene (Chen & Cleary, 1990). It was found that five strains of S. *iniae* harbor candidate *Scp*-like gene *scpA* and *scpB*. Similar to the present study, C5a peptidase encoding gene *scpA* was identified from *S. dysgalactiae* (McKenna *et al.*, 2022) and *S. pyogenes* (Chen & Cleary, 1990).

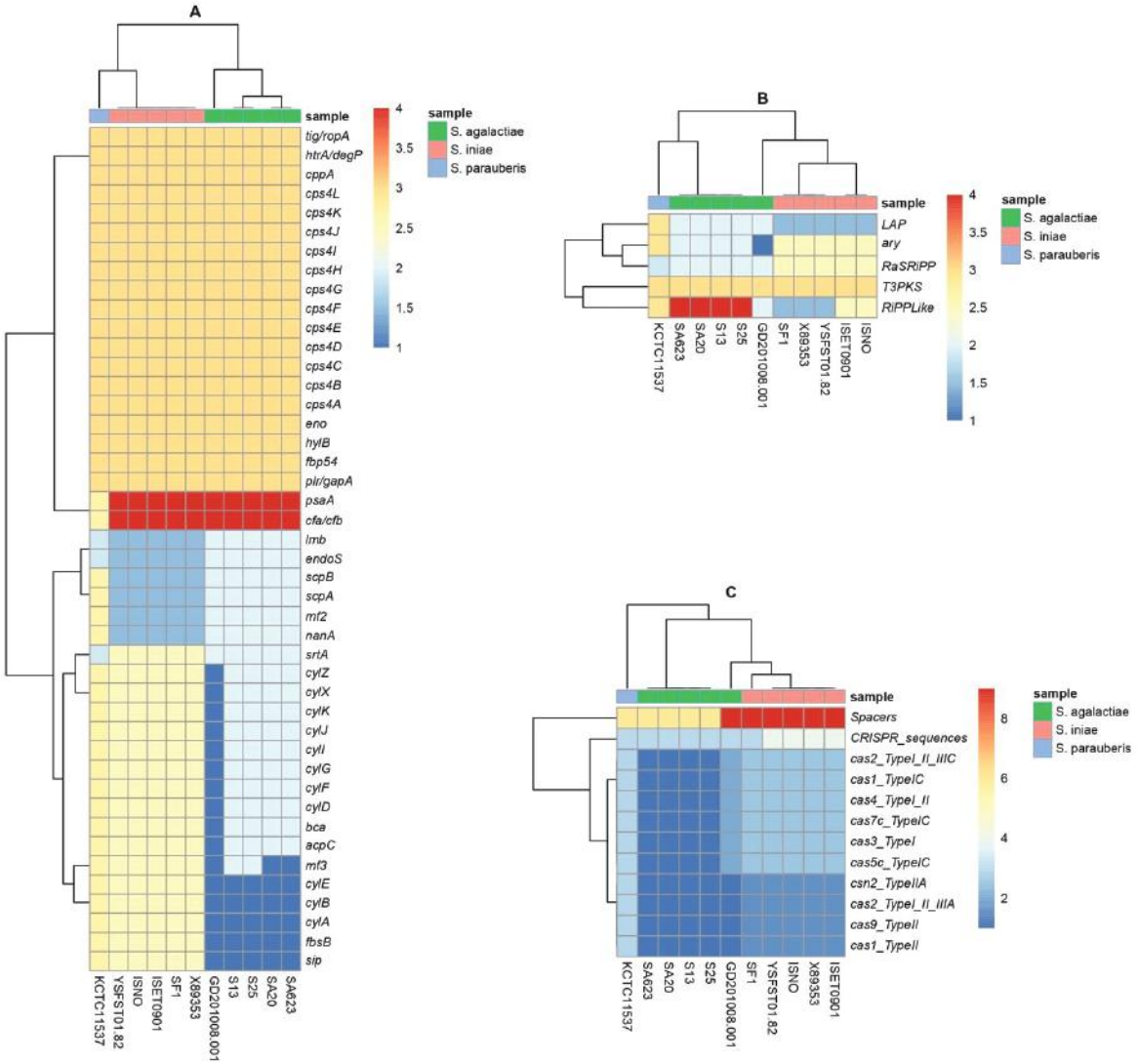

Fig. 2. An overview of genes identified from the genome sequence of 11 strains of fish pathogenic *Streptococcus*. Here, sample means the type of three Streptococcus spp. The sample Green, pink, and sky colours indicated *S. agalactiae, S. iniae*, and *S. parauberis*, respectively. A. Represented the virulence gene profile. In the colour bar, 1 to 1.9 means the presence of gene, 2 to 2.9 means the absence of gene, 3 means identical gene in 11 strains of *Streptococcus* and 4 means common gene among ten strains *S. agalactiae a*nd *S. iniae* B. Secondary metabolites (SM) obtained by antiSMASH 5.0 (Medema *et al*., 2011). The colour bar, 1 to 1.9 means presence of SM, 2 to 2.9 indicate absence of SM, 3 represent identical SM in all strains and 4 means present of SM in two area of the genome C. CRISPR-Cas obtained from CRISPRCasFinder tool (Couvin *et al.*, 2018). The color bar, 1 to 1.9 represents present of cas gene, 2 to 2.9 means absence of cas gene, 3 and 4 mean single and double CRISPR sequence, respectively; 5 and 6 mean single and multiple spacers present on the CRISPR sequences, respectively.

Five strains of *S. agalactiae* harbored fibronectin-binding proteins (*fbsB)* and Surface immunogenic protein (*sip*) which were absent in the other six strains of *Streptococcus* belonging to *S. iniae* and *S. parauberis* (Fig. 2A). In this connection similar results were reported by other studies (Kayansamruaj et al, 2015; Gor et al, 2005).

There were 11 *cyl* genes (*cyl A, cylB, cylD, cylE, cylF, cylG, cylI, cylJ, cylK, cylX,* and *cylZ*) identified from *S. agalactiae* which are required for the production of beta hemolysin/cytolysin (Fig. 2A). Although all of *cyl* genes were identified from the Chinese strain GD201008-001 of *S. agalactiae*, only three genes (*cylA, cylB,* and *cylE*)) were present in other four Brazilian strains (S25, SA20, SA623, and S13). Virulence gene *mf3* encoding mitogenic factor 3 was identified in *S. agalactiae* strains GD201008-001, SA20, and SA623. The gene encoding beta C protein (*bca*) and synthesis of sortase A gene (*srtA*) were found in the genome sequence of Chinese *S. agalactiae* strain GD201008-001 and Korean *S. parauberis* strain KCTC11537, respectively (Fig. 2A). The *cyl* gene was also identified from *S. agalactiae* (Kayansamruaj *et al.*, 2015; Shimizu *et al.*, 2020). Beta C protein encoded gene (*bca*), and synthesis of sortase-A gene (*srtA)* were found in the genome sequence of *S. agalactiae* strain GD201008-001 and *S. parauberis* strain KCTC11537. Similar genes were reported from *S. agalactiae* (Bobadilla *et al.*, 2021; Kayansamruaj *et al.*, 2015).

Antibiotic resistance genes (ARGs)

Eight ARGs were detected by using three approaches (ResFinder, CARD and ARG-ANNOT) on the genome sequences of *Streptococcus* sp. (Table 3). According to the ResFinder, a common putative ARG including *mreA* was identified from five strains of *S. agalactiae* that were resistant to the macrolide group of antibiotics such as erythromycin, azithromycin, and spiramycin. Although, ResFinder did not identify any ARG from *S. iniae*, The CARD database identified macrolide resistance gene *mreA* in the genome sequence of five strains of *S. iniae* with 73.2 % of identity according to the matching region and 99.68% of reference length sequence.

Table 3: Antibiotic resistance genes identified in the genome sequences of 11 strains of *Streptococcus* spp.

Name of genes	Bacteria Species											Drug class
	S. agalactiae					*S. iniae*					*S. parab aeruis*	
	GD 201008-001	SA 623	S 25	SA 20	S 13	89353	ISET 0901	ISNO	YSFST 01-82	SF1	KCTC 11537	
[c,r] *mreA*	+	+	+	+	+	+	+	+	+	+	-	Macrolide antibiotic (erythromycin, azithromycin and, spiramycin)
[c] *mprF*	+	+	+	+	+	-	-	-	-	-	-	Peptide antibiotic
[c] *vanY* gene in *vanB* cluster	+	+	+	+	+	+	+	+	+	+	-	Glycopeptide antibiotic
[c] *vanT* gene in *vanG* cluster	+	+	+	+	+	-	-	-	-	-	-	Glycopeptide antibiotic
[c] *vanY* gene in *vanF* cluster	-	-	-	-	-	+	+	+	+	+	+	Glycopeptide antibiotic
[c] *vanY* gene in *vanM* cluster	-	-	-	-	-	-	-	-	-	-	+	Glycopeptide antibiotic
[c] *patB*	-	-	-	-	-	+	+	+	+	+	+	Fluoroquinolone antibiotic
[c] *qacJ*	-	-	-	-	-	-	-	-	-	-	+	Disinfecting agents and antiseptics

Note: Here [c, r] mean CARD and ResFinder, respectively. Symbols + and – mean presence and absence of gene, respectively.

Similarly, the macrolide resistance *mreA* gene was reported in other clinical isolates of *S. agalactiae* (Clancy *et al.*, 1997; Clarebout *et al.*, 2001; Vidal Amaral *et al.*, 2022). A study reported that the macrolide resistance efflux *mreA* gene was a resident in the chromosome of *S. agalactiae* and was responsible for multiple metabolic functions (Clarebout *et al.*, 2001). Resemble to the current study, macrolide resistance M like genes were also found in other streptococci such as *msrA*, *mefE*, and *mefA* from *S. epidermidis* (Ross *et al.*, 1990), *S. pneumonia* (Sutcliffe *et al.*, 1996) and *S. pyogenes* (Clancy *et al.*, 1997), respectively.

Vancomycin is a widely used glycopeptide antibiotic to cure infectious diseases against gram positive bacteria. Several vancomycin resistant genes (*van*) were discovered from different types of organisms such as *vanA*, *vanB*, *vanC*, *vanD*, *vanE*, *vanM*, *vanG*, *vanX*, *vanY* etc (Khan *et al.*, 2008; Li *et al.*, 2022). Likewise, one glycopeptide ARG including *vanY* gene in *vanB* cluster was found in all of the studied strains of *S. agalactiae* and *S. iniae* through CARD analysis. Another glycopeptide ARG *vanY* gene in *vanF* cluster was detected on six genome sequences of *S. iniae* and *S. parauberis* (Table 3). Moreover, two ARGs such as the glycopeptide encoding *vanT* gene in *vanG* cluster and peptide encoding *vanT* gene were identical to the genome sequences of five strains of *S. agalactiae* (Table 3). The glycopeptide ARG *vanY* gene in *vanM* cluster and disinfectants and antiseptics resistant gene *qacJ* were only harbored by the genome sequence of *S. parauberis* KCTC11537 (Table 3). Similar to the present results, several vancomycin resistance van gene were found in different streptococcal species like *S. bovis* (*vanB*) (Poyart *et al.*, 1997), *vanG* in *S. agalactiae,* and *S. anginosus* (Srinivasan *et al.*, 2014). Efflux proteins encoded gene *patB* was identified in five *S. iniae,* and *S. parauberis* strains which belong to the ABC transporter family. In the same line of the current results, fluoroquinolone resistance gene *patB* was also found in *S. uberis* (Hassan *et al.*, 2022). Five study strains of *S. agalactiae* conserved *mprF* gene that was also found in the other study with *S. agalactiae* and *Staphylococcus aureus* (Oku *et al.*, 2004; Vidal Amaral *et al.*, 2022). The quaternary ammonium compounds (QACs) are known as disinfectants and antiseptics and they are widely used in human medicine and food industries. The quaternary ammonium compounds resistance gene *qacJ* was located in the plasmids staphylococcal species (Bjorland *et al.*, 2003). Like staphylococcal bacteria, disinfectants and antiseptics resistance gene *qacJ* gene was also found in the *S. parauberis*. However, ARG-ANNOT database was not able to identify any ARGs in the genome sequences of 11 bacterial strains in the current study.

Secondary metabolites

Secondary metabolites are produced from different biosynthetic pathways of pathogens and play an important role in different virulence activities. A total of five types of secondary metabolites were detected, including Arylpolyene (*ary*), Type III Polyketide synthases (*T3PKS*), ribosomally synthesized and Lanthipeptides that were ribosomally synthesized and post-translationally modified peptide (RiPP-like), linear azol(in)e-containing peptides (*LAPs*), and *RaS-RiPP* antimicrobial compound from the study bacteria of *Streptococcus* (Fig. 2B).

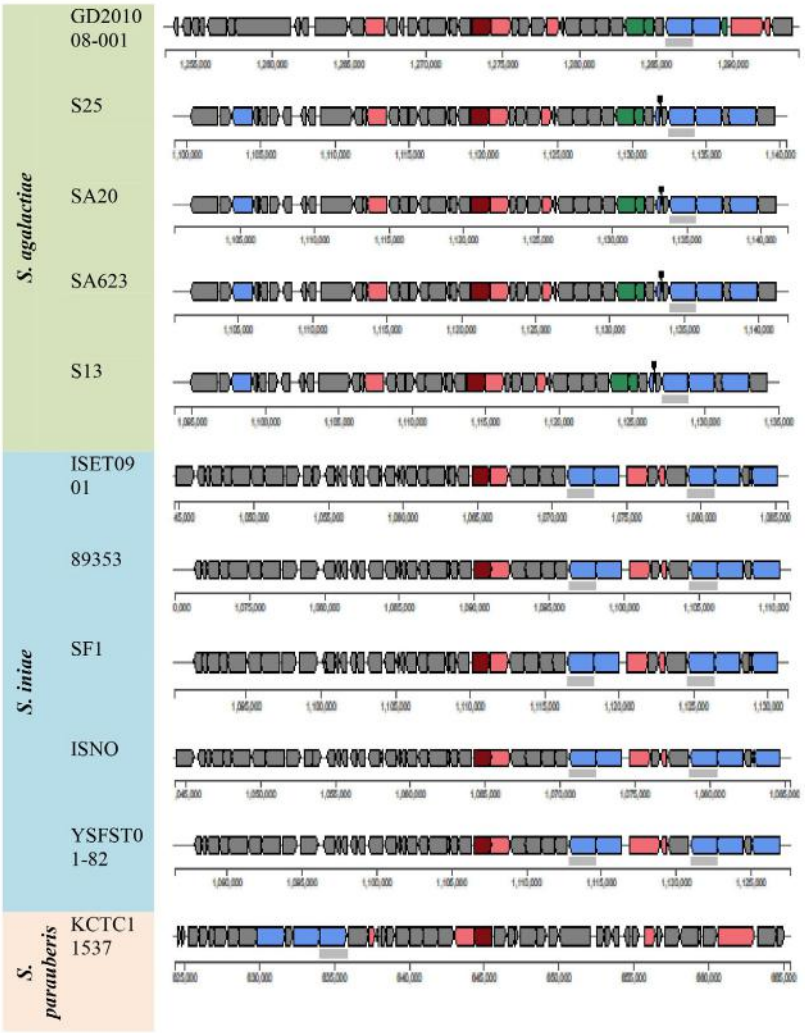

Fig. 3. Comparison of the secondary metabolites *T3PKS* biosynthetic gene clusters among the eleven fish pathogenic streptococcal strains of three species such as *S. agalactiae* (GD201008-001, S25, SA20, SA623, and S13), *S. iniae* (ISET0901, 89353, SF1, INSO, and YSFST01-82) and *S. parauberis* (KCTC11537). The arrows represent the direction of each gene's transcription. Here different functional genes are identified with different colours such as dark red indicates core biosynthetic genes; pink, blue, green, and grey colour represented additional biosynthetic genes, transport-related genes, regulatory genes, and other genes, respectively. Secondary metabolite *T3PKS* producing region in the genomes of *S. agalactiae, S. iniae,* and *S. parauberis* were identified with the antiSMASH 5.0 web tool (Medema *et al.*, 2011).

Biosynthetic gene cluster *T3PKS* was found in all eleven strains of *Streptococcus* sp. (Fig. 2B). According to the species of bacteria, position and content of functional genes were identical but core biosynthesis genes were unique in all eleven strains in *T3PKS* gene cluster (Fig. 3). Arylpolyene (*ary*) gene was detected only from the Chinese strain *S. agalactiae GD201008-001* responsible for biofilm formation and protection from oxidative stress (Fig. 2B). On the other hand, each of the Brazilian strain of *S. agalactiae* conserved two sequences encoded for RiPP-like secondary metabolites (Fig. 2B). Similarly, three strains of *S. iniae* such as 89353, YSFST01-82, and YSFST01-82 harbored a single sequence responsible for RiPP-like secondary metabolites (Fig. 2B). The *LAPs* encoded gene sequence similar to streptolysin S was obtained from five strains of *S. iniae*. *RaS- RiPP* gene clusters that can produce peptides involved in the control of a quorum sensing (QS) system in the genomes of

Streptococci. Although *RaS-RiPP* gene was found in *S. parauberis* strain KCTC11537, it was absent in other two species of *Streptococcus* (Fig. 2B).

The *aryl* polyenes involve bacterial pigments and act as carotenoids to protect bacteria against reactive oxygen species (Schöner *et al.*, 2016). The *ary* gene was also found in other strains of *S. agalactiae* isolated from Bovine Mastitis (Vidal Amaral *et al.*, 2022). Conversely, Microbial type III PKSs are involved in the biosynthesis of numerous secondary metabolites and lipid compounds that have significant biological functions and important pharmaceutical activities (Katsuyama & Ohnishi, 2012). The lanthipeptide is a noteworthy family of RiPP with lanthionine amino acid in their structure. Ripp-like products were also found in the *Enterococcus* sp. YC2–6 genome (Okoye *et al.*, 2022). Certain pathogenic bacteria of streptococci conserved a factor called linear azol(in) e-containing peptides (*LAPs*). The *LAPs* are constructed with a combination of thiazole and (methyl) oxazole heterocycles. A known *LAP* including streptolysin S was found as an integral factor of the pathogenic mechanism of *S. pyogenes (Letzel et al., 2014; Nizet et al., 2000),* as reported in the current results. Radical S-adenosylmethionine (*RaS*) is an enzyme that is involved in RiPP biosynthesis in mammalian Streptococci and develops a *RaS-RiPPs* enzyme network (Clark *et al.*, 2022). This enzyme network produces redox-active cofactors such as pyrroloquinoline quinone usually found in different gram-negative bacteria including *Klebsiella pneumonia* (Clark *et al.*, 2022). Therefore, *RaS-RiPP* may involve in the virulence of *Streptococcus.*

CRISPR/CRISPR-Cas analysis

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and their associated *Cas* protein were investigated using the web tool CRISPRCasFinder among the genome sequence of 11 streptococcal strains (Fig. 2C). We identified two types of *Cas* clusters such as *CAS-TypeIC,* and *CAS-TypeIIA* among the ten strains of *S. agalactiae,* and *S. iniae* sequences. Except *S. parauberis*, all of the studied streptococcal strains harbored the *CAS-TypeIIA* loci which consisted of four types of *Cas* genes including *cas9_TypeII*, *cas1_TypeII*, *cas2_TypeI-II-III,* and *csn2_TypeIIA*. The four Brazilian strains of *S. agalactiae* (S25, SA20, SA623, and S13) consisted *CAS-TypeIC Cas* locus possess six types of *Cas* protein including *cas3_TypeI*, *cas5c_TypeIC*, *cas7c_TypeIC*, *cas4_TypeI-II*, *cas1_TypeIC,* and *cas2_TypeI-II-III* (Fig. 2C).

Although one CRISPR sequence was detected from five strains of *S. agalactiae*, all strains of *S. iniae* conserved 2 CRISPR sequences with spacer numbers varied from 1 to 12. On the other hand, one CRISPR sequence was mined from the genome sequence of *S. parauberis KCTC11537* and no *Cas* loci was detected. The presences of CRISPR-Cas systems represent an adaptive immunity mechanism against the mobile genetic elements (MGEs) in bacteria (Lopez-Sanchez *et al.*, 2012) and contributes to the diversity of MGEs in the population. The strains in the current study conserved CRISPR-Cas system with several spacers' which may represent to the diversity of MGEs in the population.

Single nucleotide polymorphism (SNP) analysis

A phylogenetic tree based on the concatenated SNPs was constructed (Fig. 4). The CSI phylogeny pipeline identified maximum 792 SNPs positions in the shared core genome of the isolates in this study. The phylogenetic analysis revealed that the *Streptococcus* strains clustered according to bacterial species and their geographical origin. Three strains *S. agalactiae* that were isolated from Brazil, grouped into a common clade. They also form branches very closely with another two strains of *S. agalactiae* (SA623, and GD201008-001). Although *S. iniae* strains ISET0901 and ISNO were isolated from two different geographical origins (Israel and USA), no genetic diversity was observed between them; rather, they shared a common branch according to species.

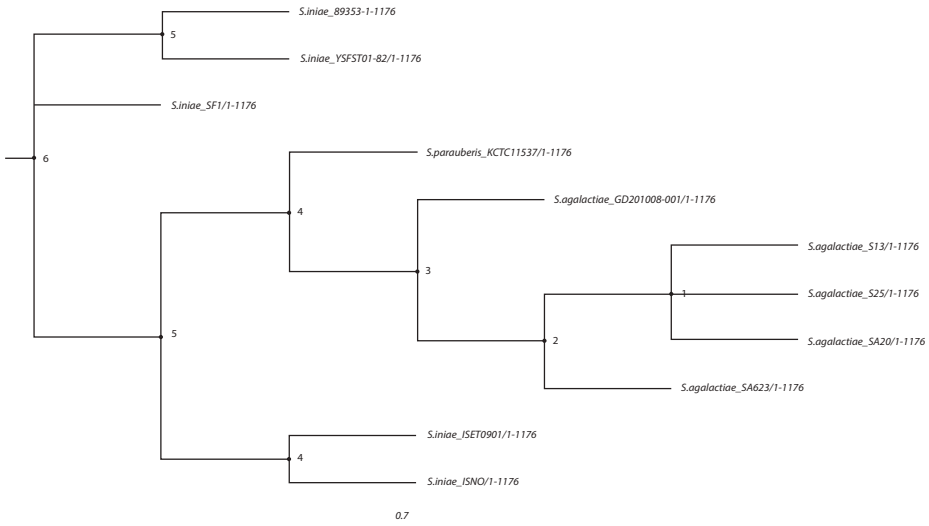

Fig. 4. The Phylogenetic tree based on SNP analysis of fish pathogenic 11 strains of *Streptococcus* sp. The tree was generated by CSIphylogeny v1.4 (Kaas *et al.*, 2014) and modified with Fig.Tree v1.3. The scale bar indicates the numbers of substitution per site.

Two strains *S. iniae* such as 89353 and YSFST01-82 clustered in a same group according to geographical origin and species and it was distinctly separated from the branch that contains Chinese strain *S. iniae* SF1. Another species *S. parauberis* KCTC3651 found separately from the other two species of *S. iniae* and *S. agalactiae*. However, from the tree topology, a significant evolutionary diversification was observed according to streptococcal species and a large SNPs difference was found among all of the isolates.

Conclusion

The current study is an *in-silico*-based comparative analysis of whole genome sequence of the fish pathogenic three streptococcal species. The results reveal that the pathogens are almost identical according to species and geographical area. Although few variations are noted from Chinese *S. agalactiae* to other four Brazilian strains in terms of secondary metabolites content and CRISPR *Cas* analysis, they are unique for virulence, ARGs and phage regions contents. Genome sequence analysis of five strains of *S. iniae* is very identical without any variation. The genome sequence of another strain of *S. parauberis* has very few similarities with the other two *Streptococcus* sp. Few virulence genes are common among all three species. In the case of phage detection analysis, species and host specificity were observed in case of *S, agalactiae,* and *S. iniae*. Only intact prophage was found in the genome sequences of *S. iniae* strain SF1, and *S. parauberis* KCTC3651 that were obtained from flounder. In the phylogenetic analysis, East Asian three strains of S. *iniae* such as 89353, YSFST01-82, and SF1 form close cluster and other two strains *S. iniae* such as ISET0901, and ISNO isolated from tilapia form group in one clade. This variation might be due to the differences in geographical locations and some extent of host specificity. Nonetheless, this *in-silico* comparative analysis of genome sequences of three species reveal the pathogenicity of *streptococcus* sp. would be closer in causing streptococcosis regardless of the origin, distribution and host.

Conflict of interest: None

Ethics approval: Note applicable

Acknowledgement: None

References

Akter, T., M. J. Foysal, M. Alam, R. Ehsan, S. I. Paul, F. Momtaz, M. A. Siddik, A. C. Y. Tay, R. Fotedar, S. K. Gupta, T. Islam and M. M. Rahman. 2021. Involvement of *Enterococcus* species in streptococcosis of Nile tilapia in Bangladesh. *Aquaculture*. 531: 735790. https://doi.org/10.1016/j.aquaculture.2020.735790

Akter, T., M. N. Haque, R. Ehsan, S. I. Paul, M. J. Foysal, A. C. Y. Tay, T. Islam and M. M. Rahman. 2023. Virulence and antibiotic-resistance genes in *Enterococcus faecalis* associated with streptococcosis disease in fish. *Sci. Rep. 13*(1): 1551.

Alikhan, N. F., N. K. Petty, N. L. Ben Zakour and S. A. Beatson. 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genom.* 12: 402. https://doi.org/10.1186/1471-2164-12-402

Arndt, D., J. R. Grant, A. Marcu, T. Sajed, A. Pon, Y. Liang and D. S. Wishart. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res*. 44(W1): W16-21. https://doi.org/10.1093/nar/gkw387

Bjorland, J., T. Steinum, M. Sunde, S. Waage and E. Heir. 2003. Novel plasmid-borne gene *qacJ* mediates resistance to quaternary ammonium compounds in equine *Staphylococcus aureus*, *Staphylococcus simulans*, and *Staphylococcus intermedius*. *Antimicrob. Agents Chemother.* 47(10): 3046-3052.

Bobadilla, F. J., M. G. Novosak, I. J. Cortese, O. D. Delgado and M. E. Laczeski. 2021. Prevalence, serotypes and virulence genes of *Streptococcus agalactiae* isolated from pregnant women with 35–37 weeks of gestation. *BMC Infect. Dis.* 21(1): 1-11.

Carattoli, A., E. Zankari, A. Garcia-Fernandez, M. V. Larsen, O. Lund, L. Villa, F. Møller Aarestrup and H. Hasman. 2014. In Silico Detection and Typing of Plasmids using PlasmidFinder and Plasmid Multilocus Sequence Typing. *Antimicrob. Agents Chemother.* 58(7): 3895-3903. https://doi.org/10.1128/Aac.02412-14

Chen, C. C. and P. Cleary. 1990. Complete nucleotide sequence of the streptococcal C5a peptidase gene of *Streptococcus pyogenes*. *J. Biol. Chem.* 265(6): 3161-3167.

Clancy, J., F. Dib-Hajj, J. W. Petitpas and W. Yuan. 1997. Cloning and characterization of a novel macrolide efflux gene, *mreA*, from *Streptococcus agalactiae*. *Antimicrob. Agents Chemother.* 41(12): 2719-2723.

Clarebout, G., C. Villers, C. and R. Leclercq. 2001. Macrolide resistance gene *mreA* of S*treptococcus agalactiae* encodes a flavokinase. *Antimicrob. Agents Chemother.* 45(8): 2280-2286.

Clark, K. A., L. B. Bushin and M. R. Seyedsayamdost. 2022. RaS-RiPPs in Streptococci and the human microbiome. *ACS bio & med. Chem. Au*. 2(4): 328-339.

Cole, J. N., J. A. Aquilina, P. G. Hains, A. Henningham, K. S. Sriprakash, M. G. Caparon, V. Nizet, M. Kotb, S. J. Cordwell, S. P. Djordjevic and M. J. Walker. 2007. Role of group A *Streptococcus HtrA* in the maturation of *SpeB* protease. *Proteomics*. 7(24): 4488-4498.

Courtney, H. S., J. B. Dale and D. Hasty. 1996. Differential effects of the streptococcal fibronectin-binding protein, *FBP54*, on adhesion of group A streptococci to human buccal cells and HEp-2 tissue culture cells. *Infect. Immun.* 64(7): 2415-2419.

Couvin, D., A. Bernheim, C. Toffano-Nioche, M. Touchon, J. Michalik, B. Néron, E. P. Rocha, G. Vergnaud, D. Gautheret and C. Pourcel. 2018. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* 46(W1): W246-W251.

de Aguiar, E. L., D. C. B. Mariano, M. V. C. Viana, L. D. J. Benevides, F. de Souza Rocha, L. de Castro Oliveira, F. L. Pereira, F. A. Dorella, C. A. G. Leal, A. F. de Carvalho and G. S. Santos. 2016. Complete genome sequence of *Streptococcus agalactiae* strain GBS85147 serotype of type Ia isolated from human oropharynx. *Stand. Genom. Sci.* 11(1): 1-8.

Facimoto, C. T., R. T. Chideroli, D. D. Goncalves, A. O. D. Carmo, E. Kalaphotakis and U. P. Pereira. 2017. Whole-Genome Sequence of *Streptococcus agalactiae* Strain S13, Isolated from a Fish Eye from a Nile Tilapia Farm in Southern Brazil. *Genome. Announc*. 5(35): e00917-00917. https://doi.org/10.1128/genomeA.00917-17

Ferretti, J. J., W. M. McShan, D. Ajdic, D. J. Savic, G. Savic, K. Lyon, C. Primeaux, S. Sezate, A. N. Suvorov, S. Kenton and H. S. Lai. 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes. Proc. Natl. Acad. Sci.* 98(8): 4658-4663.

Gor, D. O., X. Ding, D. E. Briles, M. R. Jacobs and N. S. Greenspan. 2005. Relationship between surface accessibility for *PpmA*, *PsaA*, and *PspA* and antibody-mediated immunity to systemic infection by *Streptococcus pneumoniae*. *Infect. Immun.* 73(3): 1304-1312.

Gupta, S. K., B. R. Padmanabhan, S. M. Diene, R. Lopez-Rojas, M. Kempf, L. Landraud and J. M. Rolain. 2014. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob. Agents. Chemother*. 58(1): 212-220. https://doi.org/10.1128/AAC.01310-13

Hasegawa, T., K. Torii, S. Hashikawa, Y. Iinuma and M. Ohta. 2002. Cloning and characterization of two novel DNases from *Streptococcus pyogenes*. *Arch. Microbiol.* 177: 451-456.

Hassan, J., M. A. S. Bag, A. Kabir, M. W. Ali, M. Hossain, M. T. Rahman, M.S. Islam and M.S.R. Khan. 2022. Genomic diversity of *Streptococcus uberis* isolated from clinical mastitis of cattle in selected areas of Bangladesh. *bioRxiv*. pp. 2022-12.

Issa, E., T. Salloum, B. Panossian, D. Ayoub, E. Abboud and S. Tokajian. 2019. Genome mining and comparative analysis of *Streptococcus intermedius* causing brain abscess in a child. *Pathogens.* 8(1): 22.

Iversen, K. H., L. H. Rasmussen, K. Al-Nakeeb, J. J. A. Armenteros, C. S. Jensen, R. Dargis, R., O. Lukjancenko, U. S. Justesen, C. Moser, F. S. Rosenvinge and X. C. Nielsen. 2020. Similar genomic patterns of clinical infective endocarditis and oral isolates of *Streptococcus sanguinis* and *Streptococcus gordonii*. *Sci. Rep.* 10(1): 2728.

Johnston, J. W., L. E. Myers, M. M. Ochs, W. H. Benjamin Jr, D. E. Briles and S. K. Hollingshead. 2004. Lipoprotein *PsaA* in virulence of *Streptococcus pneumoniae*: surface accessibility and role in protection from superoxide. *Infect. Immun.* 72(10): 5858-5867.

Kaas, R. S., P. Leekitcharoenphon, F. M. Aarestrup and O. Lund. 2014. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *Plos One*. 9(8): e104984. https://doi.org/10.1371/journal.pone.0104984

Kannika, K., D. Pisuttharachai, P. Srisapoome, J. Wongtavatchai, H. Kondo, I. Hirono, S. Unajak and N. Areechon. 2017. Molecular serotyping, virulence gene profiling and pathogenicity of *Streptococcus agalactiae* isolated from tilapia farms in Thailand by multiplex PCR. *J. Appl. Microbiol.* 122(6): 1497-1507.

Katsuyama, Y. and Y. Ohnishi. 2012. Type III polyketide synthases in microorganisms. *Meth. Enzymol.* (Vol. 515, pp. 359-377). Elsevier.

Kayansamruaj, P., N. Pirarat, H. Kondo, I. Hirono and C. Rodkhum. 2015. Genomic comparison between pathogenic *Streptococcus*

agalactiae isolated from Nile tilapia in Thailand and fish-derived ST7 strains. *Infect. Genet. Evol.* 36: 307-314. https://doi.org/10.1016/j.meegid.2015.10.009

Khan, M. A., M. van der Wal, D. J. Farrell, L. Cossins, A. van Belkum, A. Alaidan and J. P. Hays. 2008. Analysis of *VanA* vancomycin-resistant *Enterococcus faecium* isolates from Saudi Arabian hospitals reveals the presence of clonal cluster 17 and two new Tn 1546 lineage types. *J Antimicrob. Chemother.* 62(2): 279-283.

Lee, Y., N. Kim, H. Roh, A. Kim, H. J. Han, M. Cho and D. H. Kim. 2021. Transcriptome analysis unveils survival strategies of *Streptococcus parauberis* against fish serum. *Plos One*. 16(5): e0252200.

Legario, F. S., C. H. Choresca Jr, J. F. Turnbull and M. Crumlish. 2020. Isolation and molecular characterization of streptococcal species recovered from clinical infections in farmed Nile tilapia (*Oreochromis niloticus*) in the Philippines. *J. Fish Dis.* 43(11): 1431-1442.

Letzel, A. C., S. J. Pidot and C. Hertweck. 2014. Genome mining for ribosomally synthesized and post-translationally modified peptides (RiPPs) in anaerobic bacteria. *BMC Genom.* 15(1): 1-21.

Li, G., M. J. Walker and D. M. De Oliveira. 2022. Vancomycin resistance in enterococcus and *Staphylococcus aureus*. *Microorganisms*. 11(1): 24.

Liu, B., D. Zheng, Q. Jin, L. Chen and J. Yang. 2019. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* 47: D687-D692. https://doi.org/10.1093/nar/gky1080

Liu, G., W. Zhang and C. Lu. 2013. Identification of immunoreactive proteins of *Streptococcus agalactiae* isolated from cultured tilapia in China. *Pathog. Dis*, 69(3): 223-231. https://doi.org/10.1111/2049-632X.12084

Lopez-Sanchez, M. J., E. Sauvage, V. Da Cunha, D. Clermont, E. Ratsima Hariniaina, B. Gonzalez-Zorn, C. Poyart, I. Rosinski-Chupin and P. Glaser. 2012. The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. *Mol. Microbiol.* 85(6): 1057-1071.

Lowe, B. A., J. D. Miller, and M. N. Neely. 2007. Analysis of the polysaccharide capsule of the systemic pathogen Streptococcus iniae and its implications in virulence. *Infect. Immun.* 75(3): 1255-1264. https://doi.org/10.1128/IAI.01484-06

Lyon, W. R. and M. G. Caparon. 2003. Trigger factor-mediated prolyl isomerization influences maturation of the *Streptococcus pyogenes* cysteine protease. *J. Bacteriol.* 185(12), 3661-3667.

McArthur, A. G., N. Waglechner, F. Nizam, A. Yan, M. A. Azad, A. J. Baylay, K. Bhullar, M. J. Canova, G. De Pascale, L. Ejim, L. Kalan and G. D. Wright. 2013. The comprehensive antibiotic resistance database. *Antimicrob. Agents. Chemother*. 57(7): 3348-3357. https://doi.org/10.1128/AAC.00419-13

McKenna, S., K. K. Huse, S. Giblin, M. Pearson, M. S. Majid Al Shibar, S. Sriskandan, S. Matthews and J. E. Pease. 2022. The role of streptococcal cell-envelope proteases in bacterial evasion of the innate immune system. *J. Innate Immun.* 14(2): 69-88.

Medema, M. H., K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano and R. Breitling. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* 39: W339-W346. https://doi.org/10.1093/nar/gkr466

Nagaoka, K., S. Konno, K. Murase, T. Kikuchi, and I. Nakagawa. 2018. Complete genome sequence of *Streptococcus agalactiae* serotype III, Multilocus sequence type 335 strain HU-GS5823, isolated from a

human patient in Japan with severe invasive infection. *Microbiol. Resour. Announc*. 7(20): 10.1128/mra. 01303-01318.

Nguyen, C. D. 2015. Role of complement proteases in *Streptococcus* pathogenesis in fish. University of Queensland, Australia.

Nho, S. W., J. Hikima, I. S. Cha, S. B. Park, H. B. Jang, C. S. del Castillo, H. Kondo, I. Hirono, T. Aoki and T. S Jung. 2011. Complete genome sequence and immunoproteomic analyses of the bacterial fish pathogen *Streptococcus parauberis*. *J Bacteriol*. 193(13): 3356-3366. https://doi.org/10.1128/JB.00182-11

Nizet, V., B. Beall, D. J. Bast, V. Datta, L. Kilburn, D. E. Low and J. C. De Azavedo. 2000. Genetic locus for streptolysin S production by group A streptococcus. *Infect. Immun.* 68(7): 4245-4254.

Okoye, C. O., K. Dong, Y. Wang, L. Gao, X. Li, Y. Wu, and J. Jiang. 2022. Comparative genomics reveals the organic acid biosynthesis metabolic pathways among five lactic acid bacterial species isolated from fermented vegetables. *New Biotechnolog.* 70: 73-83.

Oku, Y., K. Kurokawa, N. Ichihashi, and K. Sekimizu. 2004. Characterization of the *Staphylococcus aureus mprF* gene, involved in lysinylation of phosphatidylglycerol. *Microbiology*. 150(1): 45-51.

Overbeek, R., R. Olson, G. D. Pusch, G. J. Olsen, J. J. Davis, T. Disz, R. A. Edwards, S. Gerdes, B. Parrello, M. Shukla, V. Vonstein, A. R. Wattam, X. Fangfang and R. Stevens. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42: D206-D214. https://doi.org/10.1093/nar/gkt1226

Park, S. Y., I. H. Kim, H. J. Yu, H. R. Paik, J. S. Son and J. H. Kim. 2021. Complete genome sequence of serotype 3 *Streptococcus suis* INT-01, isolated from a domestic pig in Korea. *J. Anim. Sci. Technol.* 63(3): 662.

Poyart, C., C. Pierre, G. Quesne, B. Pron, P. Berche and P. Trieu-Cuot. 1997. Emergence of vancomycin resistance in the genus *Streptococcus*: characterization of a *vanB* transferable determinant in *Streptococcus bovis*. *Antimicrob. Agents. Chemother.* 41(1): 24-29.

Rahman, M., M. M. Rahman, S. C. Deb, M. S. Alam, M. J. Alam and M. T. Islam. 2017. Molecular Identification of Multiple Antibiotic Resistant Fish Pathogenic *Enterococcus faecalis* and their Control by Medicinal Herbs. *Sci. Rep.* 7: 3747. https://doi.org/10.1038/s41598-017-03673-1

Ross, J., E. Eady, J. Cove, W. Cunliffe, S. Baumberg and J. Wootton. 1990. Inducible erythromycin resistance in staphlyococci is encoded by a member of the ATP-binding transport super-gene family. *Mol. Microbiol.* 4(7): 1207-1214.

Schöner, T. A., S. Gassel, A. Osawa, N. J. Tobias, Y. Okuno, Y. Sakakibara, K. Shindo, G. Sandmann G. and H. B. Bode. 2016. *Aryl* polyenes, a highly abundant class of bacterial natural products, are functionally related to antioxidative carotenoids. *Chem. Bio. Chem*. *17*(3): 247-253.

Shainheit, M. G., M. Mulé, and A. Camilli. 2014. The core promoter of the capsule operon of *Streptococcus pneumoniae* is necessary for colonization and invasive disease. *Infect. Immun. 82*(2): 694-705.

Sharma, P., H. Lata, D. K. Arya, A. K. Kashyap, H. Kumar, M. Dua, A. Ali and A. K. Johri. 2013. Role of pilus proteins in adherence and invasion of *Streptococcus agalactiae* to the lung and cervical epithelial cells. *J. Biol. Chem.* 288(6): 4023-4034.

Shimizu, A., H.Tsukagoshi, T. Sekizuka, M. Kuroda, A. Koizumi, M. Fujita, Y. Yamad and N. Saruki. 2020. Meningitis and bacteremia by nonhemolytic Group B *Streptococcus* strain: A whole genome analysis. *Microbiol. Immunol.* 64(9): 630-634.

Smith, H. E., M. Damman, J. Van Der Velde, F. Wagenaar, H. J. Wisselink, N. Stockhofe-Zurwieden and M. A. Smits. 1999. Identification and characterization of the cps locus of *Streptococcus suis* serotype 2: the capsule protects against phagocytosis and is an important virulence factor. *Infect. Immun.* 67(4): 1750-1756.

Srinivasan, V., B. J. Metcalf, K. M. Knipe, M. Ouattara, L. McGee, P. L. Shewmaker, A. Glennen, M. Nichols, C. Harris, M. Brimmage and B. Ostrowsky. 2014. *vanG* element insertions within a conserved chromosomal site conferring vancomycin resistance to *Streptococcus agalactiae* and *Streptococcus anginosus*. *MBio*. 5(4): 10.1128/mbio. 01386-01314.

Sutcliffe, J., A. Tait-Kamradt and L. Wondrack. 1996. *Streptococcus pneumoniae* and *Streptococcus pyogenes* resistant to macrolides but sensitive to clindamycin: a common resistance pattern mediated by an efflux system. *Antimicrob. Agents. Chemother.* 40(8): 1817-1824.

Teng, J. L., Y. Ma, J. H. Chen, R. Luo, C. H. Foo, T. T. Li, J.Y. Fong, W. Yao, S. S. Wong, K. S. Fung and S. K. Lau. 2022. *Streptococcus oriscaviae* sp. nov. infection associated with Guinea pigs. *Microbiol. Spectr.* 10(3): e00014-00022.

Toniolo, C., E. Balducci, M. R. Romano, D. Proietti, I. Ferlenghi, G. Grandi, F. Berti, I. M. Y. Ros and R. Janulczyk. 2015. *Streptococcus agalactiae* capsule polymer length and attachment is determined by the proteins CpsABCD. *J. Biol. Chem.* 290(15): 9521-9532.

Toranzo, A. E., B. Magarinos and J. L. Romalde. 2005. A review of the main bacterial fish diseases in mariculture systems. *Aquaculture*. 246(1-4):37-61.

Vezina, B., H. Al-Harbi, H. R. Ramay, M. Soust, R. J. Moore, T. W. Olchowy and J. I. Alawneh. 2021. Sequence characterisation and novel insights into bovine mastitis-associated *Streptococcus uberis* in dairy herds. *Sci. Rep.* 11(1): 3046.

Vidal Amaral, J. R., R. T. Jucá Ramos, F. Almeida Araújo, R. Bentes Kato, F. Figueira Aburjaile, S. de Castro Soares, A. Góes-Neto, M. Matiuzzi da Costa, V. Azevedo, B. Brenig, and S. Soares de Oliveira. 2022. Bacteriocin producing *Streptococcus agalactiae* strains isolated from bovine mastitis in Brazil. *Microorganisms*. 10(3): 588.

Vilas-Boas, L. A., S. A. Headley, K. B. Gonçalves, J. A. Scarpassa and L. G. Pretto-Giordano. 2017. Complete Genome Sequence of *Streptococcus iniae* UEL-Si1, Isolated in Diseased Nile Tilapia (*Oreochromis niloticus*) from Northern Paraná, Southern Brazil. *Genome. Announc.* 5(2): 10.1128/genomea. 01458-01416.

Wu, T., Z. Zhao, L. Zhang, H. Ma, K. Lu, W. Ren, Z. Liu, H. Chang, W. Bei, Y. Qiu and and H. Chen. 2011. Trigger factor of *Streptococcus suis* is involved in stress tolerance and virulence. *Microb. Pathog.* 51(1-2): 69-76.

Zankari, E., H. Hasman, S. Cosentino, M. Vestergaard, S. Rasmussen, O. Lund, O., F. M. Aarestrup and M. V. Larsen. 2012. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother*. 67(11): 2640-2644. https://doi.org/10.1093/jac/dks261

Zhang, D., X. Ke, L. Liu, M. Lu, C. Shi, C. and Z. Liu. 2018. *Streptococcus agalactiae* from tilapia (*Oreochromis* sp.) transmitted to a new host, bighead carp (*Aristichthys nobilis*), in China. *Aquac Int.* 26: 885-897.