Host adaptation of aquatic *Streptococcus agalactiae*

THESIS SUBMITTED TO THE UNIVERSITY OF STIRLING

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

by

CHRISTIAN M.J. DELANNOY DVM, MSc

27 MARCH 2013

INSTITUTE OF AQUACULTURE

UNIVERSITY OF STIRLING

Declaration

I declare that this thesis has been composed in its entirety by me. Except where specifically acknowledged, the work described in this thesis has been conducted by me and has not been submitted for any other degree.

Signature:

Signature of supervisor:

Date:

Abstract

Streptococcus agalactiae is a pathogen of multiple hosts. The bacterium, an aetiological agent of septicaemia and meningo-encephalitis in freshwater and saltwater fish species, is considered a major threat to the aquaculture industry, particularly for tilapia. Cattle and humans are however the main known reservoirs for *S. agalactiae*. In humans, the bacterium (commonly referred to as Group B Streptococcus or GBS) is a member of the commensal microflora of the intestinal and genito-urinary tracts, but it is also a major cause of neonatal invasive disease and an emerging pathogen in adults. In cattle, *S. agalactiae* is a well-recognized causative agent of mastitis. Numerous studies focusing on *S. agalactiae* from human and bovine origins have provided insight into the population structure of the bacterium, as well as the genome content and pathogenic mechanisms through identification of virulence determinants. Concerning *S. agalactiae* from aquatic origins, scientific information mainly focused on case reporting and/or experimental challenges, with a limited or absence of information in terms of pathogenesis, virulence determinants and genotypes of the strains involved. The objective of this study was to enhance our understanding of the molecular epidemiology, host-adaptation and pathogenicity of *S. agalactiae* in aquatic species, with particular emphasis on tilapia.

Firstly, a collection of 33 piscine, amphibian and sea mammal isolates originating from several countries and continents was assembled, with the aim of exploring the population structure and potential host specificity of aquatic *S. agalactiae*. Isolates were characterised using pulsed-field gel electrophoresis (PFGE), multi-locus sequence typing (MLST), and a standardised 3-set genotyping system comprising molecular serotypes, surface protein gene profiles and mobile genetic element profiles. Two major subpopulations were identified in fish. The first subpopulation consisted of non-haemolytic isolates that belonged to sequence type (ST) 260 or 261, which are STs that have been reported only from teleosts. These isolates exhibited a low level of genetic diversity by PFGE and clustered with other STs that

have been reported only in fish. Another common feature was the absence of all surface protein genes or mobile genetic elements targeted as part of the 3-set genotyping and that are usually found in human or bovine isolates. The second subpopulation consisted of β-haemolytic isolates recovered from fish, frogs and sea mammals, and that exhibited medium to high genetic diversity by PFGE. STs identified among these isolates have previously been identified from strains associated with asymptomatic carriage and invasive disease in humans. The human pathogenic strain ST7 serotype Ia was detected in fish from Asia. Moreover, ST283 serotype III-4 and its novel single locus variant ST491 detected in fish from Southeast Asia shared a 3-set genotype identical to that of an emerging ST283 clone associated with invasive disease of adult humans in Asia. These observations suggested that some strains of aquatic *S. agalactiae* may present a zoonotic or anthroponotic hazard. STs found among the seal isolates (ST23) have also been reported from humans and numerous other host species, but never from teleosts. This work provided an excellent basis for exploration of the virulence of selected strains in experimental challenges.

The virulence of two strains of *S. agalactiae* was experimentally investigated by intra-peritoneal infection of Nile tilapia (*Oreochromis niloticus*), using an isolate originally recovered from fish and belonging to ST260, and an isolate originating from a grey seal and belonging to ST23. The clinical signs, the *in vivo* distribution of viable bacteria and bacterial antigens, and the gross and histopathological lesions that developed during the time course of the infection were investigated. The ST260 strain was highly virulent, whereas no major clinical sign or mortalities occurred in the fish challenged with the ST23 strain. After injection, both strains however gained access to the bloodstream and viable bacteria were recovered from all organs under investigation. During the early stages of infection, bacteria were mostly found within the reticulo-endothelial system of the spleen and kidney. Thereafter, the ST260 demonstrated a particular tropism for the brain and the heart, but granulomatous inflammation and associated necrotic lesions were observed in all organs. ST23 was responsible for a mixed inflammatory response associated with the

presence of bacteria in the choroid rete and in the pancreatic tissue only. After 7 days post-challenge and for both strain, the formation or containment of bacteria within granulomata or other encapsulated structures appeared to be a major component of the fish response. However, the load of viable bacteria remained high within organs of fish infected with ST260, suggesting that, unlike ST23, this strain is able to survive within macrophages and/or to evade the immune system of the fish. This work demonstrates that the lack of report of ST23 strains in fish is possibly not due to a lack of exposure but to a lack of virulence in this host. The two strains, which differ in prevalence and virulence in fish, provide an excellent basis to investigate genomic differences underlying the host-association of distinct *S. agalactiae* subpopulations.

The genome of the ST260 strain used in challenge studies was sequenced. We therefore provided the first description for the genome sequence of a non-haemolytic *S. agalactiae* isolated from tilapia (strain STIR-CD-17) and that belongs by multi-locus sequence typing (MLST) to clonal complex (CC) 552, which corresponds to a presumptive fish-adapted subgroup of *S. agalactiae*. The genome was compared to 13 *S. agalactiae* genomes of human (n=7), bovine (n=2), fish (n=3) and unknown (n=1) origins. Phylogenetic analysis based on the core genome identified isolates of CC552 as the most diverged of all *S. agalactiae* studied. Conversely, genomes from β-haemolytic isolates of CC7 recovered from fish were found to cluster with human isolates of CC7, further supporting the possibility that some strains may represent a zoonotic or anthroponotic hazard. Comparative analysis of the accessory genome enabled the identification of a cluster of genes uniquely shared between CC7 and CC552, which encode proteins that may provide enhanced fitness in specific niches. Other genes identified were specific to STIR-CD-17 or to CC552 based on genomic comparisons; however the extension of this analysis through the PCR screening of a larger population of *S. agalactiae* suggested that some of these genes may occasionally be present in isolates belonging to CC7. Some of these genes, occurring in clusters, exhibited typical signatures of mobile genetic elements, suggesting their acquisition through horizontal gene

transfer. It is not possible to date to determine whether these genes were acquired through intraspecies transfer or through interspecies transfer from the aquatic environment. Finally, general features of STIR-CD-17 highlighted a distinctive genome characterised by an absence of well conserved insertion sequences, an abundance of pseudogenes, a smaller genomic size than normally observed among human or bovine *S. agalactiae*, and an apparent loss of metabolic functions considered conserved within the bacterial species, indicating that the fish-adapted subgroup of isolates (CC552) has undergone niche restriction.

Finally, genes encoding recognised virulence factors in human *S. agalactiae* were selected and their presence and structural conservation was evaluated within the genome of STIR-CD-17. Numerous genes were absent in STIR-CD-17, while the *cyl* operon responsible for the β-haemolysin production was found to be only partially present, indicating that their encoded proteins are not important contributors to pathogenecity of *S. agalactiae* in fish. The gene encoding an immunogenic bacterial adhesin in certain human *S. agalactiae* (gbs2018) was identified as a distinct variant, unique to fish isolates of CC552, which possibly reflects differences or adaptations in the function of the protein. Finally, a limited set of genes were found to be well-conserved in STIR-CD-17 and included *fbsA*, *pavA*, *srr1*, *cfb*, *hylB*, *ponA* and *sodA*. The *cspA* gene was also found to be well conserved, but a deletion responsible for a frameshift suggested that, if the gene is expressed, the protein would be secreted and not cell-wall anchored. The *cps* genes were also found to be well-conserved, with the exception of *cpsK*, but whether the variations in *cpsK* affect the biosynthesis of the capsule is unknown.

In conclusion this study used a multidisciplinary and sequential approach in order to enhance our understanding of the molecular epidemiology, host-adaptation and virulence of *S. agalactiae* in tilapia. This thesis has laid a firm foundation for further studies that should address the questions of epidemio-surveillance for assessment of transmission of *S.*

agalactiae between humans and fish and the evaluation of the role of putative virulence determinants, with a view to effective control of the disease in fish through prevention or vaccine development.

Acknowledgments

First and foremost I would like to thank my supervisors Dr Mags Crumlish, Prof Ruth Zadoks, Dr Mike Fontaine and Prof Jimmy Turnbull for their guidance, support and encouragements throughout this project. Thanks must also go to the University of Stirling and the Moredun Research Institute for funding my PhD. The chance to have a PhD joint between two institutes and such a diversity of supervisors enabled this project to take unexpected and exciting routes!

I would like to thank Prof. Hugh Ferguson for always being of great support and for welcoming me in his Marine Medicine lab at St George's University. I would also like to thanks the Fisheries Society of the British Isles for the travel grant which allowed me to take up the opportunity to present my work at an international conference.

Thanks must go to all who shared my enthusiasm and helped me to go forward during this project, including Polyana Da Silva, Phuoc Nguyen Ngoc, Thi Luu, Mark Dalgeish, Johanna Baily, Jeanie Finlayson, Niall Auchinachie, Fiona Muir, John Taggart, Mintu Nath, Allen Flockhart and Robert Goldstone.

Finally, this work would not have been possible without the generosity of all researchers who provided the isolates used in the study. My gratitude also goes to all my friends and colleagues at the Moredun and at the Institute of Aquaculture. You know you are far too many to fit in one page and you also know this task would not have been possible without you! A special thank you to my family for their constant support.

Thanks to you all!!!

Table of content

List of figures

List of tables

Abbreviations

µg	Microgram
µl	Microlitre
µm	Micrometer
µM	Micromolar
βH/C	Beta-haemolysin/cytolysin
ACT	Artemis comparison tool
Alp	Apha-like protein
AP	Ancillary protein
BHI	Brain heart infusion
BLAST	Basic alignment search tool
Bp	Base pair
BP	Backbone protein
BSR	Blast score ratio
CC	Clonal complex
cfu	Colony forming unit
cm	Centimeter
COG	Clusters of Orthologous Group
Contig	Contiguous sequence
CPS	Capsular polysaccharide
CRISPR	Clustered regularly interspaced short palindromic repeats
DNA	Deoxyribonucleic acid
DLV	Double-locus variant
DR	Direct repeat
e.g.	Exempli gratia (for example)
et al.	Et alias (and others)
EPS	Exopolysaccharides
FAG	Fish-associated gene
FAO	Food and Agriculture Organisation
FbsA	Fibrinogen-binding protein A
FSG	Fish-specific gene
g	Gram
GBS	Group B streptococcus
GTR	General time reversible model
h	Hours
H&E	Haematoxylin and eosin
i.e.	id est (that is)
i.p.	intra-peritoneal
IHC	Immunohistochemistry
IOA	Institute of Aquaculture
IS	Insertion sequence
L	Litres
LD_{50}	Median lethal dose
Mb	Megabase
MMC	Melano-macrophage centre
MME	Minimal mobile element
MGE	Mobile genetic element
MLST	Multi-locus sequence typing
min	Minute
MRI	Moredun Research Institute
MS	Molecular serotype
n	Number

NCBI	National Centre for Biotechnology Information
NA	Non-available
ND	Non-determined
NT	Non-typeable
OD	Optical density
ORF	Open reading frame
PAI	Pathogenicity island
PBP	Penicillin-binding protein
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFGE	Pulsed-field gel electrophoresis
PGAAP	Prokaryotic Genomes Automatic Annotation Pipeline
PI	Pilus-island
p.i.	Post-inoculation
RAPD	Random amplification of polymorphic DNA
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
SBA	Sheep blood agar
SD	Standard deviation
sec	Seconds
SHMT	Serine hydroxymethyltransferase
SLV	Single-locus variant
SPG	Surface protein gene
Srr	Serine-rich repeat protein
SSG	Strain-specific gene
ST	Sequence type
TLV	Triple-locus variant
Tm	Melting temperature
TSA	Tryptone soya agar
UK	United Kingdom
USA	United States of America
v/v	Volume/volume
V	Volt
w/v	Weight/volume

CHAPTER 1

Streptococcus agalactiae, a pathogen of major significance in tilapia culture. A general introduction.

1.1 *Streptococcus agalactiae* and its relevance as bacterial pathogen

Bacteria belonging to the genus *Streptococcus* (from the Greek *streptos*, meaning pliant or chain, and *coccos*, meaning a grain or berry) are non-motile, non-sporulating Gram-positive spherical cells of 0.5-2 μm in diameter that occur in pairs or chains when grown in liquid media (Holt et al., 1994). They are facultatively anaerobic, catalase-negative and require nutritionally rich media for growth (Holt et al., 1994). Streptococci may produce complete or partial lysis of red blood cells when inoculated onto blood agars, leading to a greenish discoloration (α-haemolysis) or complete clearing (β-haemolysis) around colonies (Holt et al., 1994). A serological classification pioneered by Rebecca Lancefield has proved useful in differentiating streptococci based on the detection of different versions of the major cell-wall polysaccharide (Lancefield group A, B, C, etc.), with the *S. agalactiae* species placed into the Lancefield group B (Lancefield, 1933), leading to its common designation as group B streptococcus (GBS). Lancefield group B solely comprises *S. agalactiae*, and conversely all members of the species present on their surface the group B polysaccharide.

Streptococcus agalactiae (from the latin *agalactia*, no milk) was initially described in bovids in 1887 as a causative agent of mastitis (Nocard and Mollereau, 1887); the bacterium remains to-date a significant cause of subclinical, clinical or chronic mastitis in many herds, affecting milk quality and causing significant economic impact in the dairy sector (Keefe, 1997). In the early 20th century, group B streptococci (GBS) were recognized in humans as commensals of the upper-respiratory, gastro-intestinal and genito-urinary tracts (Lancefield, 1933 and 1935), leading some authors to assign these bacteria as *S. opportunus* (Brown,

1947). At this time, the bacterium was rarely associated with disease in humans (Rantz, 1942), but from the 1960s GBS emerged as a major cause of human neonatal diseases, causing infections such as pneumonia, septicemia and meningitis (Dermer et al., 2004). Moreover, clinical manifestations caused by GBS, including sepsis, soft tissue infection, peri- and endocarditis, pneumonia and urosepsis (Schuchat, 1999; Farley, 1993), have been reported increasingly in human adults over the last decades, with elderly and immuno-compromised adults being at higher risk of infection (Farley and Strasbaugh, 2001).

Streptococcus agalactiae has been isolated from various other mammals, including camels, monkeys, pigs, nutrias, rats, rabbits, horses, cats, dogs, and dolphins, occasionally in association with local or systemic pathological processes (Wibawan et al., 1993; Lämmler et al., 1998; Bekele and Molla, 2001; Yildirim et al., 2002; Hetzel et al., 2003; Zappulli et al., 2005; Ren et al., 2013; Shuster et al., 2013). In addition, *S. agalactiae* infections of poikilothermic animals have been well-documented (**Table 1.1**).

1.2 Geographical and host-species distribution of piscine *S. agalactiae*

The geographic distribution of *S. agalactiae* infections in fish is primarily confined to sub-tropical and tropical regions, but *S. agalactiae* may also occur in temperate regions, either naturally during warm seasons (*e.g,* the natural outbreak in Chesapeake Bay on East coast of the USA; Baya et al., 1990), or through the import of aquatic animals (*e.g.* doctor fish imported into the UK for spa use; Verner-Jeffreys et al., 2012). The bacterium has been reported to affect poikilotherms in all continents except Antarctica and Africa (**Table 1.1**). If the absence of the *S. agalactiae* in Antarctica is understandably linked to the temperature, the absence of the bacterium from Africa is surprising and might be due to the absence of published reports, inadequate identification techniques, or non-susceptible species. Evidence, however, suggests that the bacterium occurs in tilapia farms in Africa but further investigation is required (David Rodgers, personal communication).

Table 1.1 Reports of natural *Streptococcus agalactiae* infections in fish and other poikilothermic animals. The geographic origin and host species are presented along with the outbreak type. W., Wild animal; F., Farmed animal; A, Aquarium fish; NA, Non Available.

Common name	Scientific name	Country	Type	References
North America				
Atlantic croaker	*Micropogon undulatus*	USA	W	Plumb et al., 1974
Bluefish	*Pomatomus saltatrix*	USA	W	Baya et al., 1990
Golden shiners	*Notemigonus crysoleucas*	USA	F	Robinson and Meyer, 1966
Gulf killifish	*Fundulus grandis*	USA	NA	Garcia et al., 2008; Olivares-Fuster et al., 2008
Hybrid striped bass	*Morone saxatilis x Morone chrysops*	USA	NA	Garcia et al., 2008; Olivares-Fuster et al., 2008
Menhaden	*Brevoortia patronus*	USA	W	Plumb et al., 1974
Pinfish	*Lagodon rhomboides*	USA	W	Plumb et al., 1974
Sea catfish	*Arius felis*	USA	W	Plumb et al., 1974
Sea trout	*Cynoscion regalis*	USA	W	Baya et al., 1990
Spot	*Leiostomus xanthurus*	USA	W	Plumb et al., 1974
Stingray	*Dasyatis* sp.	USA	W	Plumb et al., 1974
Striped bass	*Morone saxatilis*	USA	W	Baya et al., 1990
Striped mullet	*Mugil cephalus*	USA	W	Plumb et al., 1974
Tilapia	*Oreochromis* sp.	USA	NA	Berridge et al, 2001; Garcia et al., 2008; Olivares-Fuster et al., 2008
		Honduras	NA	Berridge et al, 2001; Garcia et al., 2008; Olivares-Fuster et al., 2008
South America				
Bullfrog	*Rana catesbeiana*	Brazil	F	Amborski et al., 1983
Tilapia	*Oreochromis* sp.	Brazil	F/NA	Salvador et al., 2005; Evans et al., 2008; Garcia et al., 2008; Olivares-Fuster et al., 2008; Mian et al., 2009
		Colombia	F	Hernandez et al., 2009; Delannoy et al., 2013
		Costa Rica	F	Delannoy et al., 2013
		Honduras	F	Delannoy et al., 2013
Whiteleg shrimp	*Litopenaeus vannamei*	NA	F	Hasson et al., 2009
Eurasia				
Bartail flathead	*Platycephalus indicus*	Kuwait	W	Jafar et al., 2009
Bullfrog	*Rana rugurosa*	Thailand	F	Delannoy et al., 2013
Catfish	*Arius thalassinus*	Kuwait	W	Jafar et al., 2009
Doctor fish	*Garra rufa*	UK	A	Verner-Jeffreys et al., 2012
Golden pompano	*Trachinotus blochii*	Malaysia	F	Amal et al., 2012
Grey mullet	*Mugus cephalus*	Israel	NA	Eldar et al., 1994
Hybrid striped bass	*Morone saxatilis x Morone chrysops*	Israel	NA	Evans et al., 2008; Olivares-Fuster et al., 2008
Klunzinger's mullet	*Liza klunzingeri*	Kuwait	W	Evans et al. 2002; Gilbert et al., 2002; Jafar et al., 2009
Seabream	*Sparus auratus*	Kuwait	F	Evans et al. 2002; Gilbert et al., 2002; Jafar et al., 2009
Silver croaker	*Otolithes argenteus*	Kuwait	W	Jafar et al., 2009
Silver grunt	*Pomaddasys stidens*	Kuwait	NA	Jafar et al., 2009
Silver pomfret	*Pampus argenteus*	Kuwait	F	Duremdez et al., 2004; Jafar et al., 2009
Tilapia	*Oreochromis* sp.	Belgium	F	Delannoy et al., 2013
		China	F	Ye et al., 2011
		Indonesia	F	Lusiastuti et al., 2012
		Israel	NA	Eldar et al., 1994
		Malaysia	F	Musa et al., 2009; Abuseliana et al., 2010; Zamri-Saad et al., 2010
		Thailand	F	Suanyuk et al., 2008
		Vietnam	F	Delannoy et al., 2013
Ya fish	*Schizothorax prenanti*	China	F	Geng et al., 2012
Yellowtail	*Seriola quinqueradiata*	Japan	NA	Eldar et al., 1994
Australia				
Crocodile	*Crocodylus porosus*	Australia	F	Bishop et al., 2007
Estuary stingray	*Dasyatis fluviorum*	Australia	W	Bowater et al., 2012
Giant grouper	*Epinephelus lanceolatus*	Australia	W	Bowater et al., 2012
Giant sea catfish	*Arius thalassinus*	Australia	W	Bowater et al., 2012
Golden ram	*Mikrogeophagus ramirezi*	Australia	A	Delannoy et al., 2013
Javelin grunter	*Pomadasys kaakan*	Australia	W	Bowater et al., 2012
Rosy Barb	*Puntius conchonius*	Australia	A	Delannoy et al., 2013
Shovelnose ray	*Aptychotrema rostrata*	Australia	W	Bowater et al., 2012
Mangrove whipray	*Himantura granulata*	Australia	W	Bowater et al., 2012

Natural infections have occasionally been reported in shrimps, frogs, and crocodiles but most reports concern fish species. A total of 34 naturally-infected fish species have been reported thus far, including freshwater, marine and estuarine species, and involving both cultured and wild populations (**Table 1.1**). Four major epizootics affecting estuarine and marine wild fish have been described: one off the South-East Coast of the USA (Gulf of Mexico) in 1972 (Plumb et al., 1974), one in the North-East coast of the USA (Chesapeake Bay) in 1988 (Baya et al., 1988), one off the Kuwait Bay (Persian Gulf) in 2001 (Evans et al., 2002), and finally one in Australia (Northern Queensland) that occurred sporadically and continuously from 2007 to 2011 (Bowater at al., 2012). Other infections due to *S. agalactiae* mostly refer to the aquaculture industry. Initially, *S. agalactiae* was responsible for a disease outbreak in cultured freshwater golden shiners in the USA (Robinson and Meyer, 1966); however it is now considered a major threat to the aquaculture industry with most recent outbreak reports involving fish species of economic importance (**Table 1.1**).

1.3 Importance of aquaculture and tilapia farming

Fish are a vital foodstuff and over three billion people are nutritionally dependent upon fish (FAO, 1997). In 2010, aquaculture accounted for 46 percent of total food fish supply for human consumption (FAO, 2010b), with tilapia (*Oreochromis* sp.) one of the most productive and internationally traded farmed species.

Tilapia is a major source of protein in numerous developing countries (FAO, 2003-2013). It presents excellent culture characteristics, such as easy reproduction in captivity, tolerance to a wide range of environmental conditions, efficient use of low protein diets, fast growth and relative resistance to stress and infection (Hassanien et al., 2005). Although native to Africa and the Middle East, tilapia has become a globally-important aquatic species, produced in nearly 100 countries worldwide (FAO, 2010b). In view of the increasing commercialization and continuing growth of the tilapia industry, the commodity is now among the most important farmed fish globally, second only to carp but surpassing the salmonid group

(FAO, 2010a). World tilapia production is led by China, followed quite distantly by Egypt, Indonesia and Thailand (FAO, 2010a). Diseases have however become a primary constraint to tilapia culture growth. In China, the main streptococcal pathogen has been considered to be *S. iniae*, but a shift towards a predominance of *S. agalactiae* infections has been described in recent years (Wang et al., 2013). In some Latin American countries, *S. agalactiae* has thus far been the only streptococcal species isolated from natural outbreaks of disease in tilapia (Jiménez et al., 2011).

1.4 *Streptococcus agalactiae* infections in tilapia

Streptococcal infections, which have increased in number during the last decade as a consequence of the intensification of aquaculture, are responsible for significant economic losses in the farmed fish industry. Streptococcosis of fish, from a clinical point of view, is a generic term used to designate similar clinical manifestations in which different species of Gram-positive cocci are involved (Becovier et al., 1997). Among the main pathogenic species responsible for these streptococcal infections is *S. agalactiae*. Recorded losses in natural outbreaks of *S. agalactiae* in farmed tilapia are variable, with mortalities as high as 83.4% reported in some Malaysian farms (Zamri-Saad et al., 2010).

1.4.1 Route of transmission

Despite the significance of *S. agalactiae* disease in farmed tilapia, relatively little is known about the route of transmission of this pathogen. Vertical transmission of the disease is considered to be unlikely, as Hernández et al. (2009) and Jiménez et al. (2011) were unable to detect the bacterium from a batch of infected broodstock-derived larvae and juvenile tilapia; it is therefore generally accepted that the transmission of the disease predominantly occurs through horizontal transfer between congeners. This hypothesis is further supported by successful experimental challenges via cohabitation (Mian et al., 2009), even so it is unclear whether horizontal transfer occurs by ingestion, through the water, or by direct contacts between congeners.

The oral route of pathogen transmission in fish has been well described, including for *S. iniae* (Bromage and Owens, 2002). It can occur through ingestion of feces from infected fish (fecal-oral route), through ingestion of contaminated water, or through cannibalism and ingestion of infected animal tissues. Rasheed et al. (1984), however, demonstrated that the administration of 8 x 10^6 colony forming unit (cfu) of pathogenic *S. agalactiae* directly into the stomach of Gulf killifish resulted in an absence of death and bacterial recovery from internal organs 24h post-infection, suggesting that an oral route of disease transmission in this fish species is unlikely. To our knowledge, the oral route of transmission has not been evaluated further using different *S. agalactiae* strains, neither has it been tested in tilapia.

Unlike other fish pathogens, no study on the survival of *S. agalactiae* in the water has been reported and it is therefore not known whether the water represents an important vehicle for the transmission of the disease. Pasnik et al. (2009) recovered *S. agalactiae* from the intestine of intra-peritoneally infected tilapia and suggested a possible shedding of the pathogen into the environment and feco-oral transmission of the bacterium. However, the presence of *S. agalactiae* in the feces themselves has not been evaluated and the sampling of water, mud or sediment surrounding infected farms resulted in an absence (Hernández et al., 2009) or rare (Amal et al., 2013) recovery of *S. agalactiae*. In theory, however, the transmission of *S. agalactiae* can occur through the water, as exemplified by numerous challenges successfully reproducing the disease by immersion or gill inoculation (Mian et al., 2009; Rasheed et al., 1984; Pereira et al., 2009). The gills and the skin (with or without injury) have been suggested to represent an important route of entry (Mian et al., 2009; Rasheed et al., 1984). Direct contact with infected congeners is also considered an important factor of transmission of the disease (Mian et al., 2009), but this has not been assessed as it is not possible to exclude the possibility of transmission through the water in experimental challenges.

In conclusion, the existing evidence suggests that the disease is transmitted horizontally between congeners but much more work is required to elucidate the route of transmission and to accurately identify the natural portal(s) of entry.

1.4.2 Pathophysiology

Streptococcus agalactiae disease has been extensively described in both natural and experimental infections of fish. The intensity and progression of the disease is somewhat variable and possibly depends on the virulence of the strain. For example, Mian et al. (2009) infected tilapia (40g) intra-peritoneally with non-haemolytic isolates recovered from natural infections in fish farms in Brazil and observed mortalities up to 100% with only 10 cfu per fish. In contrast, β-haemolytic isolates recovered from natural infections in wild mullet appeared to be less virulent, with a determined median lethal dose (LD_{50}) of 1.9 x $10^{3.3}$ cfu per fish after intra-peritoneal infection of 30g tilapia (Evans et al., 2002). Other factors influencing infection outcome may be the specific host-susceptibility and environmental factors involving water quality. For example, different tilapia species demonstrate different levels of resistance to *S. agalactiae* (Huang et al., 2013) and low dissolved oxygen levels have been associated with higher mortalities in *S. agalactiae* infected tilapia (Evans et al., 2003).

Most experimental and natural infections share a number of common features which are similar to aquatic infections caused by *S. agalactiae* or *S. iniae* (Agnew et al., 2007). In its per-acute form, the disease is characterized by rapid and sudden death with no expression of clinical changes (Eldar et al., 1995). In the sub-acute or chronic forms, the clinical or external signs classically reported refer to behavioral manifestations and eye lesions. All studies reported an abnormal swimming behavior of infected tilapia such as spiraling, spinning, whirling at the surface, head up or tail down swimming, upside down swimming, and serpentine or exaggerated movements, whereas eye lesions corresponded to unilateral or bilateral exophthalmia, corneal opacity and periorbital or intra-ocular hemorrhages (Evans

et al., 2002; Suanyuk et al., 2005; Hernández et al., 2009; Inocente Filho et al., 2009; Zamri-Saad et al., 2010). Other clinical signs may comprise anorexia, lethargy, C-shaped curvature of the spine or other skeletal deformities, hemorrhages on the mouth, fins or vent, darkening of the skin and abdominal distension (Evans et al., 2002; Pasnik et al., 2005; Hernández et al., 2009; Zamri-Saad et al., 2010).

Internally, a consistent finding is the presence of intracranial oedema accompanied with brain softening or haemorrhages (Zamri-Saad et al., 2010). In addition to the eye lesions, gross pathological changes may consist of hepatomegaly, pale-colored liver, splenomegaly, and congestion of the kidney (Suanyuk et al., 2008; Inocente Filho et al., 2009; Zamri-Saad et al., 2010). The presence of white fibrinous exudate covering the heart or epicardial opacity has also been noticed in a few studies (Hernández et al., 2009; Inocente Filho et al., 2009), as well as visceral adherence and ascites (Eldar et al., 1995).

The main histopathological findings are located in the brain and in the eyes, but since *S. agalactiae* causes systemic disease, other organs may be affected including the kidney, liver, spleen and heart. It consists generally of focal to multifocal granulomatous inflammation and eventual granuloma formation, accompanied by multifocal necrotic inflammatory lesions (Hernández et al., 2009). Degenerative changes of endothelial cells, marked congestion and formation of thrombi associated with tissue infarction have also been reported in numerous organs (Zamri-Saad et al., 2010). The pathological and histopathological findings in the brain and in the eyes correlate with the clinical signs observed: the resulting meningo-encephalitis explains the behavioural manifestations, whereas the infiltration of inflammatory cells and exudates in the periorbital tissues and in the cornea explain the exophthalmia and corneal opacity.

The pathogenesis of *S. agalactiae* in fish is not understood fully yet. No study has focused on elucidating the progression of the bacterium and associated lesions during the time-course of an infection. To-date, all studies of experimentally or naturally-infected tilapia

have only evaluated the lesions or the presence of bacteria as observed at a specific sampling point or in moribund fish. It is believed that the bacterium primarily causes a septicaemic state, and that the presence of bacteria or toxins within the blood vessels causes degenerative changes of endothelial cells, thrombi formation and associated tissue infarction, which promote rapid bacterial dissemination and subsequent colonization of numerous internal organs (Zamri-Saad et al., 2010). The primary target organ for *S. agalactiae* have been found to be the brain, followed the eye and the heart (Hernández et al., 2009). The identification of numerous putatively viable bacteria within macrophages has led researchers to hypothesize that *S. agalactiae* may use macrophages as a vehicle to cross the blood brain barrier (Hernández et al., 2009). This phenomenon is well-accepted for *S. iniae*, which has been shown to withstand the bactericidal activity of macrophages and to trigger their apoptosis (Zlotkin et al., 2003). However, the ability of *S. agalactiae* to evade the immune system of fish has not been experimentally evaluated thus far.

1.4.3 Detection and identification

Detection methods are generally based upon culture of the bacterium, whereas presumptive identification is traditionally performed using routine phenotypic tests, which may include biochemical tests (Evans et al., 2002). However, the phenotypic profiles of *S. agalactiae* strains are quite heterogeneous and may lead to misidentifications. In humans, pathogenic *S. agalactiae* are usually β-haemolytic, and their growth, phenotypic characteristics and metabolic capabilities have been well-characterized (Holt et al., 1994). Assuming the same to be true of *S. agalactiae* from other species has led to misidentification of isolates from fish which differ from "classic" *S. agalactiae* by an absence of growth at 37°C, slow growth on Brain Heart Infusion agar plates at 30°C (1mm colonies after 48h incubation), restricted metabolic capabilities (as assessed by biochemical tests), and an absence of β-haemolysis (Eldar et al., 1994). However, phenotypic tests based on Lancefield sero-grouping may also be performed, which allows accurate identification of the bacterium as belonging to Lancefield group B.

In addition to culture and phenotypic tests, molecular methods for the detection and definitive identification of *S. agalactiae* have been described and may be performed using DNA extracted directly from tissues of infected animals or bacteria propagated in the laboratory. Researchers have designed species-specific oligonucleotide primers to amplify by Polymerase Chain Reaction (PCR) the intergenic spacer region between the 16S and 23S ribosomal gene (Berridge et al., 2001; Phuektes et al., 2001), the 16S ribosomal gene (Abdulmawjood and Lämmler, 1999), the 23S ribosomal gene (Kawata et al., 2004), and the CAMP factor encoding gene *cfb* (Kong et al., 2002). Most of these primers have been designed for human or bovine *S. agalactiae* isolates and successfully used for identification of piscine *S. agalactiae* isolates from laboratory cultures or tissues (Suanyuk et al., 2008; Evans et al., 2008; Duremdez et al., 2004; Jiménez et al., 2011). In addition, an alternative more sensitive and rapid approach for nucleotide amplification, called Loop-Mediated Isothermal amplification (LAMP), has recently been developed for detection and identification of *S. agalactiae* from fish, using primers targeting the *fbsB* gene (Wang et al., 2012).

In the field, the presence of typical clinical signs and demonstration of Gram-positive cocci from internal organs constitutes a presumptive diagnostic of streptococcosis (Amal. and Zamri-Saad, 2011). Recovery and culture of the organism, followed by phenotypic and/or genotypic tests may then be used to consolidate or confirm the diagnosis. However, these techniques are time-consuming and may delay the initiation of an appropriate treatment. For a more rapid identification, DNA extraction from tissues followed by amplification of target bacterial sequences may be performed, but this would still require laboratory work. Interestingly, an optical immunoassay for detection of the Lancefield group B antigen was recently shown to be an interesting tool for the rapid in-field identification of *S. agalactiae* using nare and brain swabs (Evans et al., 2010).

1.4.4 Prevention and therapy

The strategy commonly adopted to control bacterial disease outbreaks in aquaculture consists of the use of antibiotic-medicated food. Antibiotic usage must be approved for a specific purpose depending on each country's legislation. Antibiotics like florfenicol are commercially available for the control of fish pathogens, including *S. agalactiae*, after its incorporation into feed pellets (Aquaflor®, MERK Animal Health). However, although the control of in-farm or experimental *S. iniae* infections by the use of oral antibiotics is well-documented in the literature (Stoffregen et al., 1996; Darwish et al., 2005), no such information is available for *S. agalactiae*. It is well known that diseased fish stop feeding, and any in-feed treatment therefore requires the identification of the outbreak at an early stage. Moreover, as for any antibiotherapy in farmed animals, the choice of the antibiotic should be based on or refined following the isolation of the pathogen and antibiogram testing, thus providing a targeted therapeutic treatment to assure efficacy and avoid the development of antibiotic resistance. Following treatment, a withdrawal period should be respected to avoid antibiotic residues in fish intended for human consumption. Because of the possible development of resistance to antibiotics, the limitations of antibiotic-medicated food, and the absence of antibiotics registered for use in aquaculture in some countries like Australia (Agnew and Barnes, 2007), efforts to control the disease have focused on the development of a commercial vaccine.

Despite the industry-preference for a vaccine, there are however few published studies on vaccination trials against *S. agalactiae* infections in tilapia. Eldar et al. (1995) and Evans and colleagues (Evans et al., 2004; Pasnik et al., 2005) developed injectable *S. agalactiae* vaccines, and presented those as highly efficacious against both homologous and heterologous strains. The vaccines developed by Eldar et al (1995) were composed of formalin-killed bacteria or protein extracts, using as master-seed a non-haemolytic *S. agalactiae* isolate. Vaccinated fish were challenged with the vaccine master-seed isolate and other non-haemolytic isolates originating from distinct outbreaks. Vaccination of 165±15g

tilapia provided a long lasting protection with a relative percent of survival (RPS) close to 100%; however, no indication was provided as to whether the challenge isolates were genetically distinct. Furthermore, protection against β-haemolytic *S. agalactiae* was not evaluated. Evans and colleagues (Evans et al., 2004; Pasnik et al., 2005) developed a vaccine composed of concentrated extracellular products and formalin-killed whole-cells, using a β-haemolytic strain originating from the 2001 Kuwait outbreak. Vaccination of 30g tilapia provided a long-lasting protection with a RPS of 80%; however, the challenge isolates all appeared to originate from the 2001 Kuwait outbreak after retrieval of the bacterial identification name from another publication (Garcia et al., 2008). Thus, it is unknown whether the vaccine provides cross-protection against epidemiologically unrelated β-haemolytic strains or against non-haemolytic strains. In an elegant study by Cheng et al. (2012), vaccine candidates were screened using both single and genetically-distinct strains chosen after PFGE typing. Formalin-killed vaccines were developed, and vaccinated 51±8g tilapia were experimentally-infected using single strains or a mixture of distinct strains. Their results demonstrated that an efficient vaccine could not be obtained using a single strain as master-seed. Therefore, it is likely that, as described for *S. agalactiae* in humans (Maione et al., 2005), a single commercially-available vaccine will only be possible through the appropriate combination of strains or antigens.

Autogenous vaccines have been produced and used in response to *S. agalactiae* outbreaks in individual farms, but there are no scientific publications on the efficacy or cross-protection against other strains afforded by these vaccines. Moreover, an inactivated oil-adjuvanted injection vaccine called AquaVac® STREP *Sa* (MSD Animal Health) is commercially-available in some countries, including Brazil, Costa Rica and Indonesia, to aid in the protection against a specific *S. agalactiae* biotype. The biotype nomenclature is not commonly used in scientific publications; biotype 1 corresponds to "classical" β-haemolotyic isolates, whereas biotype 2 corresponds to the non-haemolytic isolates previously called *S. difficile*. The AquaVac® STREP *Sa* is described as safe and effective

against *S. agalactiae* biotype 2 in both laboratory and field trials; however, it does not provide any protection against *S. agalactiae* biotype 1.

1.5 Typing and molecular epidemiology

In addition to biotyping, *S. agalactiae* typing is traditionaly based on the identification of distinct capsular polysaccharides, which may be achieved using phenotypic or genotypic methods, although phenotypic capsule typing has lower typeability and discriminatory power than genotypic alternatives (Lancefield, 1934; Kong et al., 2008; Yao et al., 2013). A total of 10 capsular serotypes have been identified thus far (Ia, Ib, II to IX), with serotype III further subdivided into 4 sero-subtypes (Kong et al., 2002; Slotved et al., 2007). Only 3 serotypes have been reported among piscine isolates (Ia, Ib and III), with serotypes Ia and III corresponding to β-haemolytic isolates predominantly found in Asian countries, and serotype Ib corresponding to non-haemolytic isolates predominantly found in American countries (Evans et al., 2008; Suanyuk et al., 2008).

Numerous comparative typing methods have been developed in the past and are widely used in studies of *S. agalactiae* associated with humans or bovids. Techniques such as random amplification of polymorphic DNA (RAPD) or pulsed-field gel electrophoresis (PFGE) are useful tools for outbreak investigations. Two studies have successfully applied RAPD for investigation of piscine *S. agalactiae* outbreaks; Amal et al. (2013) used the RAPD to demonstrate that infection of tilapia in Malaysian farms originated from the acquisition of juveniles from a single contaminated hatchery. Jafar et al (2009) demonstrated using RAPD that *S. agalactiae* isolates from sewage samples had similar banding patterns to those isolated from infected fish during the Kuwait outbreak of 2001, suggesting a human source of contamination of the fish. Finally, PFGE has also been used with piscine *S. agalactiae* in order to identify different genotypes for rational selection of strain candidates to be used in challenge studies and vaccine development (Pereira et al., 2009; Chen et al., 2012).

For global population studies of *S. agalactiae*, approaches using markers that are part of the core genome are preferred since these markers are presumed to evolve slowly by the random accumulation of neutral variations, providing much more reliable data for both accurate typing and phylogeny estimation. For this, multi-locus sequence typing (MLST) is generally accepted as the standard method (Urwin and Maiden, 2003). Using MLST, isolates were found to cluster into well-defined clonal complexes (CCs), with some of those corresponding to sub-populations associated with specific host-species or clinical manifestations. Among the different clonal complexes identified, CC17 corresponds to a very homogenous group adapted to humans, which is over-represented among isolates responsible for invasive infections in neonates (Brochet et al., 2006). In comparison, CC67 is considered to constitute a bovine-adapted subpopulation (Bisharat et al., 2004), but many other CCs are commonly found in cattle (Zadoks, 2011). Using MLST, Evans et al. (2008) showed that the genotype of β-haemolytic isolates responsible for the 2001 Kuwait outbreak has also been associated with human neonatal infections, further supporting that humans may have been the source of the Kuwait epizootic. In addition, Evans et al. (2008) as well as Brochet et al (2006) identified a number of new STs among several fish species, particularly in tilapia.

To enhance the discriminatory power of *S. agalactiae* typing methods, a typing scheme based on the accessory genome, also called 3-set genotyping, has been developed. It consists of the identification of the molecular serotype, plus the determination of the presence or absence of 7 surface protein encoding genes and 7 mobile genetic elements (Kong et al., 2003). This typing scheme was further extended into a 4-set genotyping system through the additional detection of 7 antibiotic resistance genes (Suanyuk et al., 2008). The advantage of this technique is that it is portable and allows for further differentiation of isolates following MLST (Sun et al., 2005). This genotyping method has been successfully applied to isolates recovered from affected tilapia in Thailand, leading to the identification

of a genotype previously reported among invasive human *S. agalactiae* (Suanyuk et al., 2008).

1.6 Project aims and hypotheses

This study, jointly funded by the University of Stirling and the Moredun Research Institute, aimed to enhance our understanding of the molecular epidemiology, host-adaptation and virulence of *S. agalactiae* in tilapia. Hypotheses addressed throughout this work are presented below, with the aim of addressing each hypothesis within 1 year of the PhD.

The first hypothesis is that *S. agalactiae* infections in fish may be caused by a subset of strains that are found in humans, whereas other strains may be unique to fish. This hypothesis is based on 2 previous studies using either MLST (Evans et al. 2008) or 3-set genotyping (Suanyuk et al., 2008) on piscine *S. agalactiae* isolates. In this thesis, a global collection of *S. agalactiae* from multiple aquatic host species will be collected and subjected to a combination of typing methods (**Chapter 2**), including the joint use of the 2 portable genotyping methods previously mentioned. In order to answer our hypothesis, the prevalence of specific genotypes obtained will be compared to previous typing reports of strains isolated from humans, bovines and fish.

The second hypothesis is that the epidemiological distribution observed (*i.e.* only a subset of the human *S. agalactiae* population is found in fish) is due to differences in virulence as opposed to differences in exposure. To assess this hypothesis, an injectable model of tilapia infection will be used in order to compare the virulence of two *S. agalactiae* strains, one belonging to a subpopulation never encountered in fish, and another one belonging to a subpopulation uniquely recovered from fish (**Chapter 3**). Differences will be measured through comparison of mortalities, bacterial load within organs and histopathology. Comparisons of mortalities with previous publications using similar *in-vivo* models will further allow the grading of virulence of specific strains within fish.

The final hypothesis is that host-adaptation and virulence of *S. agalactiae* strains in fish is driven at the genomic level by regions of gene loss, gene gain, or diversification, as

previously reported among *S. agalactiae* strains infecting bovines (Richards et al., 2011). For this, a piscine *S. agalactiae* strain will be selected based on previous *in vitro* (**Chapter 2**) and *in vivo* (**Chapter 3**) studies and its genome will be fully sequenced. Comparative genomic analyses between *S. agalactiae* genomes from different host species (human, bovine and fish) will be performed in order to identify and describe genomic differences (**Chapter 4**). Furthermore, the conservation or diversification of genes encoding well recognised virulence determinants in *S. agalactiae* of human origin will be assessed within the *S. agalactiae* genome of fish origin (**Chapter 5**).

1.7 References

Abdulmawjood, A. and Lämmler, C., 1999. Amplification of 16S ribosomal RNA gene sequences for the identification of streptococci of Lancefield group B. Res. Vet. Sci. 67, 159-162.

Abuseliana, A., Hassan, D., Saleha, A., Siti Khairani, B., Alsaid, M., 2010. *Streptococcus agalactiae* the etiological agent of mass mortality in farmed red tilapia (*Oreochromis* sp.). J. Anim. Vet. Adv. 9, 2640-2646.

Agnew, W. and Barnes, A.C., 2007. *Streptococcus iniae*: An aquatic pathogen of global veterinary significance and a challenging candidate for reliable vaccination. Vet. Microbiol. 122, 1-15.

Amal, M. and Zamri-Saad, M., 2011. Streptococcosis in tilapia (*Oreochromis niloticus*): A review. Pertanika J. Trop. Agric. Sci. 34, 195-206.

Amal, M., Zamri-Saad, M., Siti - Zahrah, A., Zulkafli, A., 2013. Transmission of *Streptococcus agalactiae* from a hatchery into a newly established red hybrid tilapia, *Oreochromis niloticus* (L.)× *Oreochromis mossambicus* (Peters), farm. J. Fish Dis. In press.

Amal, M.N.A., Zamri-Saad, M., Iftikhar, A.R., Siti-Zahrah, A., Aziel, S., Fahmi, S., 2012. An outbreak of *Streptococcus agalactiae* infection in cage-cultured golden pompano, *Trachinotus blochii* (Lacépède), in Malaysia. J. Fish Dis. 35, 849-852.

Amborski, R.L., Snider, T.G., Thune, R.L., Culley, D.D.J., 1983. A non-hemolytic, group B streptococcus infection of cultured bullfrogs, *Rana catesbeiana*, in Brazil. J. Wildl. Dis. 19, 180-184.

Baya, M., Lupiani, B., Hetrick, F.M., Roberson, B.S., Lukacovic, R., May, E., Poukish, C., 1990. Association of *Streptococcus* sp. with fish mortalities in the Chesapeake Bay and its tributaries. J. Fish Dis. 13, 251-253.

Bercovier, H., Ghittino, C., Eldar, A., 1997. Immunization with bacterial antigens: Infections with streptococci and related organisms. Dev. Biol. Stand. 90, 153-160.

Bekele, T. and Molla, B., 2001. Mastitis in lactating camels (*Camelus dromedarius*) in afar region, north-eastern ethiopia. Berl. Munch. Tierarztl. Wochenschr. 114, 169-172.

Bellais, S., Six, A., Fouet, A., Longo, M., Dmytruk, N., Glaser, P., Trieu-Cuot, P., Poyart, C., 2012. Capsular switching in group B streptococcus CC17 hypervirulent clone: A future challenge for polysaccharide vaccine development. J. Infect. Dis. 206, 1745-1752.

Berridge, B.R., Bercovier, H., Frelier, P.F., 2001. *Streptococcus agalactiae* and *Streptococcus difficile* 16S–23S intergenic rDNA: Genetic homogeneity and species-specific PCR. Vet. Microbiol. 78, 165-173.

Bishop, E.J., Shilton, C., Benedict, S., Kong, F., Gilbert, G.L., Gal, D., Godoy, D., Spratt, B.G., Currie, B.J., 2007. Necrotizing fasciitis in captive juvenile *Crocodylus porosus* caused by *Streptococcus agalactiae*: An outbreak and review of the animal and human literature. Epidemiol. Infect. 135, 1248-1255.

Brochet, M., Couvé, E., Zouine, M., Vallaeys, T., Rusniok, C., Lamy, M.C., Buchrieser, C., Trieu-Cuot, P., Kunst, F., Poyart, C., 2006. Genomic diversity and evolution within the species *Streptococcus agalactiae*. Microb. Infect. 8, 1227-1243

Bowater, R.O., Forbes-Faulkner, J., Anderson, I.G., Condon, K., Robinson, B., Kong, F., Gilbert, G.L., Reynolds, A., Hyland, S., McPherson, G., Brien, J.O., Blyde, D., 2012. Natural outbreak of *Streptococcus agalactiae* (GBS) infection in wild giant Queensland grouper, *Epinephelus lanceolatus* (Bloch), and other wild fish in northern Queensland, Australia. J. Fish Dis. 35, 173-186.

Bromage, E. and Owens, L., 2002. Infection of barramundi Lates calcarifer with *Streptococcus iniae*: Effects of different routes of exposure. Dis. Aquat. Org. 52, 199-205.

Brown, J., 1947. Bovine mastitis and human disease. In: Proceedings Ofthe Fourth International Congress of Microbiology. Rosenkilde and Bagger, Copenhagen, 1947, 198-200.

Chen, M., Wang, R., Li, L., Liang, W., Li, J., Huang, Y., Lei, A., Huang, W., Gan, X., 2012. Screening vaccine candidate strains against *Streptococcus agalactiae* of tilapia based on PFGE genotype. Vaccine 30, 6088-6092.

Darwish, A.M. and Hobbs, M.S., 2005. Laboratory efficacy of amoxicillin for the control of *Streptococcus iniae* infection in blue tilapia. J. Aquat. Anim. Health 17, 197-202.

Delannoy, C., Crumlish, M., Fontaine, M., Pollock, J., Foster, G., Dagleish, M., Turnbull, J., Zadoks, R., 2013. Human *Streptococcus agalactiae* strains in aquatic mammals and fish. BMC Microbiol. 13, 41.

Dermer, P., Lee, C., Eggert, J., Few, B., 2004. A history of neonatal group B streptococcus with its related morbidity and mortality rates in the United States. J. Pediatr. Nurs. 19, 357-363.

Duremdez, R., Al-Marzouk, A., Qasem, J., Al-Harbi, A., Gharabally, H., 2004. Isolation of *Streptococcus agalactiae* from cultured silver pomfret, *Pampus argenteus* (Euphrasen), in Kuwait. J. Fish Dis. 27, 307-310.

Eldar, A., Bejerano, Y., Bercovier, H., 1994. *Streptococcus shiloi* and *Streptococcus difficile*: Two new streptococcal species causing a meningoencephalitis in fish. Curr. Microbiol. 28, 139-143.

Eldar, A., Bejerano, Y., Livoff, A., Horovitcz, A., Bercovier, H., 1995. Experimental streptococcal meningo-encephalitis in cultured fish. Vet. Microbiol. 43, 33-40.

Eldar, A., Shapiro, O., Bejerano, Y., Bercovier, H., 1995. Vaccination with whole-cell vaccine and bacterial protein extract protects tilapia against *Streptococcus difficile* meningoencephalitis. Vaccine 13, 867-870.

Evans, J.J., Bohnsack, J.F., Klesius, P.H., Whiting, A.A., Garcia, J.C., Shoemaker, C.A., Takahashi, S., 2008. Phylogenetic relationships among *Streptococcus agalactiae* isolated from piscine, dolphin, bovine and human sources: A dolphin and piscine lineage associated with a fish epidemic in Kuwait is also associated with human neonatal infections in Japan. J. Med. Microbiol. 57, 1369-1376.

Evans, J., Klesius, P., Gilbert, P., Shoemaker, C., Al Sarawi, M., Landsberg, J., Duremdez, R., Al Marzouk, A., Al Zenki, S., 2002. Characterization of β-haemolytic group B *Streptococcus agalactiae* in cultured seabream, *Sparus auratus* L., and wild mullet, *Liza klunzingeri* (Day), in Kuwait. J. Fish Dis. 25, 505-513.

Evans, J.J., Pasnik, D.J., Klesius, P.H., 2010. A commercial rapid optical immunoassay detects *Streptococcus agalactiae* from aquatic cultures and clinical specimens. Vet. Microbiol. 144, 422-428.

Evans, J.J., Klesius, P.H., Shoemaker, C.A., 2004. Efficacy of *Streptococcus agalactiae* (group B) vaccine in tilapia (*Oreochromis niloticus*) by intraperitoneal and bath immersion administration. Vaccine 22, 3769-3773.

Evans, J.J., Shoemaker, C.A., Klesius, P.H., 2003. Effects of sublethal dissolved oxygen stress on blood glucose and susceptibility to *Streptococcus agalactiae* in nile tilapia *Oreochromis niloticus*. J. Aquat. Anim. Health 15, 202-208

FAO. 2003-2013. Fisheries and Aquaculture topics. Capture fisheries resources. Topics Fact Sheets. In: FAO Fisheries and Aquaculture Department [online]. Rome. Updated 27 September 2001.

FAO. 2010a.World aquaculture production by species group. Rome, Italy.

FAO. 2010b.World review of fisheries and aquaculture. Rome, Italy.

Farley, M.M., Harvey, C., Stull, T., Smith, J.D., Schuchat, A., Wenger, J.D., Stephens, D.S., 1993. A population-based assessment of invasive disease due to group B streptococcus in nonpregnant adults. N. Engl. J. Med. 328, 1807-1811.

Farley, M.M. and Strasbaugh, L.J., 2001. Group B streptococcal disease in nonpregnant adults. Clinical Infectious Diseases 33, 556-561.

Garcia, J.C., Klesius, P.H., Evans, J.J., Shoemaker, C.A., 2008. Non-infectivity of cattle *Streptococcus agalactiae* in nile tilapia, *Oreochromis niloticus* and channel catfish, *Ictalurus punctatus*. Aquaculture 281, 151-154.

Geng, Y., Wang, K.Y., Huang, X.L., Chen, D.F., Li, C.W., Ren, S.Y., Liao, Y.T., Zhou, Z.Y., Liu, Q.F., Du, Z.J., Lai, W.M., 2012. *Streptococcus agalactiae*, an emerging pathogen for cultured ya-fish, *Schizothorax prenanti*, in China. Transbound. Emerg. Dis. 59, 369-375.

Glibert, P.M., Landsberg, J.H., Evans, J.J., Al-Sarawi, M.A., Faraj, M., Al-Jarallah, M.A., Haywood, A., Ibrahem, S., Klesius, P., Powell, C., 2002. A fish kill of massive proportion in Kuwait bay, Arabian Gulf, 2001: The roles of bacterial disease, harmful algae, and eutrophication. Harmful Algae 1, 215-231.

Hasson, K.W., Wyld, E.M., Fan, Y., Lingsweiller, S.W., Weaver, S.J., Cheng, J., Varner, P.W., 2009. Streptococcosis in farmed *Litopenaeus vannamei*: A new emerging bacterial disease of penaeid shrimp. Dis. Aquat. Org. 86, 93.

Hernández, E., Figueroa, J., Iregui, C., 2009. Streptococcosis on a red tilapia, *Oreochromis* sp., farm: A case study. J. Fish Dis. 32, 247-252.

Hetzel, U., König, A., Yildirim, A.Ö., Lämmler, C., Kipar, A., 2003. Septicaemia in emerald monitors (*Varanus prasinus* Schlegel 1839) caused by *Streptococcus agalactiae* acquired from mice. Vet. Microbiol. 95, 283-293.

Holt, J.G., Krieg, N.R., Snealli, P.H.A., Staly, J.T., Williams, S.T., 1994. Bergey's Manual of Determinative Bacteriology (9th ed.), Williams & Wilkins, Baltimore, Maryland.

Huang, B.F., Zou, L.L., Xie, J.G., Huang, Z.C., Li, Y.W., Li, A.X., 2013. Immune responses of different species of tilapia infected with *Streptococcus agalactiae*. J. Fish Dis. In press.

Inocente Filho, C., Müller, E., Pretto-Giordano, L., Bracarense, A., 2009. Histological findings of experimental *Streptococcus agalactiae* infection in Nile tilapias (*Oreochromis niloticus*). Brazilian Journal of Veterinary Pathology 2, 12-15.

Jafar, Q.A., Sameer, A.Z., Salwa, A.M., Samee, A.A., Ahmed, A.M., Al-Sharifi, F., 2008. Molecular investigation of *Streptococcus agalactiae* isolates from environmental samples and fish specimens during a massive fish kill in Kuwait bay. Pak. J. Biol. Sci. 11, 2500-2504.

Jiménez, A., Tibatá, V., Junca, H., Ariza, F., Verjan, N., Iregui, C., 2011. Evaluating a nested-PCR assay for detecting *Streptococcus agalactiae* in red tilapia (*Oreochromis* sp.) tissue. Aquaculture 321, 203-206.

Kawata, K., Anzai, T., Senna, K., Kikuchi, N., Ezawa, A., Takahashi, T., 2004. Simple and rapid PCR method for identification of streptococcal species relevant to animal infections based on 23S rDNA sequence. FEMS Microbiol. Lett. 237, 57-64.

Ke, D., Ménard, C., Picard, F.J., Boissinot, M., Ouellette, M., Roy, P.H., Bergeron, M.G., 2000. Development of conventional and real-time PCR assays for the rapid detection of group B streptococci. Clin. Chem. 46, 324-331.

Keefe, G.P., 1997. *Streptococcus agalactiae* mastitis: A review. Canadian Vet. J. 38, 429.

Kong, F., Martin, D., James, G., Gilbert, G.L., 2003. Towards a genotyping system for *Streptococcus agalactiae* (group B streptococcus): Use of mobile genetic elements in Australasian invasive isolates. J. Med. Microbiol. 52, 337-344.

Kong, F., Gowan, S., Martin, D., James, G., Gilbert, G.L., 2002. Serotype identification of group B streptococci by PCR and sequencing. J. Clin. Microbiol. 40, 216-226.

Kong, F., Lambertsen, L.M., Slotved, H., Ko, D., Wang, H., Gilbert, G.L., August 2008. Use of phenotypic and molecular serotype identification methods to characterize previously nonserotypeable group B streptococci. J. Clin. Microbiol. 46, 2745-2750.

Lämmler, C., Abdulmawjood, A., Weiß, R., 1998. Properties of serological group B streptococci of dog, cat and monkey origin. J. Vet. Med. B 45, 561-566.

Lancefield, R.C., 1933. A serological differentiation of human and other groups of hemolytic streptococci. J. Exp. Med. 57, 571-595.

Lancefield, R.C., 1934. A serological differentiation of specific types of bovine hemolytic streptococci (group B). J. Exp. Med. 59, 441.

Lancefield, R.C. and Hare, R., 1935. The serological differentiation of pathogenic and non-pathogenic strains of hemolytic streptococci from parturient women. J. Exp. Med. 61, 335

Lusiastuti, A.M., Textor, M., Seeger, H., Akineden, Ö., Zschöck, M., 2012. The occurrence of streptococcus agalactiae sequence type 261 from fish disease outbreaks of tilapia *Oreochromis niloticus* in Indonesia. Aquacult. Res. In press.

Maione, D., Margarit, I., Rinaudo, C.D., Masignani, V., Mora, M., Scarselli, M., Tettelin, H., Brettoni, C., Iacobini, E.T., Rosini, R., 2005. Identification of a universal group B streptococcus vaccine by multiple genome screen. Science 309, 148-150.

Mian, G., Godoy, D., Leal, C., Yuhara, T., Costa, G., Figueiredo, H., 2009. Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. Vet. Microbiol. 136, 180-183.

Mian, G.F., Godoy, D.T., Leal, C.A.G., Yuhara, T.Y., Costa, G.M., Figueiredo, H.C.P., 2009. Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. Vet. Microbiol. 136, 180-183.

Musa, N., Wei, L.S., Musa, N., Hamdan, R.H., Leong, L.K., Wee, W., Amal, M.N., Kutty, B.M., Abdullah, S.Z., 2009. Streptococcosis in red hybrid tilapia (*Oreochromis niloticus*) commercial farms in Malaysia. Aquacult. Res. 40, 630-632.

Nocard, M. and Mollereau, R., 1887. Sur une mammite contagieuse des vaches laitieres. Ann. Inst. Pasteur 1, 109.

Olivares-Fuster, O., Klesius, P., Evans, J., Arias, C., 2008. Molecular typing of *Streptococcus agalactiae* isolates from fish. J. Fish Dis. 31, 277-283.

Pasnik, D.J., Evans, J.J., Klesius, P.H., 2009. Fecal strings associated with *Streptococcus agalactiae* infection in Nile tilapia, *Oreochromis niloticus*. Open Veterinary Science Journal 3, 6-8.

Pasnik, D., Evans, J., Klesius, P., 2005. Duration of protective antibodies and correlation with survival in nile tilapia *Oreochromis niloticus* following *Streptococcus agalactiae* vaccination. Dis. Aquat. Org. 66, 129.

Pereira, U., Mian, G., Oliveira, I., Benchetrit, L., Costa, G., Figueiredo, H., 2010. Genotyping of *Streptococcus agalactiae* strains isolated from fish, human and cattle and their virulence potential in Nile tilapia. Vet. Microbiol. 140, 186-192.

Phuektes, P., Mansell, P.D., Browning, G.F., 2001. Multiplex polymerase chain reaction assay for simultaneous detection of *Staphylococcus aureus* and streptococcal causes of bovine mastitis. J. Dairy Sci. 84, 1140-1148.

Plumb, J., Schachte, J., Gaines, J., Peltier, W., Carroll, B., 1974. Streptococcus sp. from marine fishes along the Alabama and northwest Florida coast of the Gulf of Mexico. Trans. Am. Fish. Soc. 103, 358-361.

Rantz, L.A., 1942. The seroligical and biological classification of hemolytic and nonhemolytic streptococci from human souces. J. Infect. Dis. 61-68.

Rasheed, V. and Plumb, J.A., 1984. Pathogenicity of a non-haemolytic group B streptococcus sp. in gulf killifish (*Fundulus grandis* Baird and Girard). Aquaculture 37, 97-105.

Ren, S.Y., Geng, Y., Wang, K.Y., Zhou, Z.Y., Liu, X.X., He, M., Peng, X., Wu, C.Y., Lai, W.M., 2013. *Streptococcus agalactiae* infection in domestic rabbits, *Oryctolagus cuniculus*. Transboundary and Emerging Diseases. In press.

Robinson, J.A. and Meyer, F.P., 1966. Streptococcal fish pathogen. J. Bacteriol. 92, 512-512.

Salvador, R., Muller, E.E., Freitas, J.C., Leonhadt, J.H., Pretto-Giordano, L.G., Dias, J.A., 2005. Isolation and characterization of *Streptococcus* spp. group B in Nile tilapias (*Oreochromis niloticus*) reared in hapas nets and earth nurseries in the northern region of parana state, Brazil. Ciência Rural 35, 1374-1378.

Schuchat, A., 1999. Group B streptococcus. The Lancet 353, 51-56.

Shuster, K.A., Hish, G.A., Selles, L.A., Chowdhury, M.A., Wiggins, R.C., Dysko, R.C., Bergin, I.L., 2013. Naturally occurring disseminated group B streptococcus infections in postnatal rats. Comp. Med. 63, 55-61.

Slotved, H., Kong, F., Lambertsen, L., Sauer, S., Gilbert, G.L., September 2007. Serotype IX, a proposed new *Streptococcus agalactiae* serotype. J. Clin. Microbiol. 45, 2929-2936.

Sommerset, I., Krossøy, B., Biering, E., Frost, P., 2005. Vaccines for fish in aquaculture. Expert Review of Vaccines 4, 89-101.

Stoffregen, D.A., Backman, S.C., Perham, R.E., Bowser, P.R., Babish, J.G., 1996. Initial disease report of *Streptococcus iniae* infection in hybrid striped (sunshine) bass and successful therapeutic intervention with the fluoroquinolone antibacterial enrofloxacin. J. World Aquacult. Soc. 27, 420-434.

Suanyuk, N., Kangheae, H., Khongpradit, R., Supamattaya, K., 2005. *Streptococcus agalactiae* infection in tilapia (*Oreochromis niloticus*). Songklanakarin Journal of Science and Technology 27, 307-319.

Suanyuk, N., Kong, F., Ko, D., Gilbert, G.L., Supamattaya, K., 2008. Occurrence of rare genotypes of *Streptococcus agalactiae* in cultured red tilapia *Oreochromis* sp. and Nile tilapia *O. niloticus* in Thailand-Relationship to human isolates? Aquaculture 284, 35-40.

Sun, Y., Kong, F., Zhao, Z., Gilbert, G.L., 2005. Comparison of a 3-set genotyping system with multilocus sequence typing for *Streptococcus agalactiae* (group B streptococcus). J. Clin. Microbiol. 43, 4704-4707.

Urwin, R. and Maiden, M.C., 2003. Multi-locus sequence typing: A tool for global epidemiology. Trends Microbiol. 11, 479-487

Verner-Jeffreys, D.W., Baker-Austin, C., Pond, M.J., Rimmer, G.S., Kerr, R., Stone, D., Griffin, R., White, P., Stinton, N., Denham, K., Leigh, J., Jones, N., Longshaw, M., Feist, S.W., 2012. Zoonotic disease pathogens in fish used for pedicure. Emerg. Infect. Dis. 18, 1006-1008.

Wang, B., Cai, S., Lu, Y., Zhang, X., Huang, Y., Jian, J., Wu, Z., 2012. A loop-mediated isothermal amplification method targeting the fbsB gene for rapid detection of *Streptococcus agalactiae*, the causative agent of streptococcosis in farmed fish. Israeli Journal of Aquaculture 64, 677.

Wang, K., Chen, D., Huang, L., Lian, H., Wang, J., Xiao, D., Geng, Y., Yang, Z., Lai, W., 2013. Isolation and characterization of *Streptococcus agalactiae* from Nile tilapia *Oreochromis niloticus* in China. African Journal of Microbiology Research 7, 317-323.

Wibawan, I.W., Lämmler, C., Smola, J., 1993. Properties and type antigen patterns of group B streptococcal isolates from pigs and nutrias. J. Clin. Microbiol. 31, 762-764.

Wilkinson, H.W., Thacker, L., Facklam, R., 1973. Nonhemolytic group B streptococci of human, bovine, and ichthyic origin. Infect. Immun. 7, 496.

Ye, X., Li, J., Lu, M., Deng, G., Jiang, X., Tian, Y., Quan, Y., Jian, Q., 2011. Identification and molecular typing of *Streptococcus agalactiae* isolated from pond-cultured tilapia in China. Fisheries Science 77, 623-632.

Yildirim, A.Ö., Lämmler, C., Weiß, R., 2002. Identification and characterization of *Streptococcus agalactiae* isolated from horses. Vet. Microbiol. 85, 31-35.

Yao, K., Poulsen, K., Maione, D., Rinaudo, C.D., Baldassarri, L., Telford, J.L., Sørensen, U.B.S., Members of the DEVANI Study Group, Kilian, M., 2013. Capsular gene typing of *Streptococcus agalactiae* compared to serotyping by latex agglutination. J. Clin. Microbiol. 51, 503-507.

Zadoks, R.N., Middleton, J.R., McDougall, S., Katholm, J., Schukken, Y.H., 2011. Molecular epidemiology of mastitis pathogens of dairy cattle and comparative relevance to humans. J. Mammary Gland Biol. Neoplasia 1-16.

Zamri-Saad, M., Amal, M.N.A., Siti-Zahrah, A., 2010. Pathological changes in red tilapias (*Oreochromis* spp.) naturally infected by *Streptococcus agalactiae*. J. Comp. Pathol. 143, 227-229.

Zappulli, V., Mazzariol, S., Cavicchioli, L., Petterino, C., Bargelloni, L., Castagnaro, M., 2005. Fatal necrotizing fasciitis and myositis in a captive common bottlenose dolphin

(*Tursiops truncatus*) associated with *Streptococcus agalactiae*. J. Vet. Diagn. Invest. 17, 617-622.

Zlotkin, A., Chilmonczyk, S., Eyngor, M., Hurvitz, A., Ghittino, C., Eldar, A., 2003. Trojan horse effect: Phagocyte-mediated *Streptococcus iniae* infection of fish. Infect. Immun. 71, 2318-2325.

CHAPTER 2

Genomic diversity of *Streptococcus agalactiae* from aquatic hosts

Christian M. J. Delannoy, Margaret Crumlish, Michael C. Fontaine, James F. Turnbull, Ruth N. Zadoks

Status: An alternative version of this manuscript has been published (Delannoy, C., Crumlish, M., Fontaine, M., Pollock, J., Foster, G., Dagleish, M., Turnbull, J., Zadoks, R., 2013. Human *Streptococcus agalactiae* strains in aquatic mammals and fish. BMC Microbiology 13, 41). This work has been orally presented at the 3rd Research Conference of the Institute of Aquaculture (Oct. 2012; Stirling, UK), the 15th International Conference on Diseases of Fish and Shellfish (Sept., 2011; Split, Croatia), and the Lunchtime Seminar series at the Institute of Aquaculture (June 2012) and the Moredun Research Institute (June 2011).

Contribution: The candidate participated in study design and acquisition of isolates, conducted the laboratory work and analysed the data under the supervison from all co-authors. Part of the laboratory work on sea mammal isolates was performed by Jolinda Pollok (MRI, UK). The candidate drafted the manuscript. All co-authors read, criticized and approved the final manuscript.

2.1 ABSTRACT

Streptococcus agalactiae is a well-recognized pathogen affecting both terrestrial and aquatic species. Whilst human and bovine isolates have been widely studied, the population structure of *S. agalactiae* infecting aquatic species remains to be investigated further. A collection of 33 piscine, amphibian and sea mammal isolates originating from several countries and continents was assembled, with the aim of exploring the population structure and potential host-specificity of aquatic *S. agalactiae*. Isolates were characterised using pulsed-field gel electrophoresis (PFGE), multi-locus sequence typing (MLST), and a standardised 3-set genotyping system comprising molecular serotypes, surface protein gene profiles and mobile genetic element profiles. Two major subpopulations were identified. The first subpopulation consisted of non-haemolytic isolates that belonged to sequence type (ST) 260 or 261, which are STs that have been reported only from teleosts. These isolates exhibited a low level of genetic diversity by PFGE and clustered with other STs that have been reported only in fish. Another common feature was the absence of surface protein genes or mobile genetic elements usually found in human or bovine isolates. The second subpopulation consisted of β-haemolytic isolates recovered from fish, frogs and sea mammals, and that exhibited medium to high genetic diversity by PFGE. STs identified among these isolates have previously been identified in strains associated with asymptomatic carriage and invasive disease in humans. Moreover, ST283 serotype III-4 and its novel single locus variant ST491 detected in fish from Southeast Asia shared a 3-set genotype identical to that of an emerging ST283 clone associated with invasive disease of adult humans in Asia. These observations suggest that some strains of aquatic *S. agalactiae* may present a zoonotic or anthroponotic hazard. STs found among the seal isolates (ST23) have also been reported from humans and numerous other host species, but never from teleosts. This work provides a rational framework for exploration of pathogenesis and host-associated genome content of *S. agalactiae* strains.

Keywords: *Streptococcus agalactiae*, Fish, Sea mammal, Aquatic epidemiology, Molecular epidemiology, MLST, PFGE.

2.2 INTRODUCTION

Streptococcus agalactiae has been documented as a cause of epizootics in fish from as early as 1966 (Robinson and Meyer, 1966). The bacterium, identified as an aetiological agent of septicaemia and meningo-encephalitis in freshwater and saltwater fish species, is now considered a major threat to the aquaculture industry, particularly for tilapia (*Oreochromis* sp.) (Hernández et al., 2009; Mian et al., 2009; Suanyuk et al., 2008; Ye et al., 2011). *Streptococcus agalactiae* has also been reported from other aquatic or semi-aquatic species, including those that are poikilothermic (ectothermic), such as frogs and crocodiles (Bishop et al., 2007; Kawamura et al., 2005), and those that are homeothermic, such as dolphins (Evans et al., 2006; Zappulli et al., 2005).

Cattle and humans are the main known reservoirs for *S. agalactiae*. In humans, the bacterium (commonly referred to as Group B Streptococcus or GBS) is a member of the commensal microflora of the intestinal and genito-urinary tracts, but it is also considered a major cause of neonatal invasive disease and an emerging pathogen in adults (Ip et al., 2006; van der Mee-Marquet et al., 2008). In cattle, *S. agalactiae* is a well-recognized causative agent of mastitis (Duarte et al., 2005; Sukhnanand et al., 2005). Numerous typing methods including randomly amplified polymorphic DNA (RAPD) typing, pulsed-field gel electrophoresis (PFGE) and ribotyping have been used for typing of bovine and human *S. agalactiae* (Duarte et al., 2005; Martinez et al., 2000; Sukhnanand et al., 2005). These comparative typing methods, although very useful for outbreak investigations, are not always sufficiently standardized and reproducible for comparison of results between distinct studies. Surveillance systems monitoring clonal spread and prevalence in populations over extended areas or periods of time require the use of definitive or library typing schemes with exchangeable databases (Van Belkum et al., 2007). Validated or promising library typing

methods for *S. agalactiae* include multi-locus sequence typing (MLST) and profiling of other genetic markers such as those implemented in the so-called 3-set genotyping sheme and comprising molecular serotypes (MS), surface protein genes (SPGs) and mobile genetic elements (MGEs) (Gilbert et al., 2006; Jones et al., 2003; Sun et al., 2005).

MLST is expected to be replaced by whole genome sequencing in the future, but it remains an important tool for molecular epidemiology because the MLST databases for individual pathogen species currently cover far more isolates than have been characterized based on whole genome sequencing (Spratt, 2012). Similarly, isolates that have been characterized by 3-set genotyping still outnumber isolates that have been characterized by whole genome sequencing, thus providing a less detailed but broader frame of reference than offered by whole genome sequences. MLST is used to define sequence types (STs) and clonal complexes (CCs). Most human isolates can be assigned to a limited number of clonal complexes, referred to as CC1, CC17, CC19 and CC23 (Jones et al., 2003; Sørensen et al., 2010). In contrast, most bovine isolates belong to another distinct CC, named CC67 (Bisharat et al., 2004). Members of predominantly host-associated CCs are occasionally identified in other host-species; for example, members of the human-associated CC23 have been reported in crocodiles and cattle (Bisharat et al., 2004; Bishop et al., 2007). The three-set genotyping allows for further differentiation of isolates belonging to the same ST (Sun et al., 2005). For example, ST283 isolates with molecular serotype III-4, C-α protein and C-α protein repeating units and the MGEs IS1381, ISSag1, and ISSag2 are associated with the emergence of GBS meningitis in adults in Southeast Asia (Ip et al., 2006).

A limited number of DNA-based typing methods have been used to analyse the genomic diversity of *S. agalactiae* isolates from fish (Olivares-Fuster et al., 2008; Pereira et al., 2010; Suanyuk et al., 2008). Evans et al. (2008) demonstrated that *S. agalactiae* epizootics in fish can be caused by at least 2 distinct subpopulations, of which only one is also associated with infections in humans. Control measures for aquatic species may therefore need to target

epidemiologically or genetically distinct subpopulations, which need to be defined further. To that aim, we characterized a collection of *S. agalactiae* isolated from aquatic host species, with the emphasis on tilapia, using PFGE, MLST and the 3-set genotyping system. Results were compared with those from aquatic and terrestrial species, in order to evaluate whether specific genotypes are associated exclusively with piscine or aquatic hosts or are shared between piscine, aquatic and terrestrial species.

2.3 MATERIALS & METHODS

2.3.1 Isolate collection and identification

A collection of 33 aquatic *S. agalactiae* isolates was assembled, including isolates from poikilothermic and homeothermic host species originating from multiple countries and continents (Table 2.2; table in appendix). Of 33 isolates, 13 represented 3 separate disease outbreaks (5, 4 and 4 isolates from outbreaks in Kuwait, Honduras and Colombia, respectively) with the remaining 20 isolates each representing a single, unrelated outbreak or death. Most fish isolates (n=18) originated from infections in farmed tilapia (*Oreochromis* sp.) from Honduras, Colombia, Costa Rica, Belgium, Thailand and Vietnam, whereas other fish isolates originated from infections in wild Klunzinger's mullets (n=5; *Liza klunzinger*) from Kuwait and ornamental fish (n=3; rosy barb, *Puntius conchonius*; golden ram, *Mikrogeophagus ramirezi*; and an undetermined fish species) from Australia. Sea mammal isolates (n=6) were recovered at post-mortem from lung swabs of 1 bottlenose dolphin (*Tursiops truncatus*) and 5 grey seals (*Halichoerus grypus*) that had stranded at various sites around the coast of Scotland. Finally, one amphibian isolate originating from an infected farmed frog (*Rana rugurosa*) in Thailand was made available for molecular characterisation.

Bacterial isolates were streaked onto 5% sheep blood agar plates (SBA; E&O Laboratories, Bonnybridge, United Kingdom) and grown aerobically at 28°C for 48h in order to assess purity and haemolysis properties. Isolates were Gram-stained and confirmed to belong to

Lancefield Group B by a group B-specific latex agglutination test (Slidex Strepto Plus B; bioMérieux, Marcy L'Étoile, France). Single colonies were then selected, used to inoculate Brain Heart Infusion (BHI) broth (Oxoid, Basingstoke, United Kingdom) and incubated with gentle shaking at 28°C for 12h (ß-haemolytic bacteria, fast growing) or 48h (non-haemolytic bacteria, slow growing). Chromosomal DNA was prepared using lysozyme and proteinase K (Furrer et al., 1991) and used to confirm species identity of *S. agalactiae* by polymerase chain reaction (PCR) using forward primer STRA-AgI (5'-AAGGAAACCTGCCATTTG-'3) and reverse primer STRA-AgII (5'-TTAACCTAGTTTCTTTAAAACTAGAA-3'), which target the 16S to 23S rRNA intergenic spacer region (Phuektes et al., 2001). PCR reactions were performed in a final 25µl volume containing 12.5µl of GoTaq Green Master Mix (Promega, Madison, USA), 0.25 µM of each primer and 2 µl of DNA template. The thermal cycling was conducted as follows: denaturation of DNA at 94°C for 5 min followed by 35 cycles of 94°C for 1 min, 58°C for 45 sec and 72°C for 30 sec, with a final step at 72°C for 7 min. The PCR products for all reactions were visualized following electrophoresis through a 1.5% (w/v) agarose gels containing Gel Red (Cambridge Bioscience, Cambridge, UK).

2.3.2 Comparative typing: PFGE

Bacterial cells were pelleted by centrifugation of 1ml of incubated BHI, re-suspended in 0.5ml of TE buffer (10mM Tris-HCl, 1mM EDTA), warmed to 56°C and mixed with 0.5ml of 2% (w/v) low-melting point agarose (Incert agarose; Lonza, Berkshire, United Kingdom) in TE buffer. The mixture was then pipetted into plug moulds. Each solidified plug was placed into 2ml of TE buffer containing 4mg of lysosyme (Sigma Aldrich, Dorset, United Kingdom) (2mg/ml) and incubated overnight at 37°C with gentle shaking. The buffer was then replaced with 2ml of ES buffer (0.5M EDTA–1% (w/v) *N*-lauroyl sarcosine [pH 8.0 to 9.3]) supplemented with 4mg of proteinase K (Promega, Southampton, United Kingdom) (2mg/ml) and incubated at 56°C for a minimum of 48h. Plugs were washed 6 times for 1h in

TE buffer at room temperature and with gentle shaking. A 1mm thick slice (4x4mm^2) from each plug was exposed to digestion with restriction endonuclease *Sma*I (20U in 100µl of fresh reaction buffer; New England Biolabs, Hitchin, United Kingdom) at 25°C overnight. PFGE was performed with a CHEF-mapper system (BioRad Laboratories, Hertfordshire, United Kingdom) in 0.5x TBE using 1% (w/v) agarose gel (Pulsed Field Certified Agarose, BioRad Laboratories, Hertfordshire, United Kingdom), a run time of 24h and switch time of 3-55 s (linear ramp) at 14°C. Patterns were observed by UV transillumination after SYBR Gold staining (Invitrogen, Paisley, United Kingdom). Computer-assisted data analysis and dendogram construction were performed with Phoretix 1D Pro software (TotalLab Ltd, Newcastle upon Tyne, United Kingdom). Similarities between PFGE patterns were also assessed visually using standard criteria (Van Belkum et al., 2007).

2.3.3 Library typing

2.3.3.1 Multilocus sequence typing

MLST consisted of the amplification by PCR and sequencing of seven housekeeping genes, namely *adhP*, *atr*, *glcK*, *glnA*, *pheS*, *sdhA*, and *tkt*, and using primers designed by Jones et al. (2003). PCRs were conducted as previously described (section 2.3.1), with exception of annealing temperature adjusted for each primer set. The PCR products were purified with the QIAquick PCR purification kit (Qiagen, Crowley, UK) and then sent to Eurofins MWG operon (Ebersberg, Germany) for sequencing using the cycle sequencing technology, which is a modification of the traditional Sanger sequencing method. Consensus sequences were trimmed in SeqMan (DNAStar, London, United Kingdom), and the *S. agalactiae* database (shttp://pubmlst.org/sa-galactiae/, Jolley et al., 2004) was used for allele and sequence type (ST) assignments. Sequences of novel alleles were submitted to the database curator for allocation of new allele numbers and STs; these are now available in the database. The unweighted pair group method in Phylodendron was used to visualize the relationship between allelic profiles (http://pubmlst.org/sagalactiae/). Analysis using the most up-to-date

complete allelic profile list (last accessed 20 September 2011) was carried out using eBURST version 3 (http://eburst.mlst.net/) in order to define relationships between STs. Finally, single locus variants (SLVs), double locus variants (DLVs) and triple locus variants (TLVs) were explored via ST query for STs that did not form part of a clonal complex in the eBURST analysis (http://pubmlst.org/sagalactiae/).

2.3.3.2 Three-set genotyping system

A 3-set genotyping system, comprising MS, SPG profiles and MGE profiles, was used. Serotyping was performed using multiplex-PCR assays (Poyart et al., 2007), with modifications as described by Evans et al. (2008). Non-typeable (NT) isolates were further investigated using other primer sets (Kong et al., 2005) and the serosubtyping of MS III isolates was performed as previously described (Kong et al., 2002). Surface-protein genes were determined by PCR and sequencing of PCR products, using primers targeting the *bca*, *bac*, *alp1*, *alp2*, *alp3* and *alp4* genes (Zhao et al., 2006). Finally, the prevalence of 7 MGEs, corresponding to 1 group II intron (GBSi1) and 6 insertion sequences (IS1381, IS861, IS1548, ISSa4, ISSag1 and ISSag2) was evaluated by PCR and amplicon identity was confirmed by sequencing of PCR products, using published primers (Kong et al., 2003; Suanyuk et al., 2008). PCRs and sequencing of products were conducted as previously described (sections 2.3.2 and 2.3.3.1). Primers used are listed in **Table 2.1**.

2.4 RESULTS

2.4.1 Isolate collection and identification

All isolates were Lancefield Group B, Gram-positive cocci appearing in pairs and chains. They were either β-haemolytic or non-haemolytic on SBA (**Table 2.2**). All were confirmed as *S. agalactiae* by species-specific PCR.

2.4.2 PFGE analysis

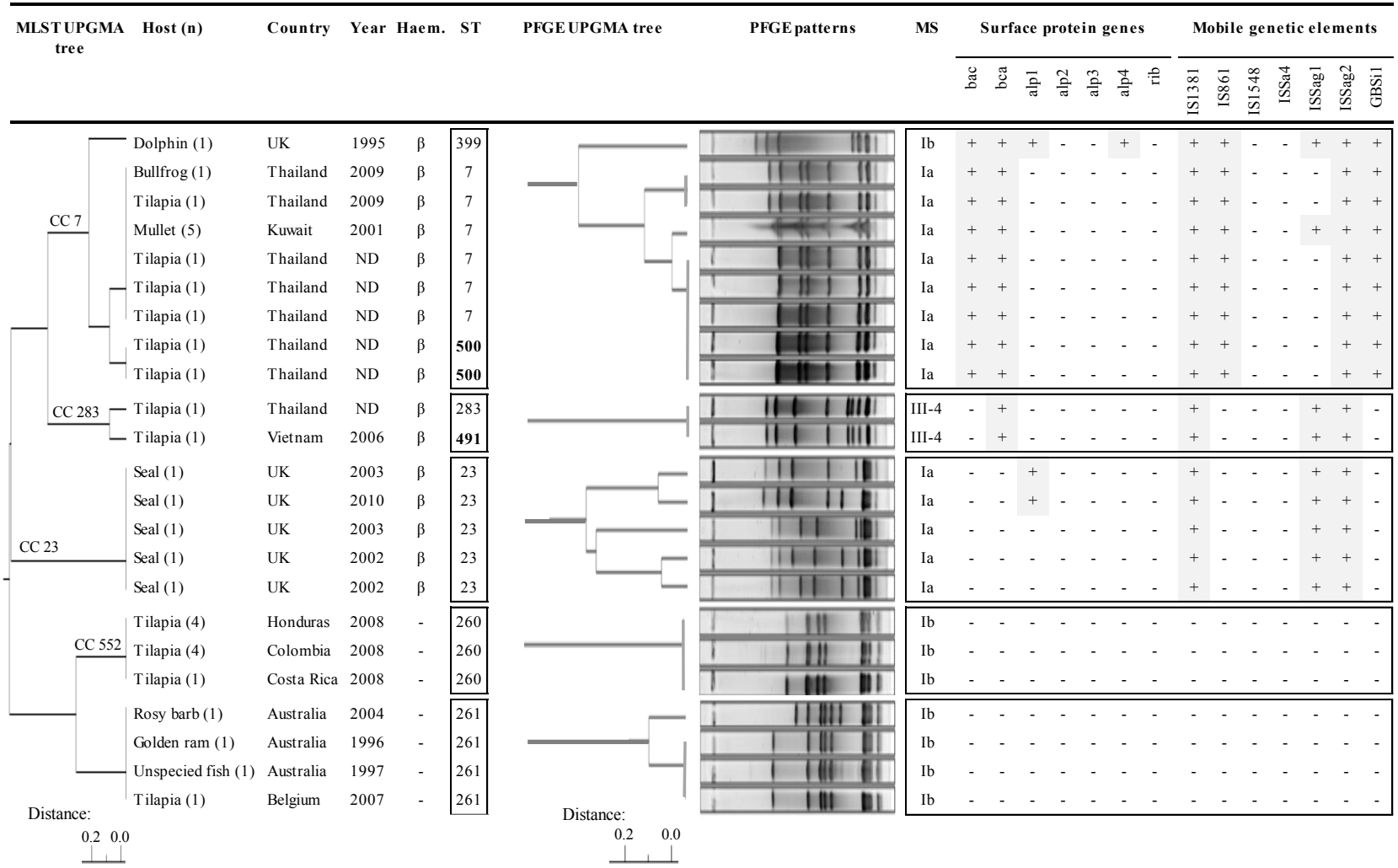
All isolates were typeable by *Sma*I macrorestriction and 13 pulsotypes were produced (**Table 2.2**). Pulsotypes were identical when multiple isolates from a single outbreak were analysed. In some cases pulsotypes were also identical between isolates from distinct outbreaks, countries or host species (*e.g.* tilapia isolates from Honduras, Colombia and Costa Rica; bullfrog and tilapia isolates from Thailand; **Table 2.2**). Despite efforts to identify potential epidemiological relationships between farms sharing the same pulsotype, *e.g*. through shared broodstock or feed companies, no such links could be identified and each outbreak is considered to be epidemiologically independent.

Table 2.1 List of primers used for the 3-set genotyping of *Streptococcus agalactiae*.

Target	Primer name	Primer sequence[Δ]	Genbank no.	Tm (˚C)	Reference
Molecular serotypes (n=9) - First set					
Ia (*cpsH*)	cps1aHS	7943 GGTCAGACTGGATTAATGGTATGC 7966	AB028896	61.00	Poyart et al., 2007
	cps1aHA	8463 GTAGAAATAGCCTATATACGTTGAATGC 8436	AB028896	60.70	Poyart et al., 2007
Ib (*cpsJ-K*)	cps1bJS	5627 TAAACGAGAATGGAATATCACAAACC 5652	AB050723	58.50	Poyart et al., 2007
	cpsIbKA	6396 GAATTAACTTCAATCCCTAAACAATATCG 6368	AB050723	59.60	Poyart et al., 2007
II (*cpsK*)	cps2KS	7449 GCTTCAGTAAGTATTGTAAGACGATAG 7475	AY375362	60.40	Poyart et al., 2007
	cps2KA	7845 TTCTCTAGGAAATCAAATAATTCTATAGGG 7816	AY375362	59.90	Poyart et al., 2007
III (*cpsI-J*)	cps1a/2/3IS	8273 TCCGTACTACAACAGACTCATCC 8295	AF163833	60.60	Poyart et al., 2007
	cps1a/2/3JA	10098 AGTAACCGTCCATACATTCTATAAGC 10073	AF163833	60.10	Poyart et al., 2007
IV (*cpsN*)	cps4NS	8991 GGTGGTAATCCTAAGAGTGAACTGT 9015	AF355776	61.30	Poyart et al., 2007
	cps4NA	9568 CCTCCCCAATTTCGTCCATAATGGT 9544	AF355776	63.00	Poyart et al., 2007
V (*cpsO*)	cps5OS	9852 GAGGCCAATCAGTTGCACGTAA 9873	AF349539	60.30	Poyart et al., 2007
	cps5OA	10552 AACCTTCTCCTTCACACTAATCCT 10529	AF349539	59.30	Poyart et al., 2007
VI (*cpsI*)	cps6IS	8276 GGACTTGAGATGGCAGAAGGTGAA 8299	AF337958	62.70	Poyart et al., 2007
	cps6IA	8762 CTGTCGGACTATCCTGATGAATCTC 8738	AF337958	63.00	Poyart et al., 2007
VII (*cpsM*)	cps7MS	5089 CCTGGAGAGAACAATGTCCAGAT 5111	AY376403	60.60	Poyart et al., 2007
	cps7MA	5459 GCTGGTCGTGATTTCTACACA 5439	AY376403	57.90	Poyart et al., 2007
VIII (*cpsJ*)	cps8JS	7016 AGGTCAACCACTATATAGCGA 7035	AY375363	55.90	Poyart et al., 2007
	cps8JA	7297 TCTTCAAATTCCGCTGACTT 7278	AY375363	53.20	Poyart et al., 2007
Molecular serotypes (n=9) - Second set					
Ia (*cpsH*)	IacpsHS	8176 ATACAGTTGTCGTAAAGAAGAAAAC 8200	AB028896	58.42	Kong et al., 2005
	IacpsHA	8536 TGTTTAGCTTTCCTACCAATATTAG 8512	AB028896	58.46	Kong et al., 2005
Ib (*cpsH*)	IbcpsHS	2986 TTTAGAAGTCCAGAATTTCATAGAGTC 3012	AB050723	60.82	Kong et al., 2005
	IbcpsHA	3265 CAAAGAAAGCCATTGCTCTCTG 3244	AB050723	64.91	Kong et al., 2005
II (*cpsK*)	IIcpsKS	7269 CTCCAGATGGTCTTTGTGAC 7288	AY375362	59.68	Kong et al., 2005
	IIcpsKA	7700 AAAATTGGTATATTTCCTCTTGAC 7677	AY375362	58.08	Kong et al., 2005
III (*cpsH*)	IIIcpsHS	7509 CCACATATGAGAATAAGACTTGC 7531	AF163833	59.15	Kong et al., 2005
	IIIcpsHA	7874 CCTAGTGATAGTACTTTGGTTTCTG 7850	AF163833	58.86	Kong et al., 2005
IV (*cpsH*)	IVcpsHS	7634 ATAGCCTTTTGACAGGTAGGTT 7655	AF355776	59.56	Kong et al., 2005
	IVcpsHA	7958 TGTAAATCATCTACACCCCC 7939	AF355776	58.67	Kong et al., 2005
V (*cpsH*)	VcpsHS	7616 GATGTTCTTTTAACAGGTAGATTACAC 7642	AF349539	59.04	Kong et al., 2005
	VcpsHA	7945 CTTTTTTATAGGTTCGATACCATC 7922	AF349539	58.56	Kong et al., 2005
VI (*cpsH*)	VIcpsHS	7648 TGTTTTTCTTACAAAGTGGAGTC 7670	AF337958	58.83	Kong et al., 2005
	VIcpsHA	7926 CCTGTTTTGTTTGATAGCTTCTC 7904	AF337958	60.66	Kong et al., 2005
VII (*cpsM*)	VIIcpsMS	5359 GTGCAATTAGAGGACAAAAATTTA 5382	AY376403	59.92	Kong et al., 2005
	VIIcpsMA	5651 CATCGAATCAGGAAAAATAGAT 5630	AY376403	58.56	Kong et al., 2005
VIII (*cpsJ*)	VIIIcpsJS	5687 ATCTCATGGCATGTCTGG 5704	AY375363	59.56	Kong et al., 2005
	VIIIcpsJA	5998 CATTCGAATAAACAATCTTATTGC 5975	AY375363	60.23	Kong et al., 2005
Molecular sero-subtypes (n=1)					
cpsE-G	cpsES3	6020 GTTAGATGTTCAATATATCAATGAATGGTCTATTTGGTCAG 6060	AF163833	66.40	Kong et al., 2002
	cpsGA1	6809 GTGTGTGATAACAATCTCAGCTTC 6781	AF163833	59.30	Kong et al., 2002
Surface protein genes (n=7)					
Protein Cα (*bca*)	bcaS	313 GCTTACATAGATTTATATGATGTAAAATTAGG 344	M97256	59.20	Zhao et al., 2006
	bcaA	637 CAGTACGACTTTCTTCCGTC 618	M97256	57.30	Zhao et al., 2006
Cα-like protein 2 (*alp2*)	alp23S	1254 CATGGAAGTGACAATTATGAAAG 1276	AF208158	55.30	Zhao et al., 2006
	alp2A	1603 CCACTGTAACTTTTACAGGAACTTC 1579	AF208158	59.70	Zhao et al., 2006
Cα-like protein 3 (*alp3*)	alp23S	1254 CATGGAAGTGACAATTATGAAAG 1276	AF208158	55.30	Zhao et al., 2006
	alp3A	1847 CTTTTGAACCATCTGGGTAAG 1827	AF208158	55.90	Zhao et al., 2006
Protein R (*rib*)	ribS	219 AGATACTGTGTTTGCAGCTGAAGTAA 244	U58333	60.10	Zhao et al., 2006
	ribA	577 CTATTTTATCTCTCAAAGCTGAAG 554	U58334	55.90	Zhao et al., 2006
Cα-like protein 4 (*alp4*)	alp4S	158 TGTTAGCAGCTGAAGTAGTTGAAG 181	AJ488912	59.30	Zhao et al., 2006
	alp4A	350 GCATAAACTTTTGAACCTTGTG 329	AJ488912	54.70	Zhao et al., 2006
Protein epsilon (*alp1*)	epsS	304 CTGTGTTTGCAGCTGAGGTG 323	U33554	58.80	Zhao et al., 2006
	epsA	719 CAGTACATCTTTTCGACTATCATCG 695	U33554	59.70	Zhao et al., 2006
Protein Cβ (*bac*)	bacS	1337 AAGGCTATGAGTGAGAGCTTGGAG 1360	X59771	62.70	Zhao et al., 2006
	bacA	1960 CTGCTCTGGTGTTTTAGGAACTTG 1937	X59772	61.00	Zhao et al., 2006
Mobile genetic elements (n=7)					
IS*1548*	IS1548S	143 CTATTGATGATTGCGCAGTTGAATTGGATAGTCGTC 178	Y14270	68.30	Kong et al., 2003
	IS1548A3	930 CGTGTTACGAGTCATCCCAATACCACGTAACTTATGCC 893	Y14270	71.60	Kong et al., 2003
IS*1381*	IS1381S1	272/818 CTTATGAACAAATTGCGGCTGATTTTGGCATTCACG 307/853	AF064785/AF367974	68.30	Kong et al., 2003
	IS1381A	881/1424 CTAAAATCCTAGTTCACGGTTGATCATTCCAGC 849/1392	AF064785/AF367974	67.00	Kong et al., 2003
ISS*a4*	ISSa4S	326 CGTATCTGTCACTTATTTCCCTGCGGGTGTCTCC 359	AF165983	71.90	Kong et al., 2003
	ISSa4A2	780 CGTAAAGGAGTCCAAAGATGATAGCCTTTTTGAACC 745	AF165983	68.30	Kong et al., 2003
IS*861*	IS861S	445 GAGAAAACAAGAGGGAGACCGAGTAAAATGGGACG 479	M22449	70.60	Kong et al., 2003
	IS861A2	1020 CAAACTCCGTCACATCGGTATAGCACTTCTCATAGG 985	M22449	70.60	Kong et al., 2003
GBSi1	GBSi1S1	721 CATCTCGGAACAATATGCTCGAAGCTTACAAGCAAGTG 758	AJ292930	70.50	Kong et al., 2003
	GBSi1A2	1161 CCAGGGACATCAATCTGTCTTGCGGAACAGTATCG 1127	AJ292930	71.80	Kong et al., 2003
IS*Sag1*	ISSag1S	449 GGCGACTGGTCGTTTAATCTCTGTTGAATTCTAGTTTC 486	AF329276	69.50	Suanyuk et al., 2008
	SSag1A	792 CTATCTCAGATTGTCAGGACACGGATTTTGTACTAGAC 755	AF329276	69.50	Suanyuk et al., 2008
IS*Sag2*	ISSag2S	1910 CAAATTCAATTTTGAAAGCACGATACGTACTTTCAGC 1946	AF329276	66.10	Suanyuk et al., 2008
	ISSag2A	2289 CCCTATTCCCAACCACTTAGACAGGCAATTTAAGC 2255	AF329276	69.50	Suanyuk et al., 2008

[Δ]Numbers represent the base positions at which primer sequences start and finish (numbering refers to the GenBank accession numbers for the corresponding genes).

Table 2.2 Overview of *S. agalactiae* origin, isolate number (n) and results of phenotypic and genotypic characterization



Results include analysis of haemolysis (Haem), multilocus sequence typing (MLST), pulsed-field gel electrophoresis (PFGE), molecular serotyping (MS), and profiling of surface protein genes and mobile genetic elements. Trees for MLST and PFGE results were constructed using unweighted pair group method analysis (UPGMA). Boxes enclose major clonal complexes (CCs) or sequence types (STs). STs shown in bold were first identified in the current study. ND: not determined.

2.4.3 MLST and eBURST analysis

Among the 33 *S. agalactiae* isolates, 8 STs were observed (**Table 2.2**), including 2 new STs (ST491 and ST500) that differed from existing STs through the presence of a novel allele. The new STs were single locus variants of known STs and resulted from single nucleotide changes.

STs were identical when multiple isolates from a single outbreak were analysed. In several cases, STs were also identical between isolates from epidemiologically unrelated animal deaths, outbreaks, countries or host species (*e.g.* ST261 was associated with three epidemiologically unrelated cases from Australia; ST260 was found in tilapia from Honduras, Colombia and Costa Rica; ST7 was found in bullfrog and tilapia from Thailand and mullet from Kuwait; **Table 2.2**).

E-burst analysis (**Figure 2.1**) showed that piscine isolates from Asia and the Middle-East (ST7 and its SLV ST500; ST283 and its SLV ST491) belonged to 2 related subgroups, CC7 and CC283. The amphibian isolate from Asia (ST7) and the bottlenose dolphin isolate from the UK (ST399) also belonged to CC7. All grey seal isolates from the UK belonged to ST23, which is not related to ST7 or ST283. Piscine isolates from Latin America (ST260) were part of the previously unreported CC552, which also includes ST257, ST259, ST552 and ST553. Based on analysis of DLVs and TLVs, ST261 and its SLV, ST246, may also be related to CC552, whilst ST258 is a TLV of C552 (**Figure 2.2**). STs in this cluster (CC552, ST246 and ST261, ST258) have only been reported from fish (Brochet et al., 2006; Evans et al., 2008; Figueiredo et al., 2011; this study). None of the STs in this cluster share SLVs, DLVs or TLVs with isolates of human, bovine or other non-piscine origin.

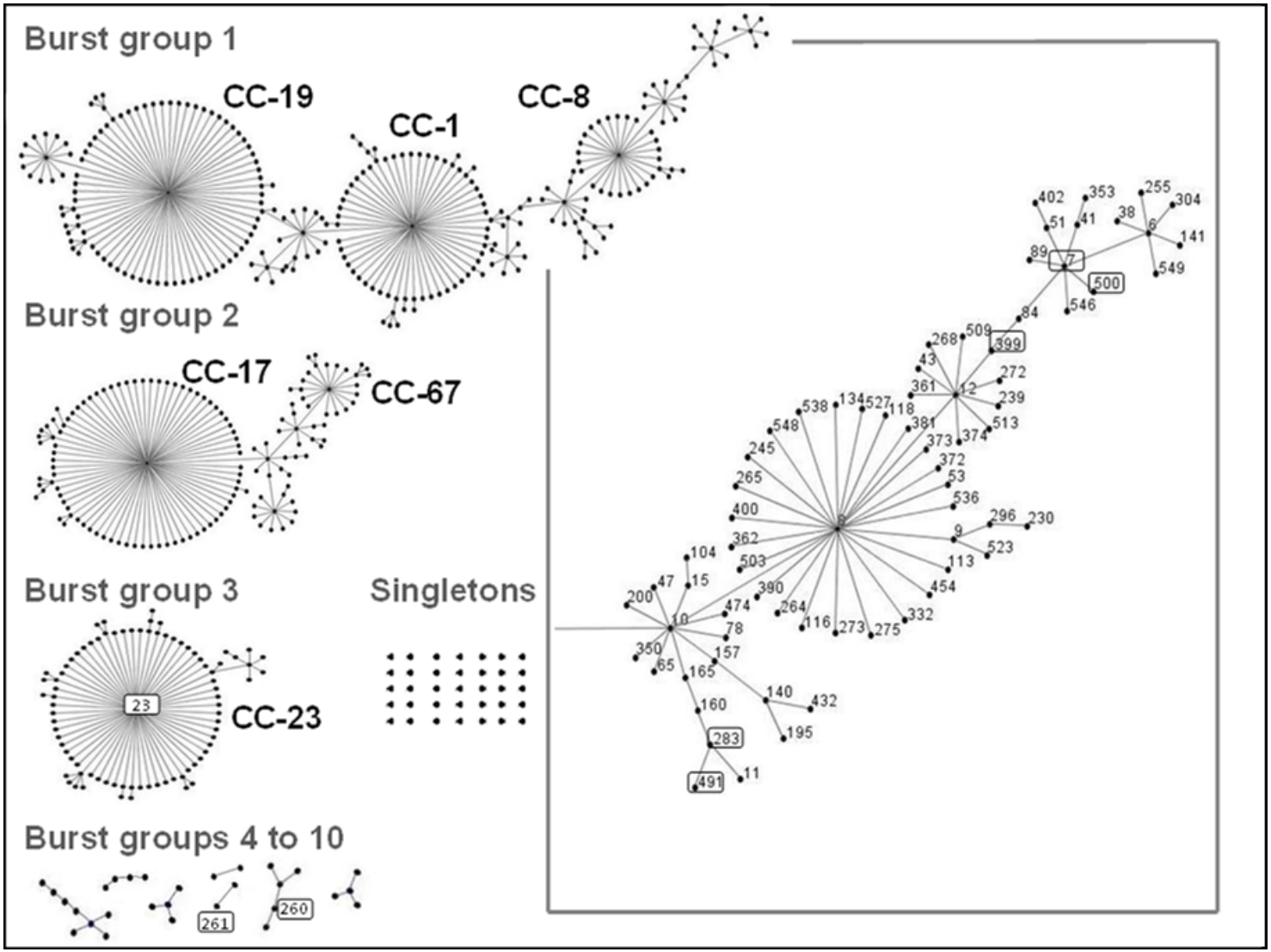


Figure 2.1 Population snapshot of *Streptococcus agalactiae*. All sequence types (STs) from the public MLST database (http://pubmlst.org/sagalactiae/; last accessed 20 September 2011; updated version available in Delannoy et al., 2013) are displayed as a single eBURST diagram. Burst groups and major clonal complexes (CCs) are identified. STs identified in the current study are shown (framed numbers). CC8 and related STs are enlarged for clarity (inset).

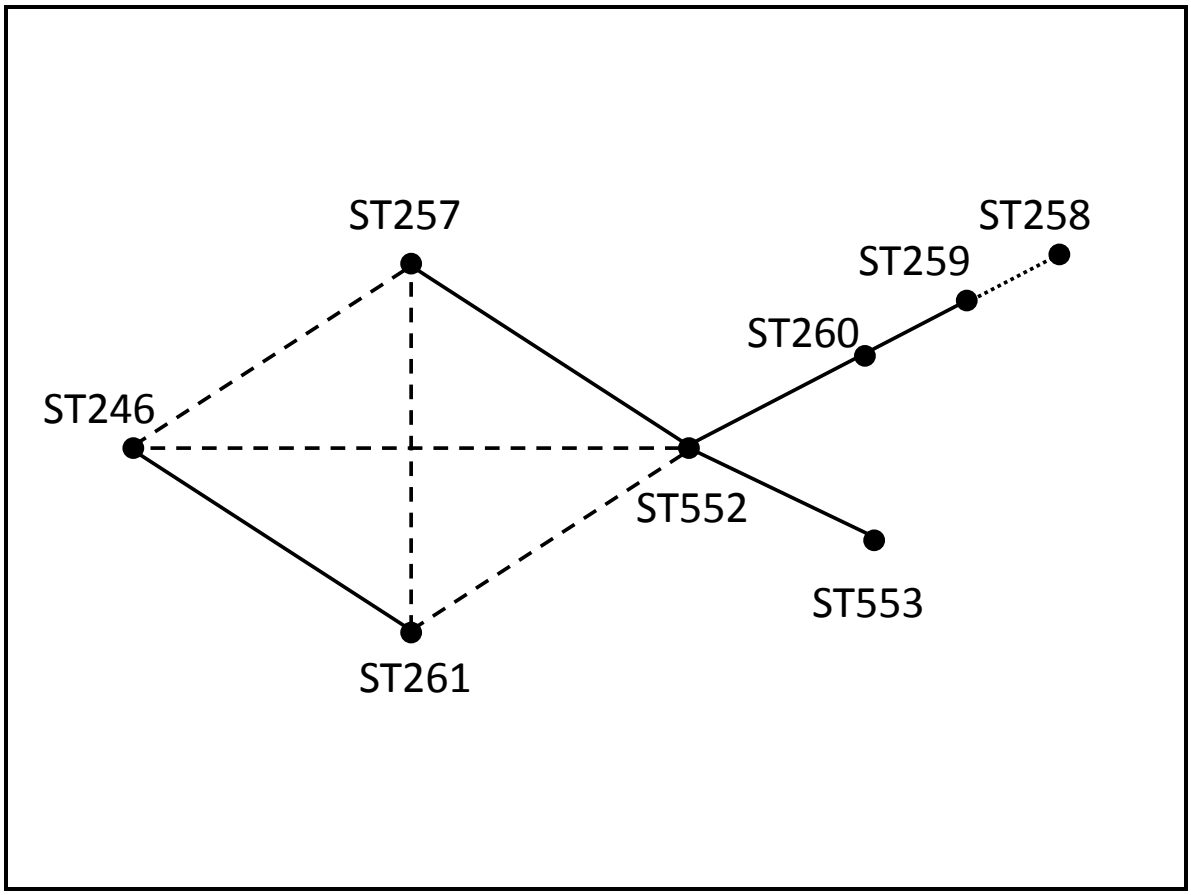


Figure 2.2 Relations between clonal complex 552 and other sequence types (STs) that have only been identified in fish. Continuous, dashed and dotted lines represent single, dual and triple locus variants, respectively, *i.e.* STs that differ at one, two or three of seven loci. STs in this cluster differ from all other STs in the database by more than 3 alleles.

2.4.4 Three-set genotyping system

Using primers from Poyart et al. (2007), isolates were identified as serotype Ia, Ib or NT. Further investigation of NT isolates with additional primer sets (Kong et al., 2002 and 2005) showed that the NT isolates belonged to serotype III sero-subtype 4. Based on the combination of MS, SPGs and MGEs, seven 3-set genotypes were distinguished (**Table 2.2**). Three set genotypes were identical when multiple isolates from a single outbreak were analysed. Piscine and amphibian isolates from Asia and the Middle-East and all mammalian isolates were positive for MGE IS1381 and ISSag2. IS861 was always found in combination with GBSiI and *vice versa* but rarely in combination with ISSag1, which was found in all mammalian isolates tested but only 7 of 27 non-mammalian isolates. When the Cβ protein gene (*bac*) was present, it was always found in association with the Cα protein gene (*bca*) and only in CC7 but *bca* was present in the absence of *bac* in CC283 (**Table 2.2**). Piscine isolates from Latin America (n=6), Australia (n=3) and Europe (n=1), all shared serotype Ib (**Table 2.2**) but none of the surface protein genes and mobile genetic elements investigated in this study was detected in any of these isolates.

2.4.5 Comparison across typing methods

All β-haemolytic isolates belonged to CCs that are also found in humans and carried at least 1 SPG or MGE (**Tables 2.2** and **2.3**). By contrast, all non-haemolytic isolates belonged to STs that are unique to fish and did not carry any of the SPGs and MGEs examined in this study (**Tables 2.2** and **2.3**). Each CC correlated with a PFGE cluster, although MLST could be more discriminatory than PFGE and *vice versa*. Concordance between PFGE and MLST was good. Within β-haemolytic isolates, PFGE clusters matched CCs and within non-haemolytic isolates, PFGE clusters matched STs. In some instances, PFGE was more discriminatory than MLST. For example, ST261 contained a PFGE type and subtype that differed from each other by 3 bands and ST23 comprised as many as 5 PFGE patterns. In other instances, MLST was more discriminatory than PFGE, *e.g.* in differentiation of

isolates in CC283. The number of PFGE types (**Table 2.2**) appeared to correlate with bacterial population size and host range of the corresponding ST or CC: the largest number of PFGE patterns were found in ST23 which is a very common ST and associated with a wide host range; an intermediate number of PFGE patterns were found in CC7, which encompasses less commonly found STs and is associated with a narrower range of host species; and the lowest number of PFGE patterns was found in ST260 and ST261, which have been reported in a limited number of isolates, all of which originate from teleosts (**Table 2.3**). Results from 3-set genotyping were also in agreement with origin, phenotyping, MLST and PFGE results. For example, the two closely related PFGE clusters within CC23 were associated with two closely related virulence gene profiles (**Table 2.2**). Within CC7, the dolphin isolate was the most divergent isolate based on PFGE typing, number of virulence genes and serotype. Within CC7, fish isolates from Kuwait had an extra MGE compared to the remainder of the fish isolates from Thailand (**Table 2.2**).

Table 2.3 Number of epidemiologically unrelated *S. agalactiae* reports in aquatic, mixed or terrestrial host species for STs identified in the current study.

MLST		Host species °						Key references
CC	ST	Aquatic	Mixed			Terrestrial		
		Fish	Amph	Rep	Mam	Bov	Hum	
7	7	10*	1	-	1*	1	19	Jones et al. (2003); Hery-Arnaud et al. (2007); Evans et al. (2008); Ye et al., 2011
	500	2	-	-	-	-	-	/
	399	-	-	-	1	-	1	Lartigue et al. (2009)
283	283	1	-	-	-	-	9	Ip et al. (2006); Salloum et al. (2010)
	491	1	-	-	-	-	-	/
23	23	-	-	1	4	7	>500	Bisharat et al. (2004); Brochet et al. (2006); Bishop et al. (2007); Hery-Arnaud et al., 2007
552	260	4	-	-	-	-	-	Evans et al. (2008)
NA	261	5	-	-	-	-	-	Evans et al. (2008)

° Amph, Amphibians; Rep, aquatic Reptiles; Mam, sea Mammals; Bov, Bovines; Hum, Humans

* The numerous isolates recovered from the 2001 fish kill of Kuwait bay are accounted as 1 single epidemiological event. The sea mammal isolate originated from a dolphin found dead during this outbreak.

2.5 DISCUSSION

Within the piscine *S. agalactiae* assessed in this study, 2 major subgroups were distinguished: the first one comprised β-haemolytic strains, with STs commonly associated with humans; the second one comprised non-haemolytic strains, with STs that appear to be unique to fish (**Table 2.2**).

Non-haemolytic *S. agalactiae* have previously been isolated from fish and frogs and have temporarily been misclassified as a new species, *Streptococcus difficile* (Kawamura et al., 2005). It is now widely recognized as a cause of disease in fish and has been reported from the USA, Honduras, Brazil and Israel (Evans et al., 2008; Mian et al., 2009; Robinson and Meyer, 1966). In our study, non-haemolytic piscine *S. agalactiae* were recovered from epizootics in Latin America (Honduras, Colombia, and Costa Rica), from an outbreak in Europe (Belgium) and from 3 unrelated ornamental fish in Australia. All our isolates from Latin America were of ST260 and serotype Ib, showing a wide dissemination of this specific clone in the region. The Belgian and Australian isolates were triple locus variants of ST260, namely ST261, and also of serotype Ib. ST260 and ST261 were first identified in non-haemolytic piscine *S. agalactiae* from Honduras and Israel, respectively (Evans et al., 2008). Like ST552, ST553, ST246, ST257, ST258 and ST259, they have never been reported from animals other than teleost fish (**Table 2.3**). These 8 piscine STs are genetically related and differ from all known STs associated with terrestrial animals at more than 3 loci, suggesting the existence of a fish-adapted CC of *S. agalactiae* (Figure 2). In the laboratory, members of the group that includes ST260 and ST261 do not grow well at 37°C, which may explain their current absence from homeothermic species. Among the ST260 and ST261 isolates from our study, we identified none of the SPGs and MGEs that are commonly found in human or bovine isolates (Table 1), providing additional evidence to support the idea that these strains are genetically distinct from human and bovine *S. agalactiae*. It is unknown whether fish-specific SPGs and MGEs exist in *S. agalactiae*.

The other isolates recovered in this study from fish, amphibians and sea mammals were all β-haemolytic. Our piscine isolates from the 2001 Kuwait bay outbreak belonged to ST7, as previously reported by Evans et al. (2008) based on analysis of other isolates from that outbreak. The ST7 strain from Kuwait is closely related to CC7 isolates from Thailand, with PFGE patterns and 3-set genotypes differing by 1 band and 1 MGE, respectively. PFGE and 3-set genotyping results showed that all ST7 and ST500 isolates from Thailand, which originated from tilapia and a bullfrog, share two closely related PFGE patterns and a single 3-set genotype (Ia-*bca*-*bac*-IS1381-IS861-ISSag2-GBSi1). This 3-set genotype has previously been detected among tilapia isolates from Thailand (Suanyuk et al., 2008) but we could find no reports of its occurrence among human or bovine isolates.

Sequence type 7 is part of a large group of human-associated clonal complexes, which includes CC1 and CC10 (**Figure 2.1**), but ST7 itself is relatively rare among isolates associated with colonization or infection in humans. By contrast, ST7 appears to be one of the major *S. agalactiae* clones associated with disease in fish (**Table 2.3**). Human ST7 can cause disease and death in fish, as shown by experimental challenge of tilapia with an ST7, serotype Ia isolate from a case of human neonatal meningitis (Evans et al., 2009). Fish-to-human transmission of *S. agalactiae* has never been documented, but does occur with *Streptococcus iniae*, a bacterial fish pathogen that is closely related to *S. agalactiae* (Facklam et al., 2005). It remains to be established if and how human-to-fish or fish-to-human transmission of *S. agalactiae* occurs under natural conditions.

Sequence type 283 and its SLV ST491, found among isolates from Thailand and Vietnam respectively, share a single PFGE pattern and 3-set genotype (III-*bca*-IS1381-ISSag1-ISSag2). ST283 and ST491 have not previously been reported from fish, but isolates with this 3-set genotype were recently found among piscine isolates from Thailand (Suanyuk et al., 2008). In humans, ST283 has been associated with meningitis and with skin, soft tissue and osteo-articular infections in adults (Ip et al., 2006; Salloum et al., 2010). Our piscine

ST283 isolates share the same 3-set genotype as ST283 isolates associated with invasive disease in human adults in Hong Kong (Ip et al., 2006). ST283 has evident invasive potential in human adults and in fish but it has not been identified among human-colonizing strains (Ip et al., 2006; van der Mee-Marquet et al., 2008). Ip et al. (2006) postulated that the emergence of the ST283 clone might be due to a process similar to the emergence of human ST17, which is a neonatal hyper-invasive clone that has been suggested to have evolved from the bovine-associated ST67 (Bisharat et al., 2004). The evolutionary linkage between ST17 and ST67 is controversial, however, as other authors argue that the common ancestor is more closely related to ST17 than to ST67 (Brochet et al., 2006), or even that the two STs are not closely related at all (Sørensen et al., 2010). Based on the population snapshot for *S. agalactiae*, ST283 belongs to a cluster of human-adapted CCs (**Figure 2.1**), making direct inter-species transmission, similar to that described for ST7 more likely than a long term evolutionary linkage between human and fish isolates of ST283.

The distinction between human-associated and fish-specific subpopulations of *S. agalactiae* in aquatic species is important for several reasons, including the assessment of the risk of zoonotic or anthroponotic transmission, investigation of biological mechanisms of host-adaptation, and development of disease control measures. The haemolotyic, human-associated and non-haemolytic, fish-specific subpopulations described in our study are compatible with biotypes 1 and 2, respectively. The biotype nomenclature is not commonly used in scientific publications, but a commercial vaccine against biotype 2 *S. agalactiae* in tilapia (AquaVac® STREP Sa, MSD Animal Health) has recently become available. This vaccine does not provide cross-protection against biotype 1 strains of *S. agalactiae*, providing further evidence that fish are affected by two distinct subpopulations of the pathogen.

Isolates from aquatic mammals in our study belonged either to ST399, which is part of CC7, or to ST23 (**Figure 2.1**). Unlike the dolphin isolate described by Evans et al. (2006), our

isolate was not associated with a fish kill but was obtained from a single stranding. It was recovered in the same geographic area as the seal isolates, all of which belonged to ST23. The number of isolates from aquatic mammals is too small to determine whether the association of dolphin isolates with CC7 and seal isolates with ST23 is due to coincidence or to the biology of aquatic mammals or *S. agalactiae*. ST399 has been reported once before, from a human isolate (Lartigue et al., 2009). The dolphin ST399 isolate had serotype Ib, whereas the rest of our isolates from CC7 as well as the human ST399 isolate shared serotype Ia. This suggests that capsular serotype switching may have occurred within this CC. ST23, which was found in seals, is the founder of a major CC mostly human-associated (**Figure 2.1**) that has however been identified in numerous aquatic and terrestrial hosts, including cattle, dogs and crocodiles (Brochet et al., 2006; Bishop et al., 2007) (**Table 2.3**). Within ST23, molecular serotypes Ia and III predominate, with serotype Ia linked to humans and serotype III primarily found in dairy cattle (Brochet et al., 2006; Sørensen et al., 2010). All seal isolates included in the current study had serotype Ia, suggesting a human origin. Among ST23 isolates in our study, two 3-set genotypes were distinguished (Table 1). One of these genotypes, Ia-alp1-IS1381-ISSag1-ISSag2, has previously been reported from ST23 isolates associated with necrotizing fasciitis in crocodiles (Bishop et al., 2007). Based on the screening of a subset of elements from the 3-set genotyping system (ISSag1 and 2 not screened) Ia-alp1-IS1381 also appears to be the most common profile found amongst human ST23. Despite its occurrence in a wide range of host species, ST23 has not been described in fish. It is unknown whether this is due to the limited number of fish isolates characterized to date or to host or pathogen biology.

2.6 CONCLUSION

Within aquatic *S. agalactiae*, 3 subgroups were distinguished. The first subgroup comprises non-haemolytic isolates that belonged to ST260 or ST261. Those STs have been reported only from teleosts and exhibited a low level of genetic diversity as detected by PFGE. A common feature is the absence of surface protein genes or mobile genetic elements

commonly found in human or bovine isolates. Although only ST260 is currently defined as a member of CC552, both ST260 and ST261 appear to be part of a fish-specific cluster of *S. agalactiae* strains. The second subgroup, which was found in fish, amphibia and sea mammals, comprises β-haemolytic isolates that exhibit medium to high genetic diversity by PFGE and 3-set genotyping. Their corresponding STs, together with some of the 3-set genotyping profiles, have been reported to cause disease in humans which raises the possibility of inter-species transmissions. The third group consists of ST23 isolates, which were found in all seals that were investigated. ST23 has been found in numerous terrestrial and aquatic species, including homeothermic and poikilothermic species, but to date it has never been reported from fish. This work provides a rational framework for exploration of pathogenesis and host-associated genome content of *S. agalactiae* strains.

2.7 REFERENCES

Bisharat, N., Crook, D.W., Leigh, J., Harding, R.M., Ward, P.N., Coffey, T.J., Maiden, M.C., Peto, T., Jones, N., 2004. Hyperinvasive neonatal group B streptococcus has arisen from a bovine ancestor. J. Clin. Microbiol. 42, 2161-2167.

Bishop, E.J., Shilton, C., Benedict, S., Kong, F., Gilbert, G.L., Gal, D., Godoy, D., Spratt, B.G., Currie, B.J., 2007. Necrotizing fasciitis in captive juvenile *Crocodylus porosus* caused by *Streptococcus agalactiae*: An outbreak and review of the animal and human literature. Epidemiol. Infect. 135, 1248-1255.

Brochet, M., Couvé, E., Zouine, M., Vallaeys, T., Rusniok, C., Lamy, M., Buchrieser, C., Trieu-Cuot, P., Kunst, F., Poyart, C., Glaser, P., 2006. Genomic diversity and evolution within the species *Streptococcus agalactiae*. Microb. Infect. 8, 1227-1243.

Delannoy, C., Crumlish, M., Fontaine, M., Pollock, J., Foster, G., Dagleish, M., Turnbull, J., Zadoks, R., 2013. Human *Streptococcus agalactiae* strains in aquatic mammals and fish. BMC Microbiol. 13, 41.

Duarte, R.S., Bellei, B.C., Miranda, O.P., Brito, M.A.V.P., Teixeira, L.M., 2005. Distribution of antimicrobial resistance and virulence-related genes among Brazilian group B streptococci recovered from bovine and human sources. Antimicrob. Agents Chemother. 49, 97-103.

Evans, J.J., Pasnik, D.J., Klesius, P.H., Al-Ablani, S., 2006. First report of *Streptococcus agalactiae* and *Lactococccus garviae* from a wild bottlenose dolphin (*Tursiops truncatus*). J. Wildl. Dis. 42, 561-569.

Evans, J.J., Bohnsack, J.F., Klesius, P.H., Whiting, A.A., Garcia, J.C., Shoemaker, C.A., Takahashi, S., 2008. Phylogenetic relationships among *Streptococcus agalactiae* isolated

from piscine, dolphin, bovine and human sources: A dolphin and piscine lineage associated with a fish epidemic in Kuwait is also associated with human neonatal infections in Japan. J. Med. Microbiol. 57, 1369-1376.

Evans, J.J., Klesius, P.H., Pasnik, D.J., Bohnsack, J.F., 2009. Human *Streptococcus agalactiae* isolate in Nile tilapia (*Oreochromis niloticus*). Emerg. Infect. Dis. 15, 74-776.

Facklam, R., Elliott, J., Shewmaker, L., Reingold, A., 2005. Identification and characterization of sporadic isolates of *Streptococcus iniae* isolated from humans. J. Clin. Microbiol. 43, 933-937.

Figueiredo, H.C.P, Godoy, D.T., Carvalho-Castro, G.A., Pereira, U.P., Leal, C.A.G., Leite, R.C. 2011. Virulence determinants in fish isolates of *Streptococcus agalactiae,* Diseases of Fish and Shellfish, Abstract of the 15th EAFP International Conference, 12–15 September 2011, Split, Croatia, O-086.

Furrer, B., Candrian, U., Hoefelein, C., Luethy, J., 1991. Detection and identification of listeria monocytogenes in cooked sausage products and in milk by in vitro amplification of haemolysin gene fragments. J. Appl. Microbiol. 70, 372-379.

Gilbert, G.L., Zeng, X., Zhao, Z., Gidding, H., Kong, F., 2006. Genotyping of group B Streptococcus and risk assessment for invasive infection. Int. Congr. Ser. 1289, 66-70.

Hernández, E., Figueroa, J., Iregui, C., 2009. Streptococcosis on a red tilapia, *Oreochromis* sp., farm: A case study. J. Fish Dis. 32, 247-252.

Hery-Arnaud, G., Bruant, G., Lanotte, P., Brun, S., Picard, B., Rosenau, A., van der Mee-Marquet, N., Rainard, P., Quentin, R., Mereghetti, L., 2007. Mobile genetic elements provide evidence for a bovine origin of clonal complex 17 of *Streptococcus agalactiae*. Appl. Environ. Microbiol. 73, 4668-4672.

Ip, M., Cheuk, E.S.C., Tsui, M.H.Y., Kong, F., Leung, T.N., Gilbert, G.L., 2006. Identification of a *Streptococcus agalactiae* serotype III subtype 4 clone in association with adult invasive disease in Hong Kong. J. Clin. Microbiol. 44, 4252-4254.

Jolley, K., Chan, M., Maiden, M., 2004. mlstdbNet - distributed multi-locus sequence typing (MLST) databases. BMC Bioinformatics 5, 86.

Jones, N., Bohnsack, J.F., Takahashi, S., Oliver, K.A., Chan, M., Kunst, F., Glaser, P., Rusniok, C., Crook, D.W.M., Harding, R.M., Bisharat, N., Spratt, B.G., 2003. Multilocus sequence typing system for group B streptococcus. J. Clin. Microbiol. 41, 2530-2536.

Kawamura, Y., Itoh, Y., Mishima, N., Ohkusu, K., Kasai, H., Ezaki, T., 2005. High genetic similarity of *Streptococcus agalactiae* and *Streptococcus difficilis*: *S. difficilis* Eldar et al. 1995 is a later synonym of *S. agalactiae* Lehmann and Neumann 1896 (approved lists 1980). Int. J. Syst. Evol. Microbiol. 55, 961-965.

Kong, F., Ma, L., Gilbert, G.L., 2005. Simultaneous detection and serotype identification of *Streptococcus agalactiae* using multiplex PCR and reverse line blot hybridization. J. Med. Microbiol. 54, 1133-1138.

Kong, F., Martin, D., James, G., Gilbert, G.L., 2003. Towards a genotyping system for *Streptococcus agalactiae* (group B streptococcus): Use of mobile genetic elements in Australasian invasive isolates. J. Med. Microbiol. 52, 337-344.

Kong, F., Gowan, S., Martin, D., James, G., Gilbert, G.L., 2002. Serotype identification of group B streptococci by PCR and sequencing. J. Clin. Microbiol. 40, 216-226.

Lartigue, M., Hery-Arnaud, G., Haguenoer, E., Domelier, A., Schmit, P., van der Mee-Marquet, N., Lanotte, P., Mereghetti, L., Kostrzewa, M., Quentin, R., 2009. Identification of *Streptococcus agalactiae* isolates from various phylogenetic lineages by matrix-assisted laser desorption ionization-time of flight mass spectrometry. J. Clin. Microbiol. 47, 2284-2287.

Martinez, G., Harel, J., Higgins, R., Lacouture, S., Daignault, D., Gottschalk, M., 2000. Characterization of *Streptococcus agalactiae* isolates of bovine and human origin by randomly amplified polymorphic DNA analysis. J. Clin. Microbiol. 38, 71-78.

Mian, G.F., Godoy, D.T., Leal, C.A.G., Yuhara, T.Y., Costa, G.M., Figueiredo, H.C.P., 2009. Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. Vet. Microbiol. 136, 180-183.

Olivares-Fuster, O., Klesius, P.H., Evans, J., Arias, C.R., 2008. Molecular typing of *Streptococcus agalactiae* isolates from fish. J. Fish Dis. 31, 277-283.

Pereira, U.P., Mian, G.F., Oliveira, I.C.M., Benchetrit, L.C., Costa, G.M., Figueiredo, H.C.P., 2010. Genotyping of *Streptococcus agalactiae* strains isolated from fish, human and cattle and their virulence potential in Nile tilapia. Vet. Microbiol. 140, 186-192.

Phuektes, P., Mansell, P.D., Browning, G.F., 2001. Multiplex polymerase chain reaction assay for simultaneous detection of staphylococcus aureus and streptococcal causes of bovine mastitis. J. Dairy Sci. 84, 1140-1148.

Poyart, C., Tazi, A., Réglier-Poupet, H., Billoët, A., Tavares, N., Raymond, J., Trieu-Cuot, P., 2007. Multiplex PCR assay for rapid and accurate capsular typing of group B streptococci. J. Clin. Microbiol. 45, 1985-1988.

Robinson, J.A. and Meyer, F.P., 1966. Streptococcal fish pathogen. J. Bacteriol. 92, 512.

Salloum, M., van der Mee-Marquet, N., Domelier, A., Arnault, L., Quentin, R., 2010. Molecular characterization and prophage DNA contents of streptococcus agalactiae strains isolated from adult skin and osteoarticular infections. J. Clin. Microbiol. 48, 1261-1269.

Sørensen, U.B.S., Poulsen, K., Ghezzo, C., Margarit, I., Kilian, M., July/August 2010. Emergence and global dissemination of host-specific *Streptococcus agalactiae* clones. MBio 1, e00178-10-e00178-18.

Spratt, B.G., 2012. The 2011 garrod lecture: From penicillin-binding proteins to molecular epidemiology. J. Antimicrob. Chemother. 67, 1578-1588.

Suanyuk, N., Kong, F., Ko, D., Gilbert, G.L., Supamattaya, K., 2008. Occurrence of rare genotypes of *Streptococcus agalactiae* in cultured red tilapia *Oreochromis* sp. and Nile tilapia *O. niloticus* in Thailand-Relationship to human isolates? Aquaculture 284, 35-40.

Sukhnanand, S., Dogan, B., Ayodele, M.O., Zadoks, R.N., Craver, M.P.J., Dumas, N.B., Schukken, Y.H., Boor, K.J., Wiedmann, M., 2005. Molecular subtyping and characterization of bovine and human *Streptococcus agalactiae* isolates. J. Clin. Microbiol. 43, 1177-1186.

Sun, Y., Kong, F., Zhao, Z., Gilbert, G.L., 2005. Comparison of a 3-set genotyping system with multilocus sequence typing for *Streptococcus agalactiae* (group B streptococcus). J. Clin. Microbiol. 43, 4704-4707.

Van Belkum, A., Tassios, P.T., Dijkshoorn, L., Haeggman, S., Cookson, B., Fry, N.K., Fussing, V., Green, J., Feil, E., Gerner-Smidt, P., Brisse, S., Struelens, M., for the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) Study Group on Epidemiological Markers (ESGEM), 2007. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clin. Microbiol. and Infec. 13, 1-46.

van der Mee-Marquet, N., Fourny, L., Arnault, L., Domelier, A., Salloum, M., Lartigue, M., Quentin, R., 2008. Molecular characterization of human-colonizing *Streptococcus agalactiae* strains isolated from throat, skin, anal margin, and genital body sites. J. Clin. Microbiol. 46, 2906-2911.

van der Mee-Marquet, N., Domelier, A.-S., Salloum, M., Violette, J., Arnault, L., Gaillard, N., Bind, J.-L., Lartigue, M.-F., Quentin, R., 2009. Molecular Characterization of Temporally and Geographically Matched *Streptococcus agalactiae* Strains Isolated from Food Products and Bloodstream Infections. Foodborne Pathog. Dis. 6, 1177-1183.

Zappulli, V., Mazzariol, S., Cavicchioli, L., Petterino, C., Bargelloni, L., Castagnaro, M., 2005. Fatal necrotizing fasciitis and myositis in a captive common bottlenose dolphin (*Tursiops truncatus*) associated with *Streptococcus agalactiae*. J. Vet. Diagn. Invest. 17, 617-622.

Zhao, Z., Kong, F., Gilbert, G.L., 2006. Reverse line blot assay for direct identification of seven *Streptococcus agalactiae* major surface protein antigen genes. Clin. Vaccine Immunol. 13, 145-149.

Ye, X., Li, J., Lu, M., Deng, G., Jiang, X., Tian, Y., Quan, Y., Jian., Q., 2011. Identification and molecular typing of *Streptococcus agalactiae* isolated from pond-cultured tilapia in China. Fish. Sci. 77, 623-632.

CHAPTER 3

Virulence and pathogenesis of fish-associated and non fish-associated subtypes of *Streptococcus agalactiae* in experimental challenges of Nile tilapia (*Oreochromis niloticus*)

Christian M.J. Delannoy, Ruth N. Zadoks, Jimmy Turnbull, Michael C. Fontaine, Margaret Crumlish.

Status: This manuscript is in preparation for submission in a relevant peer-reviewed journal. Part of the work has been presented in the Lunchtime Seminar series at the Institute of Aquaculture (June 2012; Stirling, UK).

Contribution: The candidate designed the study and conducted the experimental challenges with technical assistance from Niall Auchinachie (IOA, UK), Phuoc Nguyen Ngoc (IOA, UK) and Thi Luu (IOA, UK). Gross pictures were taken by Denny Conway (IOA, UK). Guidance and support for immunohistochemistry was provided by Mark Dagleish (MRI, UK) and Jeanie Finlayson (MRI, UK). Guidance for histopathological interpretations was provided by Prof. Hugh W. Ferguson (St George's University, West Indies) and Johanna Baily (MRI, UK). Biostatistic analyses resulted from collaboration with Mintu Nath (BioSS, UK). The candidate wrote the present manuscript with corrections and editing from Prof. Hugh W. Ferguson and listed co-authors.

3.1 ABSTRACT

Streptococcus agalactiae, a causative agent of septicaemia and meningo-encephalitis in fish, is recognised as a major threat for tilapia culture worldwide. In this study, the virulence of two strains of *S. agalactiae* was experimentally investigated by intra-peritoneal infection of Nile tilapia (*Oreochromis niloticus*), using an isolate originally recovered from fish and belonging to sequence type (ST) 260, and an isolate originating from a grey seal and belonging to ST23. Isolates of ST260 have only been reported from poikilotermic hosts, whereas isolates of ST23 have been recovered from numerous host species but never from teleosts. The clinical signs, the *in vivo* distribution of viable bacteria and bacterial antigens, and the gross and histopathological lesions that develop during the time course of the infection were investigated. The ST260 strain was highly virulent, whereas no major clinical sign or mortalities occurred in the fish challenged with the ST23. After injection, both strains however gained access to the bloodstream and viable bacteria were recovered from all organs under investigation. During the early stages of infection, bacteria were mostly found within the reticulo-endothelial system of the spleen and kidney. Thereafter, the ST260 demonstrated a particular tropism for the brain and the heart, but granulomatous inflammation and associated necrotic lesions were observed in all organs. ST23 was responsible for a mixed inflammatory response associated with the presence of bacteria in the choroid rete and in the pancreatic tissue mainly. After 7 days post-challenge and for both strain, the formation or containment of bacteria within granulomata or other encapsulated structures in various organs appeared to be a major component of the fish response. However, the load of viable bacteria remained high within organs of fish infected with ST260, suggesting that, unlike ST23, this strain is able to survive within macrophages and/or to evade the immune system of the fish. This work demonstrates that the lack of report of ST23 strains in fish is possibly not due to a lack of exposure (*e.g.* absence from the aquatic environment) but to a lack of virulence. The two strains, which differ in prevalence

and virulence in fish, provide an excellent basis to investigate genomic differences underlying the host-association of distinct *S. agalactiae* subpopulations.

Keywords: Challenge, *Streptococcus agalactiae*, tilapia, ST260, ST23

3.2 INTRODUCTION

Streptococcus agalactiae is a pathogenic bacterium affecting cultured and wild fish worldwide in a wide range of habitats including freshwater, estuarine and marine environments (Robinson et al, 1966; Eldar et al, 1994; Evans et al, 2008; Suanyuk et al, 2008; Musa et al, 2009; Mian et al, 2009). The bacterium is a cause of septicaemia and meningo-encephalitis in fish and associated pathological lesions have been described in both naturally acquired (Bowater et al., 2012; Baya et al., 1990; Zamri-Saad et al., 2010; Hernández et al., 2009) and experimental infections (Eldar et al., 1995; Rasheed et al., 1985; Suanyuk et al., 2008). *Streptococcus agalactiae*-associated lesions of necrotising fasciitis have also been reported in other aquatic or semi-aquatic hosts such as sea mammals and crocodiles, respectively (Bishop et al, 2007; Zappulli et al, 2005).

As well as aquatic animal species, *S. agalactiae* is also a significant pathogen in terrestrial species. Initially it was described in bovids in 1887 as a causative agent of mastitis (Nocard and Mollereau, 1887). However, in humans, the bacterium has highly variable roles from being a commensal organism of the intestinal and genito-urinary tracts to a pathogen responsible for invasive and often fatal infections in neonates (Yang et al., 2010). The genomic diversity of *S. agalactiae* isolates from human and bovine origins has been investigated using a broad range of molecular typing techniques. Library typing methods have recently been applied to isolates recovered from aquatic species with the aim of studying population structure and potential transmission between aquatic and terrestrial animal species. Studies were based on a housekeeping gene-based typing method, *i.e*. multi-locus sequence typing (MLST; Ye et al., 2011, Evans et al., 2008), or methods of typing genes from the accessory genome referred to as "3-set genotyping". The latter approach

consists of the identification of molecular serotypes and the profiling of surface protein genes and mobile genetic elements (Suanyuk et al, 2008). More recently, both techniques were simultaneously used to characterise isolates from a wide range of aquatic species (Chapter 2). This work showed that fish can be infected with 2 major and distinct bacterial subpopulations; the first was comprised of isolates with genotypes previously reported among humans and the second of isolates with genotypes that have, thus far, been recovered only from poikilotherms, primarily teleosts.

Isolates with genotypes found in both humans and fish species are beta-haemolytic, of serotype Ia or III-4, and by MLST classification belong to the human-adapted clonal complexes (CCs) 7 (ST7 and ST500) or 283 (ST283 and ST491). These strains have been responsible for major clinical disease outbreaks in wild mullet in Kuwait (ST7) and in farmed tilapia (*Oreochromis* sp.) in China (ST7), Thailand (ST7, ST500 and ST283) and Vietnam (ST491) (Chapter 2; Ye et al., 2011; Evans et al., 2008). Isolates of ST7 have also been recovered from a bottlenose dolphin (*Tursiops truncatus*) found dead during a fish epidemic in Kuwait (Evans et al., 2008). In experimental challenge studies, ST7 isolates recovered from either human, fish or dolphin sources can cause clinical disease and death in tilapia after intra-peritoneal injection, suggesting that ST7 strains are virulent in tilapia independent of their original host (Evans et al., 2002, 2006 and 2009). Evans et al. (2002 and 2006) used isolates recovered either from mullet, seabream or a dolphin recovered during a natural outbreak to experimentally infect 30-40g tilapia intra-peritoneally. This research determined that inoculation of 10^7 colony forming units (cfu) per fish of these strains later confirmed as ST7 (Evans et al., 2008) would cause 60 to 100% mortalities. In comparison, infection of tilapia using different doses (10^2 to 10^7 cfu per fish) of a ST7 of human origin caused clinical signs and mortalities ranging from 0 to 50% mortalities but a linear dose response was not seen (Evans et al., 2009).

Isolates with genotypes found exclusively in fish species are non-haemolytic, of serotype Ib, and belong by MLST classification to CC552 (ST246, 257, 258, 259, 260, 261). ST261 strains have been isolated from ornamental fish in Australia and from tilapia cultured in Belgium, Israel and Indonesia (Chapter 2; Evans et al., 2008; Lusiastuti et al., 2012). Using a strain originally named *S. difficile* and later found to be a non-haemolytic *S. agalactiae* of genotype ST261 (Kawamura et al., 2005; Evans et al., 2008), clinical disease has been reproduced experimentally in tilapia (Eldar et al., 1995). Other STs belonging to CC552 have been reported from outbreaks of clinical disease in tilapia in Latin American countries (Honduras, Colombia, Costa Rica, and Brazil) and in hybrid striped bass in Israel (Chapter 2; Evans et al., 2008). No experimental challenge has been reported with other STs belonging to CC552 such as ST260.

In a recent study, another ST (ST23) more frequently associated with human and cattle was repeatedly isolated from the lungs of Scottish grey seals (Chapter 2). Isolates belonging to ST23 have also been associated with an outbreak of necrotizing fasciitis in crocodiles (Bishop et al, 2007) and they have been isolated from dogs (Brochet et al., 2006). It would appear, therefore, that ST23 may be able to infect or colonise a wide variety of host species including aquatic animals, but unlike the ST7, it has never been reported in clinical disease of teleosts. It is unknown whether this is due to the relatively low number of fish isolates characterised by MLST to date, to lack of exposure of fish to ST23, or to biological constrains, *i.e.* the inability of ST23 to cause disease in fish even when exposure occurs.

Although the pathology of *S. agalactiae* infection in fish has been described previously, there are no reports of the development and progression of lesions during the course of the disease. Moreover, most pathological studies provided limited or no information on the genotype of the strain involved. The aims of this study were (1) to investigate the *in vivo* virulence of 2 different strains of *S. agalactiae* that differ in term of prevalence in tilapia (*Oreochromis niloticus*), one a strain belonging to a subpopulation that has uniquely been

associated with fish and the other a strain corresponding to a subpopulation never encountered in fish, and (2) to describe the clinical signs, the *in vivo* bacterial distribution, and the gross and histopathological lesions that develop during the time course of the experimental infection. Additionally, the tissue distribution of the bacterium during the challenge study was visualised by immunohistochemistry (IHC).

3.3 MATERIALS & METHODS

3.3.1 Bacterial strains

3.3.1.1 Bacterial strain selection

The *S. agalactiae* strains STIR-CD-17 and MRI Z1-201 used in this study were isolated from aquatic animals and selected based on previous epidemiological, phenotypic and genotypic characterisation (Chapter 2). STIR-CD-17 was isolated in 2008 during a clinical disease outbreak affecting farmed tilapia (*Oreochromis* sp.) in Honduras, whereas MRI Z1-201 was recovered by lung swab from a grey seal (*Halichoerus grypus*) found dead in 2003 in Scottish coastal waters. Post-mortem examination of the seal identified a systemic infectious process as the cause of death, but it is not certain whether the strain recovered and used in this study was directly involved. These strains are phenotypically and genotypically distinct. STIR-CD-17 is non-haemolytic, of serotype Ib and, based on multi-locus sequence typing (MLST), belongs to ST260 and clonal complex (CC) 552, which corresponds to a cluster of strains associated exclusively with disease in poikilotherms. MRI Z1-201 is beta-haemolytic, of serotype Ia, and belongs to ST23 and CC23, representing a subpopulation of *S. agalactiae* that has been reported from many host species but never from teleosts. Based on the standardized 3-set genotyping, STIR-CD-17 was negative for all surface protein genes and mobile genetic elements screened, whereas MRI Z1-201 contains genes encoding an alpha-like protein (*alp1*) and 3 insertion sequences (IS1381, ISSag1and ISSag2). MRI Z1-201 has the same genotype as a strain previously reported as a cause of necrotizing

fasciitis in crocodiles (Bishop et al, 2007). Both STIR-CD-17 and MRI Z1-201 will be referred to as ST260 and ST23 respectively in the remaining chapter.

3.3.1.2 Bacterial identification

Bacteria were recovered from frozen glycerol stocks onto 5% sheep blood agar plates (SBA; blood agar base [Oxoid Ltd, Basingstoke, UK] containing 5% [v/v] of defibrinated sheep blood [E&O laboratories Ltd, Bonnybridge, UK]) and grown aerobically at 28˚C for 72h to assess purity and haemolytic properties. Upon recovery, strain identity was further confirmed as Gram positive by a standard Gram staining protocol and as Lancefield Group B by a group B-specific latex agglutination test (Slidex Strepto Plus B; bioMérieux, Marcy L'Étoile, France). Isolates were further confirmed as *S. agalactiae* by polymerase chain reaction (PCR) based on amplification of the 16S to 23S rRNA intergenic spacer region using primers from Phuektes et al. (2001). Briefly, sample templates were prepared by suspending a single 48h growth colony into 50µl of TE buffer (10mM Tris-HCl, 1mM EDTA). PCR reactions were then performed as previously described (Chapter 2, section 2.3.1) with exception of the initial denaturation step which lasted for 15min for optimal cell lysis.

3.3.1.3 Inocula preparation for experimental infections

Single colonies from a pure culture were selected and aseptically inoculated into 4 ml of tryptone soya broth (TSB; Oxoid Ltd, Basingstoke, UK) and incubated aerobically for 8h (the ST23 strain, fast growing) or 24h (the ST260 strain, slow growing) at 28˚C with gentle shaking (140 rpm). A 36 ml volume of TSB was then seeded with 4 ml of each culture and incubated overnight (16h, 28˚C, 140 rpm). The bacterial suspensions were then centrifuged at 3,500 rpm for 15 min and the supernatant was removed; this process had to be repeated numerous times to enhance bacterial concentration for ST260 strain since this bacterium produced a fragile cell pellet. The cell pellets were resuspended in sterile 0.85% saline and the OD_{600nm} adjusted to give an OD value of 1 corresponding to approximately 10^9 viable

colony forming units (cfu) per ml for the ST260 strain and 10^8 viable cfu/ml for the ST23 strain, as determined using the method of Miles and Misra, (1938). Serial 10-fold dilutions were eventually performed using sterile 0.85% saline to obtain the desired bacterial concentrations for the challenge study. The exact concentration for all inocula was confirmed using the viable cell method of Miles and Misra, (1938).

3.3.2 The fish

3.3.2.1 Origin and screening of the fish

Nile tilapia (*Oreochromis niloticus*) were reared in the Tropical Aquarium at the Institute of Aquaculture (University of Stirling, UK) and maintained in a re-circulating water system in aquaria at 28±2°C under constant aeration and filtration. The fish were fed twice daily with commercial pellets (Skretting Trout Nutra 25) and kept on a 12h light: 12h dark cycle. Prior to use for bacterial challenge, a subset of the fish (n=3) were sampled randomly for bacterial recovery. Bacterial recovery was performed as previously described (Crumlish et al., 2010) and consisted of the insertion of a sterile plastic bacteriological loop (Fisher Scientific, Loughborough, UK) into the kidney followed by the inoculation of a tryptone soya agar plate (TSA; Oxoid Ltd, Basingstoke, United Kingdom). Plates were incubated at 28˚C for 72h and then examined for the presence of bacterial colonies. No growth was observed and apparently healthy animals weighing 40±5 g were used for the subsequent challenge study.

3.3.2.2 General live animal procedures and maintenance of infected fish

Apparently healthy fish were transferred to the Aquatic Research Challenge Facility (Institute of Aquaculture, University of Stirling, UK) where they were subjected to bacterial challenge by intra-peritoneal (i.p.) injection. For this, the fish were first lightly anaesthetized by immersion in a benzocaine bath (Sigma-Aldrich, Irvine, UK) and then each fish received the required inoculum by i.p. injection. For i.p. injection, the needle mounted on a 1ml syringe was inserted cephalad into the midline of the abdomen just posterior to the pectoral fins and 0.1ml of the inoculum was injected intra-peritoneally. Fish were fasted for 24h

prior to injection and for 12h following the injection, at which time daily feeding was resumed.

All fish from different experimental groups were kept in separate 10L aquaria with separate flow-through water systems, a temperature of 28 ± 2˚C and a 12h light: 12h dark cycle. Fish were monitored at least 3 times daily for signs of disease and death. All moribund and dead fish were removed and moribund fish were euthanized with an overdose of benzocaine. All animal experiments were conducted at the Institute of Aquaculture (University of Stirling, UK) in accordance with the Animals (Scientific Procedures) Act 1986.

3.3.3 Experimental challenge studies

3.3.3.1 Bacterial isolates passages

Each strain was passaged through fish prior to challenge in order to enhance their virulence post-storage (Eldar et al., 1995). Inocula prepared as previously described and containing a high concentration of bacteria were injected into a single fish. This procedure was performed once with ST260 as the fish died within 3 days post-inoculation (p.i.). Fish challenged with ST23 had no mortality and the procedure was successively repeated 3 times with the fish euthanized after 3 days p.i. These fish were sampled for bacterial recovery from a range of organs including the brain, eyes, spleen, heart, kidney and liver onto tryptone soya agar (TSA; Oxoid Ltd, Basingstoke, UK). One colony isolated from the brain after the 1st passage (ST260) or the 3rd passage (ST23) was sub-cultured onto SBA and used for further experimental infections after its identity was confirmed as previously described (section 3.3.1.2.).

3.3.3.2 Pre-challenge studies

Two pre-challenge studies were performed; in one, fish were challenged with ST260 and in the other, fish were challenged with ST23.

A dose-dependent experimental challenge was performed using strain ST260 to compare survival rates of tilapia infected with different bacterial doses and to provide a description of the clinical signs. For this, 3 different doses (10^2, 10^5 and 10^7 cfu per fish) were used to challenge 10 fish per dose and 10 further fish were administered 0.85% sterile saline solution as negative controls. Each treatment group consisted of 10 fish per tank. The experiment lasted 16 days, and during this time moribund fish were euthanized as well as the fish remaining at day 16 p.i. Euthanized fish were aseptically sampled for bacterial recovery from the kidney (Crumlish et al. 2010). For statistical analysis, Kaplan-Meier curves were used to compare survival rates of tilapia challenged with different bacterial doses; differences between doses were compared using a log rank test in Graph Pad Prism version 5 (Graph Pad Software, San Diego, California, USA).

For ST23, the objectives were to evaluate the development of morbidity or mortality and successful bacterial recovery from tilapia challenged with a single high bacterial dose. Based on results from the passages (section 3.3.3.1), a dose-dependent challenge study was not considered necessary for this strain. In this study, 20 fish were injected with a single bacterial concentration (1 x 10^7 cfu per fish) and allocated in 2 groups (10 fish per tank). Additionally, 10 negative control fish received 0.1ml of 0.85% sterile saline by i.p. injection and were allocated to a separate tank. For the fish receiving ST23, 1 group of fish (n=10) was sampled as described previously at day 7 p.i., while the second group of fish (n=10) and negative control fish (n=10) were sampled at day 16 p.i.

The identity of bacterial isolates recovered from both studies was confirmed as *S. agalactiae* based on previously described methods (section 3.3.1.2.).

3.3.3.3 Challenge studies

3.3.3.3.1 Challenge study design

After the pre-challenge studies, 2 further challenge studies were performed: in one, fish were challenged with ST260 and in the other, fish were challenged with ST23. A total of

135 fish were used per challenge study, with 120 fish administered the appropriate bacterial strain and 15 fish injected with 0.85% sterile saline solution as negative controls. The fish receiving ST260 were injected with 10^5 cfu/fish, whereas the fish receiving ST23 were injected with 10^7 cfu/fish. The fish density was 10 fish per tank for the challenged animals and 15 fish per tank for the negative controls. A total of 13 tanks were therefore used for each challenge study. Each study was terminated 14 days after bacterial challenge.

During the challenge studies, moribund fish were euthanized and examined grossly. At necropsy, selected tissues were sampled from brain, heart, liver, spleen, kidney and eye and preserved in 10% buffered formalin for histopathology. To study the progression of infection and pathology, all fish from 3 tanks, plus 3 negative control fish, were sacrificed at day 1, 3, 7 and 14 p.i. Sacrificed fish were processed in the same manner as moribund fish to obtain pathological information and histopathology material. In addition, 3 fish were randomly selected from each tank for viable bacterial counts from the blood (0.1ml/fish collected with heparinised syringes [Sigma Aldrich, Dorset, UK]), brain, heart, liver, spleen, kidney and eye.

3.3.3.3.2 Bacterial counts from organs

Samples of aseptically harvested blood, brain, heart, liver, spleen, kidney and eye were placed in sterile tubes, weighed and stored on ice until all samples had been collected at any one time-point. Organs were then manually homogenized using disposable pestles (VWR International Inc., Poole, UK) and diluted 1:10 with sterile 0.85% saline solution (w/v). The homogenates were further diluted serially in 10-fold steps to 10^{-6}. Triplicate samples (20µl) from each serial dilution (10^{-6} to 10^{-1}) were dropped onto TSA plates, allowed to dry at room temperature and then incubated aerobically at 28˚C for 48h. Total viable bacterial counts were performed using a dilution level giving, when possible, a range of 10 to 50 colonies per spot and the bacterial load per gram of tissue or per millilitre of blood was

calculated for each organ from each fish. This was then used to determine the mean and standard deviation (SD) for each organ at each of the sampling time points.

For statistical analysis, all viable colony counts data were $\log_{10}$ transformed. A linear mixed model was built using the 12th Edition of the GenStat statistical package (Payne et al., 2009). In this model, tanks and fish within tanks were used as random effects whereas day, organs and treatments (ST260 and ST23) were used as fixed effects. To evaluate the mean difference of bacterial counts (1) between organs at each time point of the same treatment, (2) between time points for a specified organ and treatment, or (3) between treatments for the same organs at the same time point, t-test were conducted. The P-values obtained from the t-tests were adjusted using a false discovery rate (FDR) approach to take into account multiple comparisons of treatment means.

3.3.3.3.3 Isolates identity confirmation and strain typing

The identity of bacterial isolates recovered was confirmed as *S. agalactiae* based on previously described methods (section 3.3.1.2.). Colonies confirmed as non-*S. agalactiae* were further identified to the genus level using standard Gram-staining protocol and biochemical tests (API 20E; bioMerieux, Marcy l'Etoile, France), followed by partial 16s rDNA sequencing of PCR product. For PCRs, sample templates were prepared by suspending a single 48h growth colony into 50µl of TE buffer (10mM Tris-HCl, 1mM EDTA). PCR amplification of a 700 base segment from the 16s rDNA region was then performed. Reactions were performed in a final 25 µl volume containing 12.5µl of GoTaq Green Master Mix (Promega, Madison, USA), 0.25 µM of primers PEU7 (5'-GCAAACA-GGATTAGATACCC-3') (Rothman et al., 2002) and DG74 (5'-AGGAGGTGATCCAA-CCGCA-3') (Greisen et al., 1994) and 2 µl of the sample template, with the rest of the volume made up of DNA-free water. The cycling conditions were as follows: cell lysis and initial denaturation of DNA at 94°C for 15min followed by 30 cycles of 94°C for 1min, 50°C for 1min and 72°C for 1.5min, with a final step at 72°C for 7min. The PCR products

were visualized on a 1.5% (w/v) agarose gel containing gel red (Cambridge Bioscience, Cambridge, UK), purified with the QIAquick PCR purification kit (Qiagen, Crowley, UK) and then sent to Eurofins MWG operon (Ebersberg, Germany) for sequencing using primers PEU7 and P3SH (5'-CTACGGTTA-CCTTGTTACGACTT-3') (Ralyea et al., 1998). Microbial identification was accomplished through Basic Local Alignment Search Tool (BLASTN; available from: http://www.ncbi.nlm.ni-h.gov/BLAST; accessed 25 Nov. 2012) analysis to compare the obtained sequence data with 16s rDNA sequence data stored in GenBank (Altschul et al., 1990).

The original challenge strains (ST260 and ST23) and a subset of confirmed *S. agalactiae* colonies recovered during the challenge studies (39 colonies from the ST260-challenged fish; 21 colonies from for the ST23-challenged fish) were further typed by pulsed field gel electrophoresis (PFGE) as described in Chapter 2 (section 2.4.2).

3.3.3.3.4 Histopathology

Samples were fixed for a minimum of 24h in 10% neutral buffered formalin and routinely processed prior to being embedded in paraffin wax by a maximum of 4 days post-fixation This maximum fixation time was deliberately chosen to avoid interferences with antigen recognition by immunohistochemistry (IHC; Ramos-Vara, 2005). Formalin-fixed paraffin-wax-embedded (FFPE) tissue sections (4 μm thick) were mounted on glass microscope slides, dried overnight at 37˚C, dewaxed in xylene and rehydrated through graded alcohols. Sections were stained with haematoxylin and eosin (H&E). Duplicate sections were subjected to Gram staining and IHC to aid in the localisation of the bacteria in the tissues.

3.3.3.3.5 Immunohistochemistry

Semi-serial tissue sections from all samples were mounted on Superfrost^{TM} Plus slides (Menzel-Gläser, Braunschweig, Germany), dewaxed and rehydrated as previously described (section 3.3.3.3.4), and subjected to IHC to determine the presence and distribution of *S. agalactiae*. For antigen retrieval, slides were placed in 0.2% citrate buffer pH 6.0 and

autoclaved at 121˚C for 10 min, allowed to cool to 50˚C and then washed in water for 5 min. Sections were then treated with 3% hydrogen peroxide in methanol (v/v) for 20 min to block endogenous peroxidase activity and washed in phosphate-buffered saline (PBS). Non-specific binding of the secondary antibody was blocked by incubation with 25% normal goat serum (Vector Laboratories, Peterborough, UK) diluted in PBS for 30 min at room temperature. Subsequently to this, rabbit anti-*S. agalactiae* polyclonal antibody (code AB53584; Batch No. GR56848-1; Abcam, Cambridge, UK) diluted 1/200 in PBS was applied to all slides and incubated for 18h at 4°C. Slides were then washed 3 times for 5 min in PBS prior to adding 100 µl of goat anti-rabbit antibody conjugated to horseradish peroxidase (HRP) (Dako UK Ltd., Cambridgeshire, UK) for 30 min at room temperature. Slides were washed as above and then incubated with Nova Red chromagen (Vector Laboratories, Peterborough, UK) for 12 min. Finally, slides were washed in water, counterstained with haematoxylin Z (Cellpath, Newton, UK), "blued" with Scott's tap water substitute (Cellpath, Newton, UK), dehydrated in graded alcohols, cleared and mounted. Negative controls were prepared by substituting the primary antibody with normal rabbit serum produced locally at the Moredun Research Institute (Penicuik, UK) at the same dilution.

3.4 RESULTS

3.4.1 Pre-challenge studies

Fish challenged with ST260 had a dose-dependent mortality response (**Figure 3.1**). Most deaths occurred between days 1 and 7 p.i. and at termination of the experiment, only 1 fish remained. No clinical signs of disease were observed from the fish that died within 48h. The first clinical signs appeared 2 to 5 days p.i. depending on the dose administered and consisted of lethargy and anorexia, which were always followed by signs of ataxia 24 to 48h prior to death. Most fish remaining after day 7 (fish groups receiving 10^2 and 10^5 cfu/fish) exhibited uni- or bilateral exophthalmia together with corneal opacity and peri-ocular

haemorrhage. *Streptococcus agalactiae* was recovered from the kidney of 100% of euthanized tilapia.

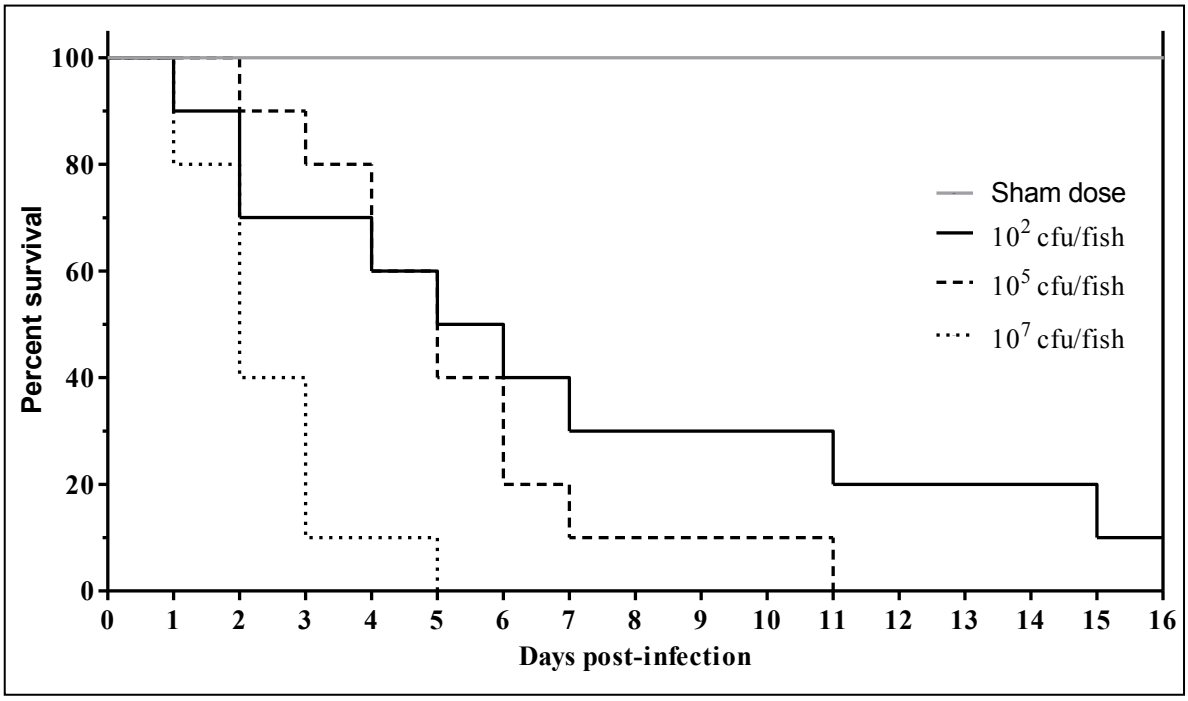


Figure 3.1 Kaplan-Meier curves comparing survival of tilapia following intra-peritoneal injection of *S. agalactiae* strain ST260 at doses corresponding to 10^2, 10^5 and 10^7 cfu/fish. No mortality was observed for the negative control fish (sham dose). Curves are significantly different as shown by the log rank (Mantel-Cox) test ($P < 0.05$).

Among fish challenged with ST23, no mortality or clinical signs were observed after 7 days p.i. and 10 fish were then sacrificed. Bacteria were recovered from the kidneys of 30% of these fish. After day 7, one dead fish was observed but this was attributed to fighting behaviour within this tank. Due to cannibalism leading to an absence of carcass, sampling for bacterial recovery was not possible for the later fish. No bacteria were recovered from the kidneys of the remaining fish sacrificed at 16 days p.i.

No mortalities or morbidity were recorded in the control animals which did not display any clinical signs nor were any bacteria recovered from these fish.

3.4.2 Challenge studies

3.4.2.1 Isolates confirmation and strain typing

All bacterial isolates recovered were confirmed as *S. agalactiae*, with one exception. One of the ST260-infected fish sacrificed at day 14 p.i. and selected for colony counts was found to be co-infected with another bacterial species. These bacteria were present in most organs investigated at a concentration of 10^4 to 10^8 cfu/g of tissue, with the exception of the eye and the heart where they were under the limit of detection. After 24h incubation on TSA, corresponding colonies were white cream and larger (5mm) than those expected for *S. agalactiae* ST260**.** These bacteria were found to be Gram-negative rods with a 16S rDNA sequence identical to numerous *Aeromonas* species (*Aeromonas hydrophila*, *Aeromonas sobria*, *Aeromonas veronii*, etc.). Biochemical tests were in agreement with the sequencing result, with the closest profile corresponding to an *Aeromonas hydrophila* (Buller, 2004).

Among the confirmed *S. agalactiae* isolates recovered and selected for PFGE, pulsotypes identical to those of the respective challenge strains were obtained (**Figure 3.2**).

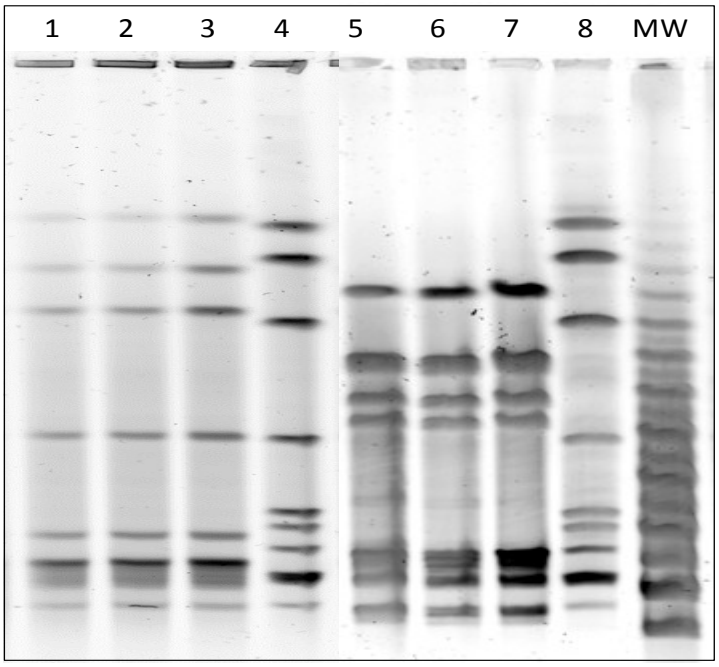


Figure 3.2 Representative PFGE patterns of *S. agalactiae* isolates from fish challenged with strain ST23 (lanes 1 and 2) or strain ST260 (lanes 5 and 6). The original strains ST23 and ST260 used for the challenge experiment were also analyzed and used as comparison (lane 3 and lane 7 respectively). *Streptococcus agalactiae* strain STIR CD-25 was run as marker (lane 4 and lane 8). MW, molecular weight marker (MidRange II PFG Marker, New England Biolabs, Herts, UK).

3.4.2.2 Bacterial counts from organs

Bacterial load of organs and number of fish for which bacterial numbers were above the limit of detection for the specified organ are represented in **Figure 3.3** and **Figure 3.4.** All calculations with regard to bacterial counts refer to all organs at a single time-point, including some organs that may have been culture negative. Detection limit for the count was of 167 cfu/ml of blood and varied between 417-1670 cfu/g for the organs depending of their weight (from 0.4g to 0.01g).

3.4.2.2.1 Fish challenged with ST260

The i.p. injection resulted in systemic dissemination into most internal organs and 24h p.i. all fish sampled had a bacterial load above the limit of detection in the spleen, kidney, liver and blood. The total viable bacterial count in the spleen, as expressed by a general mean of 3.3×10^9 cfu per gram of tissue, was significantly higher than the mean bacterial load of other organs ($p < 0.05$) and surpassed these by 1.5 to 4 logs (**Figure 3.3A**). In some fish, streptococci were not detected in the brain or the eye (**Figure 3.4A2, B2**). From day 1 to day 3, the number of bacteria increased significantly (2 to 4 logs; $p < 0.01$) in the brain, eye and heart (**Figure 3.3A and Figure 3.4A1, B1, F1**). At day 3, all organs from all fish contained bacteria above the limit of detection (**Figure 3.4A2 to G2**), with the highest load in the brain and the heart (**Figure 3.3A**). From day 3 to day 7, the number of viable bacteria remained high, but with a general tendency towards a slight decrease in bacterial concentration. A significant decrease of viable bacteria was, however, observed in the spleen ($p < 0.01$) (**Figure 3.3A** and **Figure 3.4C1**). The highest concentration of bacteria at day 7 was found, as in day 3, in brain and heart, but this difference was statistically significant only at day 7 ($p < 0.01$). Finally, between day 7 and day 14, the bacterial load remained high for brain and heart and increased significantly ($p < 0.05$) in all other organs with the exception of blood (**Figure 3.3A** and **Figure 3.4B1 to E1**). The eye was the organ with the most variability at day 14, with results spanning from 0 to 10^{10} cfu per gram of tissue (**Figure 3.4B1, B2**).

3.4.2.2.2 Fish infected with ST23

The i.p. injection of strain ST23 resulted in systemic dissemination into most internal organs and at 24h p.i. all fish sampled had bacterial loads above the limit of detection in spleen, kidney and liver (**Figure 3.3B**). Bacterial counts in spleen and kidney, as expressed by a general mean of 0.42 and 0.11 x 10^9 of cfu per gram of tissue, were significantly higher than the mean bacterial load of all other organs ($p < 0.01$), except the liver which also contained a high bacterial concentration. From day 1 to day 3, the mean viable bacterial count significantly decreased ($p < 0.05$) in all organs of the sampled fish but not in the blood (**Figure 3.3B**). At day 3, the lowest mean count was found in the eye where only 1 fish out of 9 sampled had recoverable bacteria above the limit of detection (**Figure 3.4B2**). The highest mean bacterial counts were recorded in the heart, spleen and kidney (**Figure 3.3B**), although the number of positive livers (8 of 9) was higher than the number of positive hearts or other organs (**Figure 3.4 E2**). From day 3 to day 7, the mean bacterial counts significantly decreased in most organs except brain, heart and eye. At day 7, recovery of bacteria was lower and varied considerably between fish: out of the 9 fish sampled, 5 fish had bacteria below the limit of detection in all organs, 1 fish had bacterial load above the limit of detection in all organs except brain and eye, 1 fish had bacterial load above the limit of detection in the brain and the eye only, 1 fish had bacterial load above the limit of detection in the heart and kidney only, and 1 fish had bacterial load above the limit of detection in the eye only. Finally, at day 14, viable bacteria were recovered from eye or kidney from 2 fish only (**Figure 3.4B2, D2**).

3.4.2.2.3 Comparison between strains

At day 1 p.i., bacterial concentrations were significantly different between the blood samples of fish infected with ST260 and ST23 ($p < 0.01$), whereas no statistical differences were observed when comparing bacterial concentrations within other organs between challenge strains. At day 3 p.i., bacterial concentrations were significantly different between challenge strains for all organs ($p < 0.05$) except for the kidney (**Figure 3.4A2 to G2**).

Finally, at day 7 and day 14 p.i., most organs of fish challenged with ST23 were cleared of bacteria and therefore had significantly lower means of bacterial loads in all organs (**Figure 3.4A2 to G2**).

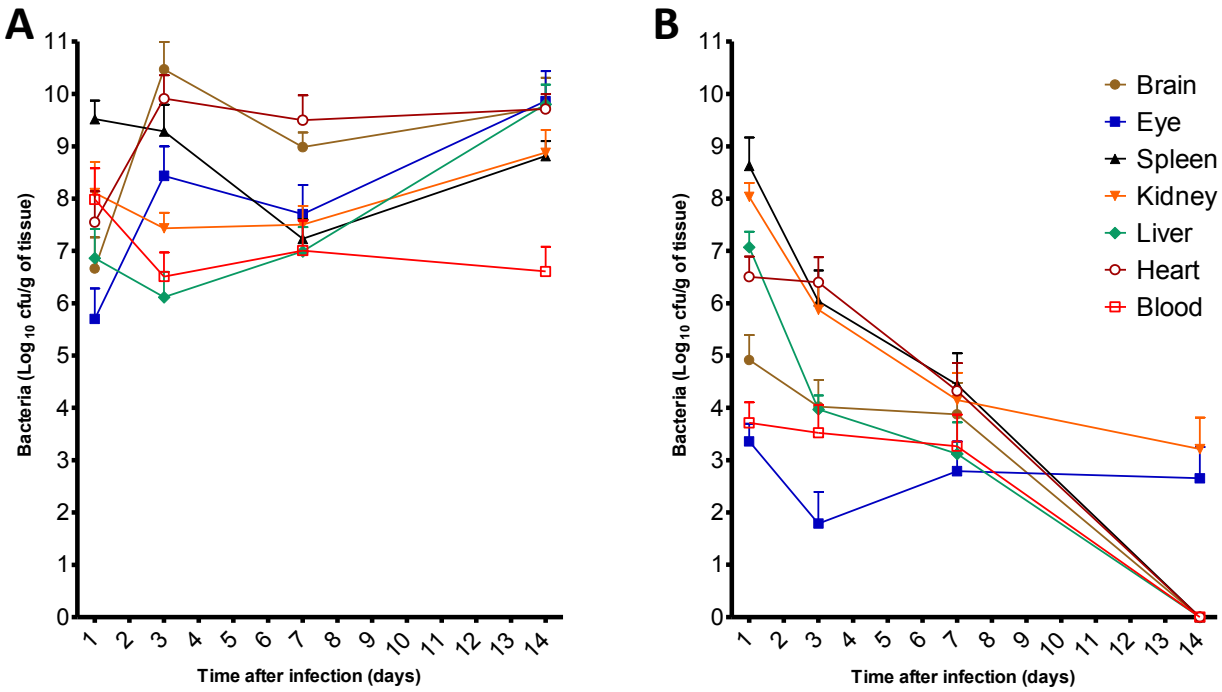
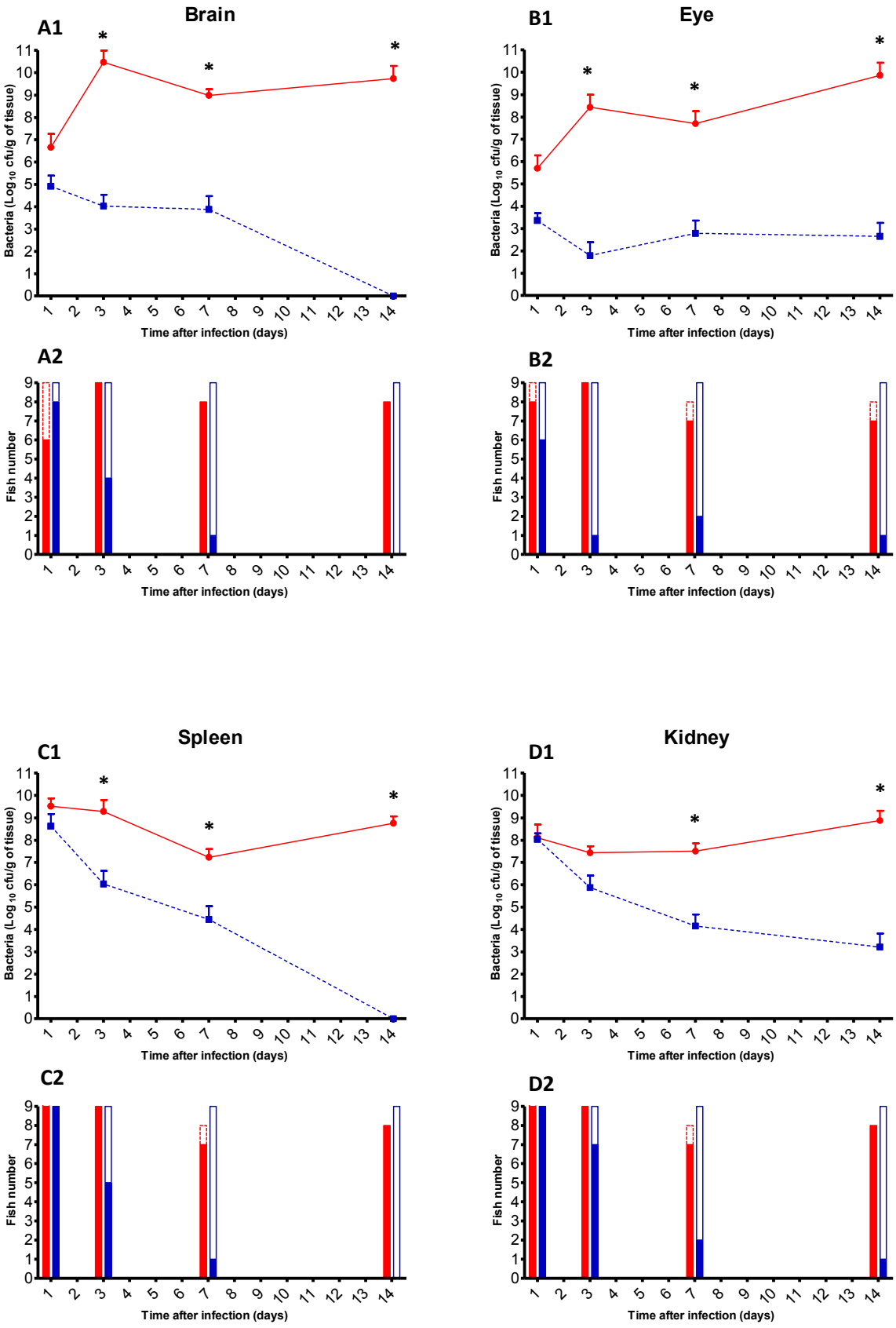


Figure 3.3 Bacterial load of *Streptococcus agalactiae* strains ST260 (A) and ST23 (B) in selected organs of Nile tilapia (*Oreochromis niloticus*) intra-peritoneally infected. Mean counts (Log_{10}) and standard deviation from fish sampled at days 1, 3, 7 and 14 post-infection are presented in the form of cfu per g of tissue or cfu per ml of blood**.**


Brain
A1
Bacteria (Log10 cfu/g of tissue)
Time after infection (days)
A2
Fish number
Time after infection (days)
Eye
B1
Bacteria (Log10 cfu/g of tissue)
Time after infection (days)
B2
Fish number
Time after infection (days)
Spleen
C1
Bacteria (Log10 cfu/g of tissue)
Time after infection (days)
C2
Fish number
Time after infection (days)
Kidney
D1
Bacteria (Log10 cfu/g of tissue)
Time after infection (days)
D2
Fish number
Time after infection (days)

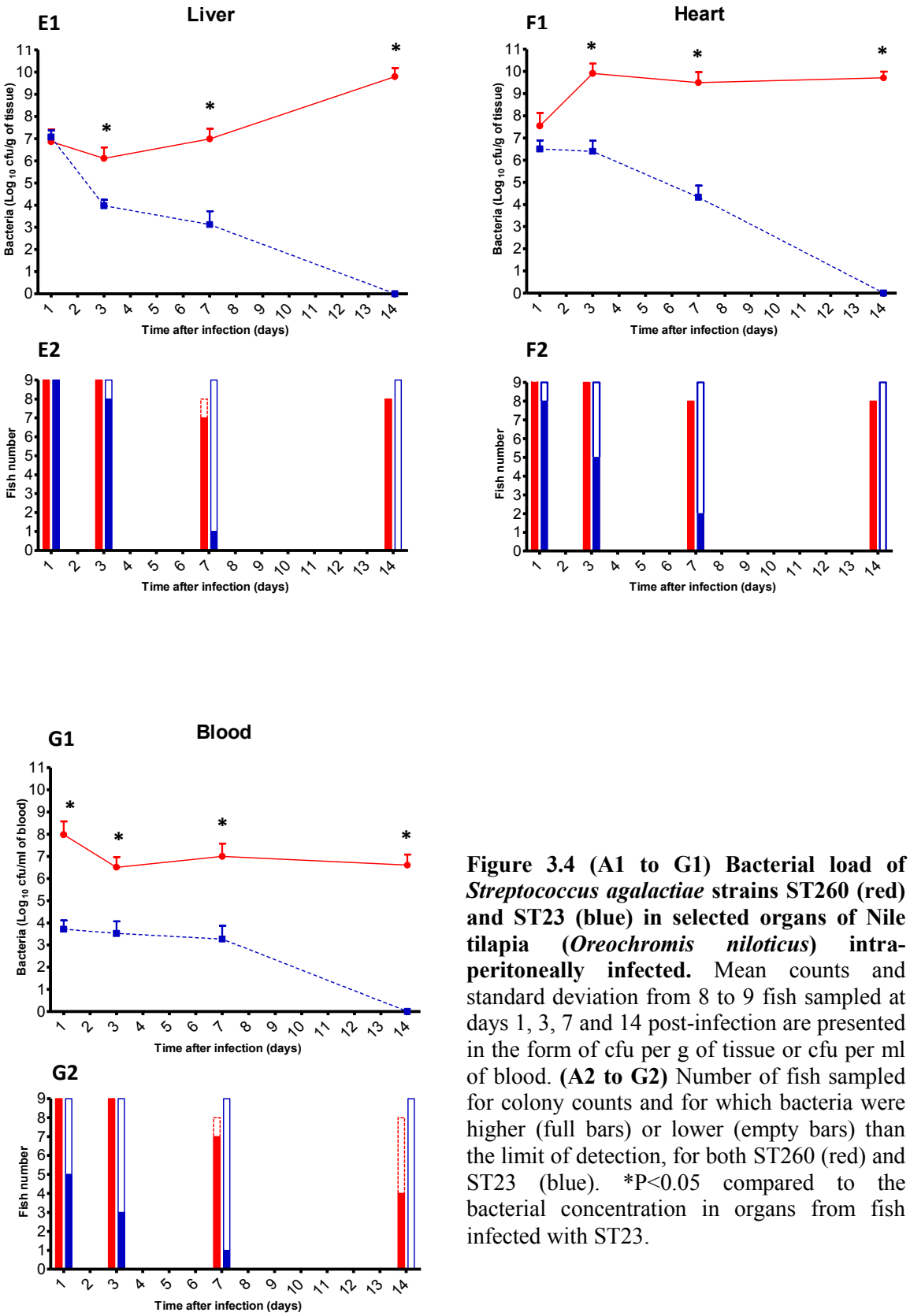


Figure 3.4 (A1 to G1) Bacterial load of *Streptococcus agalactiae* strains ST260 (red) and ST23 (blue) in selected organs of Nile tilapia (*Oreochromis niloticus*) intra-peritoneally infected. Mean counts and standard deviation from 8 to 9 fish sampled at days 1, 3, 7 and 14 post-infection are presented in the form of cfu per g of tissue or cfu per ml of blood. **(A2 to G2)** Number of fish sampled for colony counts and for which bacteria were higher (full bars) or lower (empty bars) than the limit of detection, for both ST260 (red) and ST23 (blue). *P<0.05 compared to the bacterial concentration in organs from fish infected with ST23.

3.4.2.3 Gross pathological changes

The number of fish with gross lesions in the selected organs and for each time point is presented in **Table 3.1**.

Table 3.1 Number of fish with apparent gross pathological changes in selected organs observed at each sampling point.

Time post-inoculation	Fish no.[a]	Number of fish with gross lesions					
		Brain	Eye	Spleen	Kidney	Liver	Heart
ST260							
Day 1	30	0/30	0/30	0/30	0/30	0/30	0/30
Day 3	30	15/30	2/30	2/30	0/30	2/30	10/30
Day 7	18	14/18	5/18	2/18	0/18	0/18	10/18
Day 14	11	9/11	2/11	1/11	0/11	2/11	10/11
ST23							
Day 1	30	0/30	0/30	0/30	0/30	0/30	0/30
Day 3	30	0/30	0/30	0/30	0/30	0/30	0/30
Day 7	30	0/30	0/30	0/30	0/30	0/30	0/30
Day 14	30	0/30	0/30	0/30	0/30	0/30	0/30

[a] The number of fish listed refers to live fish that could be sacrificed for examination.

3.4.2.3.1 Fish infected with ST260

No gross lesion was observed in any fish at day 1 p.i. From day 3 onwards, organs most frequently affected were brain and heart (**Table 3.1**). Affected brains were oedematous and occasionally haemorrhagic, with numerous cases of marked brain softening and liquefaction among the moribund fish and in the fish sampled at day 7 and day 14 p.i. (**Figure 3.5a**). Heart lesions consisted of focal zones of epicardial opacity at day 3, whereas most affected hearts at day 7 and day 14 appeared hypertrophic with the epicardium and the tunica adventitia of the bulbus arteriosus covered by a layer of fibrino-purulent exudate (**Figure 3.5b**). Two fish at day 14 also presented with severe pyo-pericadium, while another fish had milder lesions similar to those observed in fish at day 3. Eye lesions were occasional (**Table 3.1**) and consisted of mild (day 7) to severe (day 14) exophthalmia, which was associated at day 14 with corneal opacity (**Figure 3.5c**) and peri-ocular haemorrhage (**Figure 3.5d**). No lesion of the liver was observed with the exception of sub-capsular petechial haemorrhage in

2 fish at day 3 and a mottled appearance in 1 fish at day 14. One fish also presented at day 14 with a liver broken down in pieces possibly due to the extreme ascites and abdominal distension. No gross changes of the spleen were observed with the exception of occasional splenomegaly.

Most fish that died or were sacrificed after day 7 had a mild to severe peritonitis (**Figure 3.5d**), which was associated with ascitic fluid ranging from translucid to purulent (**Figure 3.5e**). In the case of 2 fish sacrificed at day 14, the swimming capabilities of the fish were severely impaired possibly due to the abdominal distension caused by the ascites. Finally, an unusual finding in 2 fish sampled at day 14 was the presence of abscess-like lesions of 2 to 5 mm on the oral commissure, the base of the fins, or on the jaws. No gross lesions were observed in the control fish.

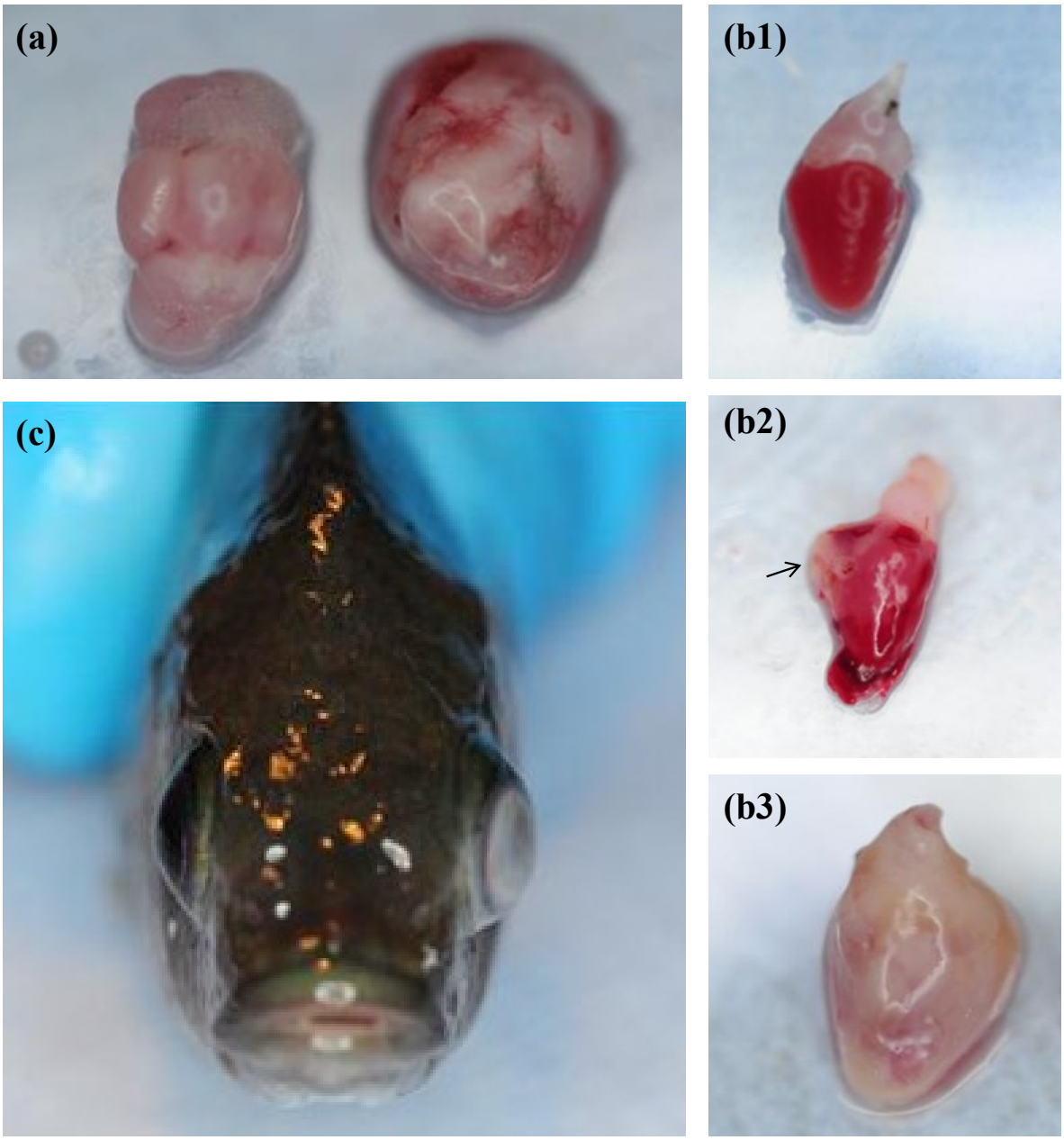


Figure 3.5 Gross external and internal pathological changes as observed in fish experimentally infected with *S. agalactiae* ST260. (a) Brain of negative control fish (left) and infected fish (right) at sampling day 7. The brain from the infected fish is highly haemorrhagic and it appears liquefied with loss of structure. **(b)** Apparently normal heart from control group (b1) and infected fish at day 3 (b2) and day 7 (b3) p.i. In the fish at day 3 p.i., a focal zone of epicardial opacity can be observed (arrow), whereas the heart is hypertrophic and covered by fibrino-purulent material in the fish at day 7. **(c)** Unilateral exophthalmia and corneal opacity as observed in a fish sampled at day 7 p.i.

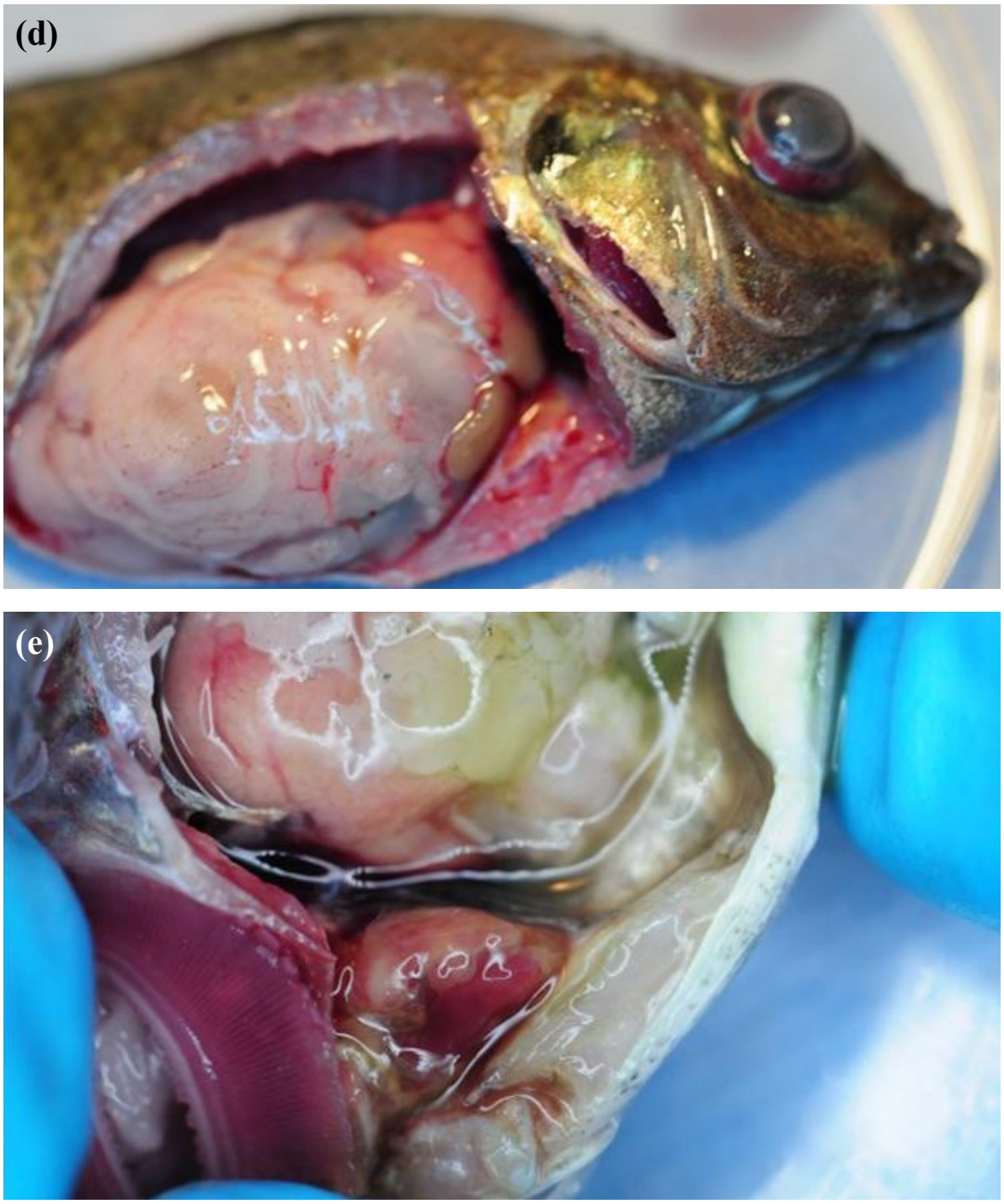


Figure 3.5 (continued). Gross external and internal pathological changes as observed in fish experimentally infected with *S. agalactiae* ST260. (d) Fish sampled at day 14 showing severe peritonitis. Note the eye that is exophthalmic and accompanied by corneal opacity and peri-ocular haemorrhages. **(e)** Fish sampled at day 14 showing the presence of fibrino-purulent material on the surface of the heart and the presence of translucent ascitic fluid within the abdominal cavity.

3.4.2.3.2 Fish challenged with ST23

No gross pathological change was observed in any of the fish and at any of the time points in the group of fish receiving strain ST23 (**Table 3.1**). The only exception was the presence of abscess-like lesions of approximately 2 to 5 mm diameter on the oral commissure, on the jaws or at the base of the fins in 2 fish sampled at day 7 p.i. and 8 fish sampled at day 14 p.i. (**Figure 3.6**).

Figure 3.6 Abscess-like lesion at the base of the pectoral fin as observed in a fish experimentally infected with *S. agalactiae* ST23.

3.4.2.4 Histopathology and immunohistochemistry

3.4.2.4.1 Fish challenged with ST260

Brain. Granulomatous meningitis ranged from minimal at day 1 to mild at day 3 (**Figure 3.7a**). At day 7 and onwards the meningitis was severe and associated with haemorrhage and vasculitis (**Figure 3.7b**). Granulomatous ventriculitis (**Figure 3.7c**) was also observed as were extensive areas of malacia extending through the cortex from the meninges (**Figure 3.7b**). Some barely encapsulated aggregates of yellow-pigmented macrophages were present scattered through the meningitis and ventriculatis. These were considered to be melano-macrophage centres (MMCs), MMC-like structures, or early granulomata (**Figure 3.7d**). At

the last sampling day, identical lesions were found with additional septic thrombi within the meninges of 4 fish out of 11 (**Figure 7e**). Two fish at day 14 however showed milder lesions similar to those observed at day 3 p.i.

With IHC, no bacteria were found at day 1 p.i. Macrophages within the meninges and meningal blood vessels were laden with bacteria from day 3 onwards. Bacteria were also observed within macrophages present in the ventricle of fish from day 7 onwards. At day 14 p.i., labelling of bacterial antigens was predominantly but not uniquely found within the MMC-like structures.

Eye. No histopathological change was observed at day 1 p.i. From day 3 onward, a marked granulomatous inflammation of the choroid of all fish was noted (**Figure 3.8a**). This observation was usually accompanied by a mild to severe inflammation of the scleral connective tissue, retrobulbar adipose tissue, optic nerve sheaths and occulo-motor muscles (**Figure 3.8b**). In some cases and more predominantly toward the end of the follow-up period, these lesions further evolved toward a pan-ophthalmitis with inflammatory reaction also involving the retina, iris, cornea and chambers, which led to a disruption of the internal structure (**Figure 3.8c**). Macrophage aggregates within the choroid rete appeared yellow pigmented. Occasionally these aggregates were loosely encapsulated at day 7 and 14 suggesting the formation of granulomata or MMC-like structures (**Figure 3.8d**).

Using IHC, no bacteria were observed at day 1 p.i. However, at day 3 bacteria were present in high number within the macrophages. At day 7 and day 14 p.i., bacteria were still present in high numbers but bacterial load was predominantly within MMC-like structures within the rete (**Figure 8e**). Numerous free bacteria were also seen within heavily infected eyes.

Spleen. The spleen exhibited major changes from day 1 and onwards. At day 1, the ellipsoids increased in size due to an accumulation of macrophages within and surrounding the Schweigger-Seidel sheaths, and there was, therefore, an apparent shift from a

predominance of red pulp in healthy fish (**Figure 3.9a**) towards a predominance of white pulp in infected animals (**Figure 3.9b**). Between day 3 and day 7, however, ellipsoidal diameter decreased to a size closer to normal, while numerous newly forming melano-macrophage aggregates were observed adjacent to the ellipsoids (**Figure 3.9c and d**). At day 14 p.i., numerous MMCs were additionally observed. These were well organized, encapsulated and surrounded by lymphoid cells (**Figure 3.9e**). Finally, blood vessels from 1 fish at day 7 and 1 fish at day 14 contained septic thrombi showing an early degree of organization and associated with areas of local tissue infarction.

Using IHC, bacterial antigens were detected from day 1 when they were apparently associated with the endothelial lining of the ellipsoidal axial vessel (**Figure 3.9b, inset**). At day 3 and day 7, bacteria were observed associated with the endothelial lining of the arteriole but also within most reticular macrophages surrounding the ellipsoids and within newly forming MMCs (**Figure 3.9d, inset**). Bacteria at day 14 were found mostly within forming and well defined MMCs. Finally, at day 7 and day 14, numerous bacteria (free and within macrophages) were observed within septic thrombi and associated infarcted tissue.

Hepato-pancreas. No major change of hepatic tissue was observed microscopically at day 1 with the exception of discrete vasculitis in most fish. Fish sampled at day 3, 7 and 14 p.i, all had mild to severe vasculitis (**Figure 3.10a**). Moreover, blood vessels from 4 fish out of the 11 still alive at day 14 contained septic thrombi showing early degree of organization, and which were associated with areas of local tissue infarction (**Figure 3.10b**). Finally, a mild infiltration of mononuclear cells within and surrounding the pancreatic tissue was present in all fish and at all sampling points (**Figure 3.10c**).

With IHC, a few *S. agalactiae*-laden phagocytic cells were present in the lumen of the vascular system at all time points. At day 14, numerous extracellular and phagocytosed bacteria were found within septic thrombi and infarcted tissues. Moreover, at all time points

and for most fish, a concentration of bacterial antigens was found to be within and surrounding pancreatic tissue (**Figure 3.10c, inset**).

Kidney. No particular lesion was found in the kidney of any fish at any of the time points. With IHC, bacteria were visible on rare occasions at day 1, day 3 and day 7 p.i. These were localized within peritubular vessels and mostly appeared endocytosed by macrophages or endothelial cells (**Figure 3.11a**). Additionally, in some cases at day 7 p.i., positive antigen labelling was also observed in MMCs. Finally, at day 14 a large number of free and endocytosed bacteria were observed scattered within peritubular portal vessels but no MMCs were found **(Figure 3.11b**).

Heart. No pathological change was observed at day 1 p.i., with the exception of minimal to moderate infiltration of inflammatory cells within the connective tissue at the base of the ventricle (**Figure 3.12a**). At day 3, lesions had progressed, and mild to severe pericarditis of the bulbus and at the base of the ventricle was observed, whereas pericarditis was minimal to mild for the rest of the ventricle. Moreover from day 3 onwards, mild to severe endocarditis of the bulbus arteriosus was observed, whereas at the level of the ventricle the endocarditis was always minimal. At day 7 and day 14, the epicarditis was severe, affecting the bulbus, atrium and ventricle (**Figure 3.12b**). Within the ventricle, fibrinoid necrosis of the compact layer of the myocardium was observed in all fish, while in some cases (5 fish at day 7; 6 fish at day 14) a layer of proteinaceous/purulent exudate was also seen external to the layer of inflammatory cells (**Figure 3.12c**). In the case of moribund fish that were euthanized between day 7 and day 14, large mats of bacteria were observed penetrating through the bulbus arteriosus (**Figure 3.12d**). Lesions at day 14 additionnally showed the presence of vasculitis of coronary arteries (**Figure 3.12e**) and numerous early granulomata within the epicardium (**Figure 3.12f**).

Using IHC, small numbers of phagocyted *S. agalactiae* were seen, but in only one of the fish sampled at day 1, localized within the connective tissue at the base of the ventricle. From day 3 onwards, large numbers of bacteria-laden macrophages and extracellular bacteria were seen within the pericardial sac. *Streptococcus agalactiae* was detected rarely from the endocardium or myocardium. At day 7 and day 14, high bacterial loads were seen occupying the inflammatory layer (mostly within macrophages) and the proteinacious layer (mostly free bacteria) of the pericardial cavity. At day 14, however, positive labeling was mostly concentrated within granulomata (**Figure 3.12g**).

Other. Abscess-like lesions of the skin were sampled for histopathological examination. These appeared to be large granulomatous lesions containing mixed inflammatory cells (macrophages predominantly) and including multinucleated giant cells. These structures were well delimited by epithelioid cells and fibroblasts and surrounded by lymphoid cells. Mononuclear inflammatory cells had also invaded the muscles tissues adjacent to the granulomata.

In IHC, positive labelling within macrophages was present focally at the periphery of the granulomata (externally and internally)

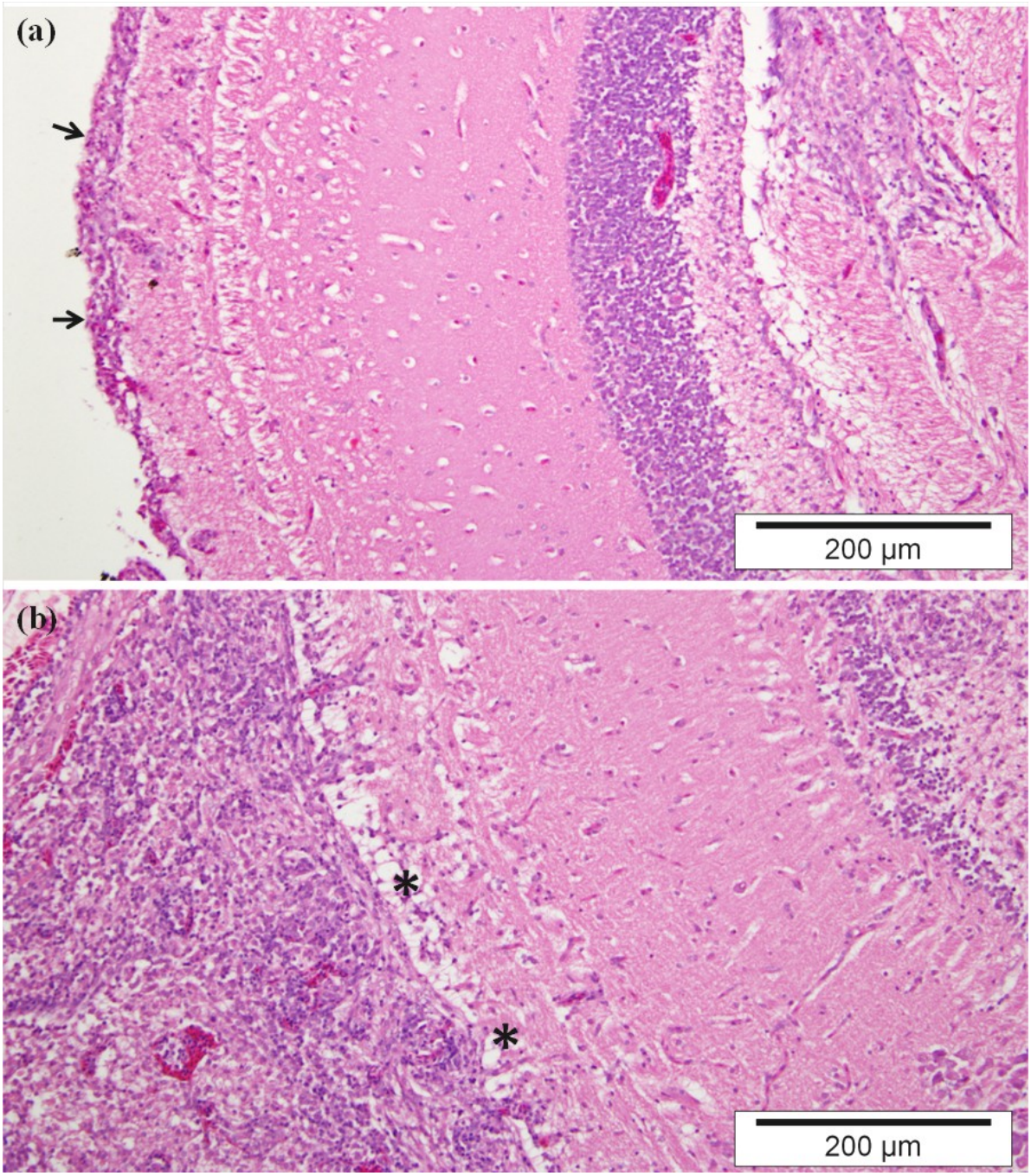


Figure 3.7 Histopathological changes in brain of fish experimentally infected with *S. agalactiae* ST260. (a) Optic tectum of a fish at day 3 p.i. with mild granulomatous meningitis (arrows). **(b)** Optic tectum of a fish at day 7 p.i. with severe meningitis. Note the loss of substance (malacia) in the stratum fibrosum marginale (*).

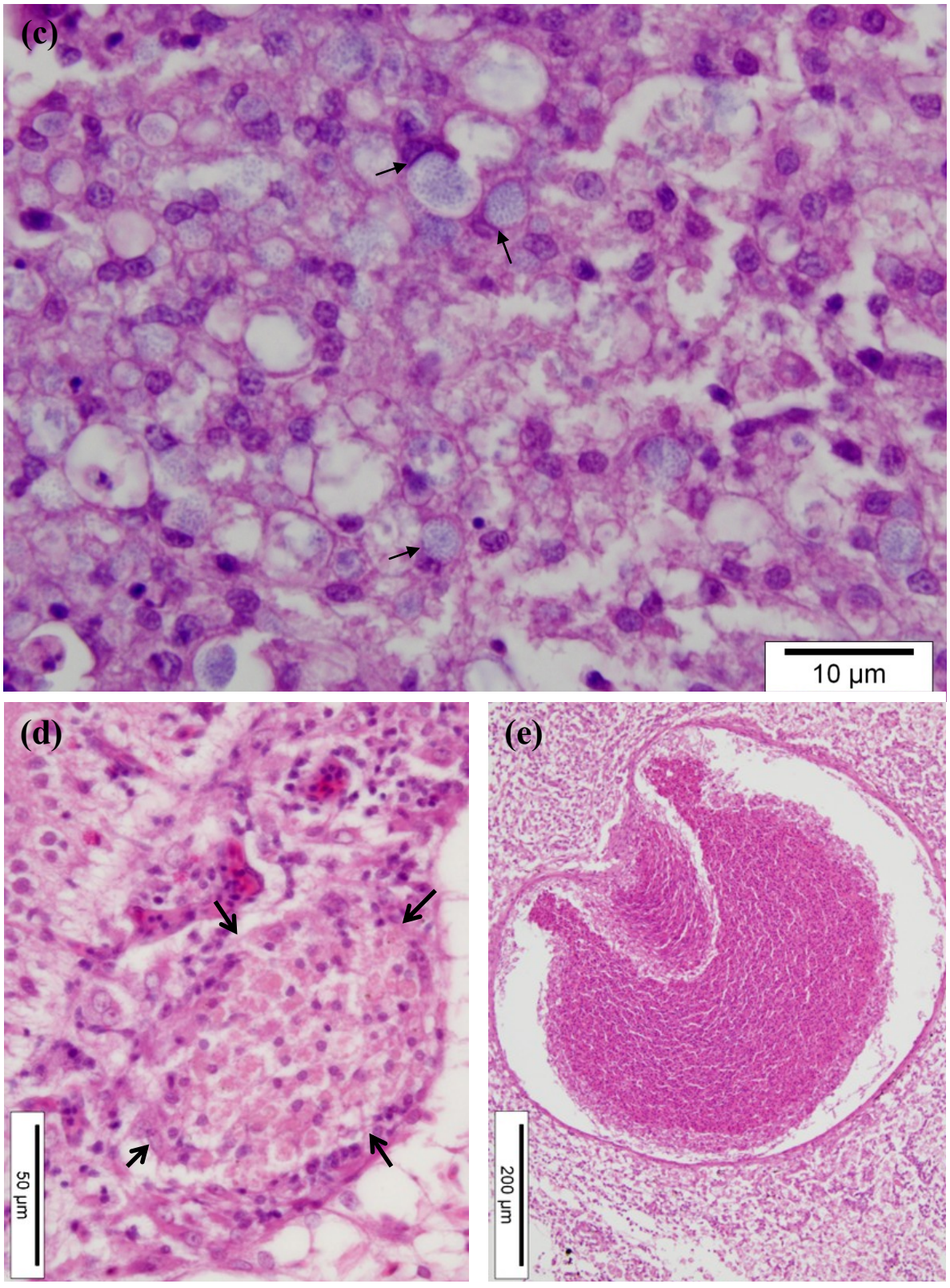


Figure 3.7 (continued). Histopathological changes in brain of fish experimentally infected with *S. agalactiae* ST260. (c) Macrophages (arrows) as observed within the ventricle of a fish at day 7 p.i. Note the presence of numerous bacteria (pale blue dots) engulfed within the macrophages **(d)** Aggregate of yellow pigmented macrophages (delimited by arrows) were observed at day 14 within the meninges of this fish, suggesting the formation of MMC-like structures. Labelling by IHC was more preponderant within these structures (data not shown). **(e)** Large thrombus as observed in the meninges of a fish sampled at day 14 p.i.

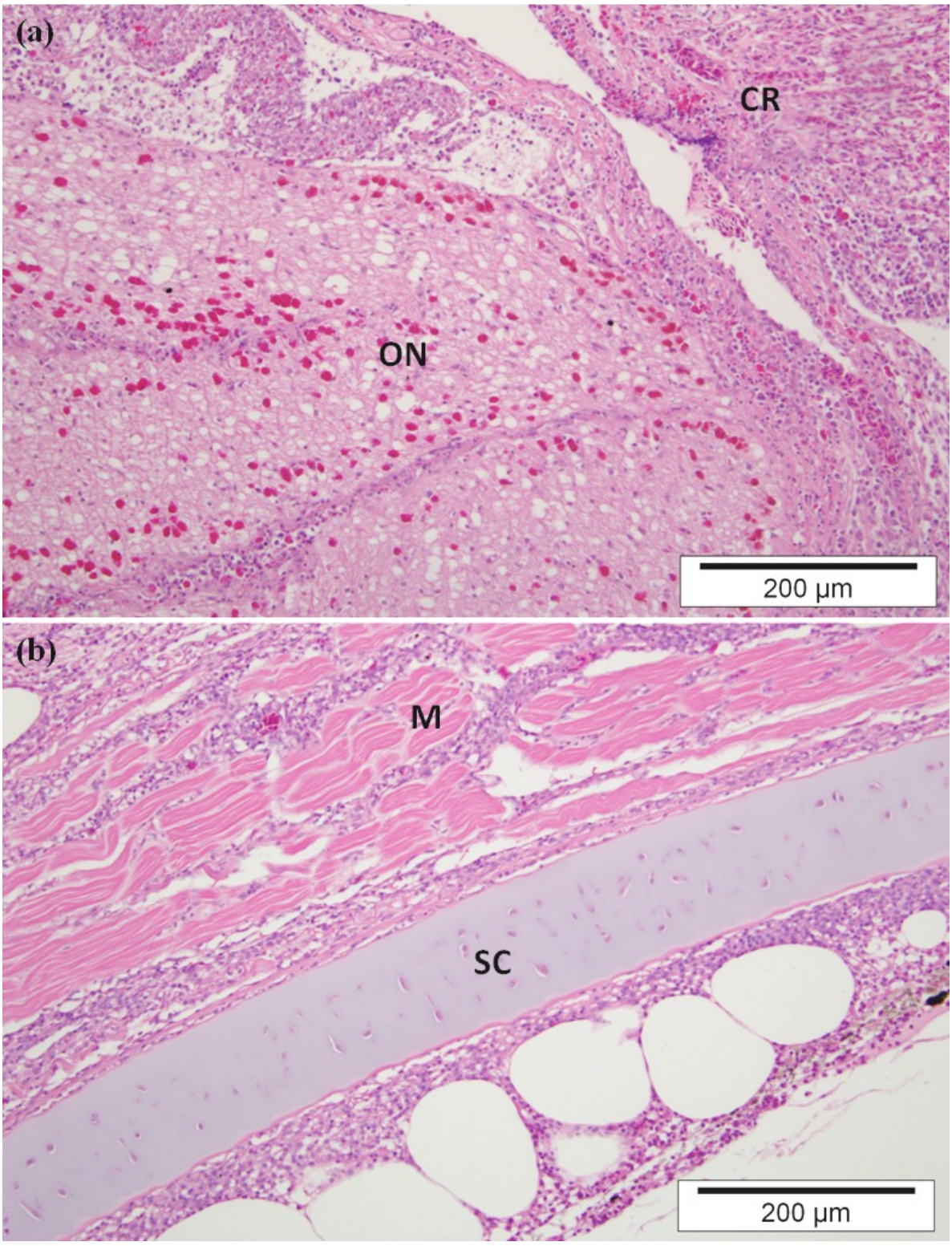


Figure 3.8 Histopathological changes in eye of fish experimentally infected with *S. agalactiae* ST260. (a) Fish at day 3 p.i. with severe granulomatous inflammation of the choroid rete and connective tissue surrounding the sheaths of the optic nerves. **(b)** Fish at day 3 p.i. with severe granulomatous inflammation surrounding the scleral cartilage and also infiltrating the external oculo-motor muscles. ON, optic nerve; SC, scleral cartilage; M, oculo-motor muscles; CR, choroid rete.

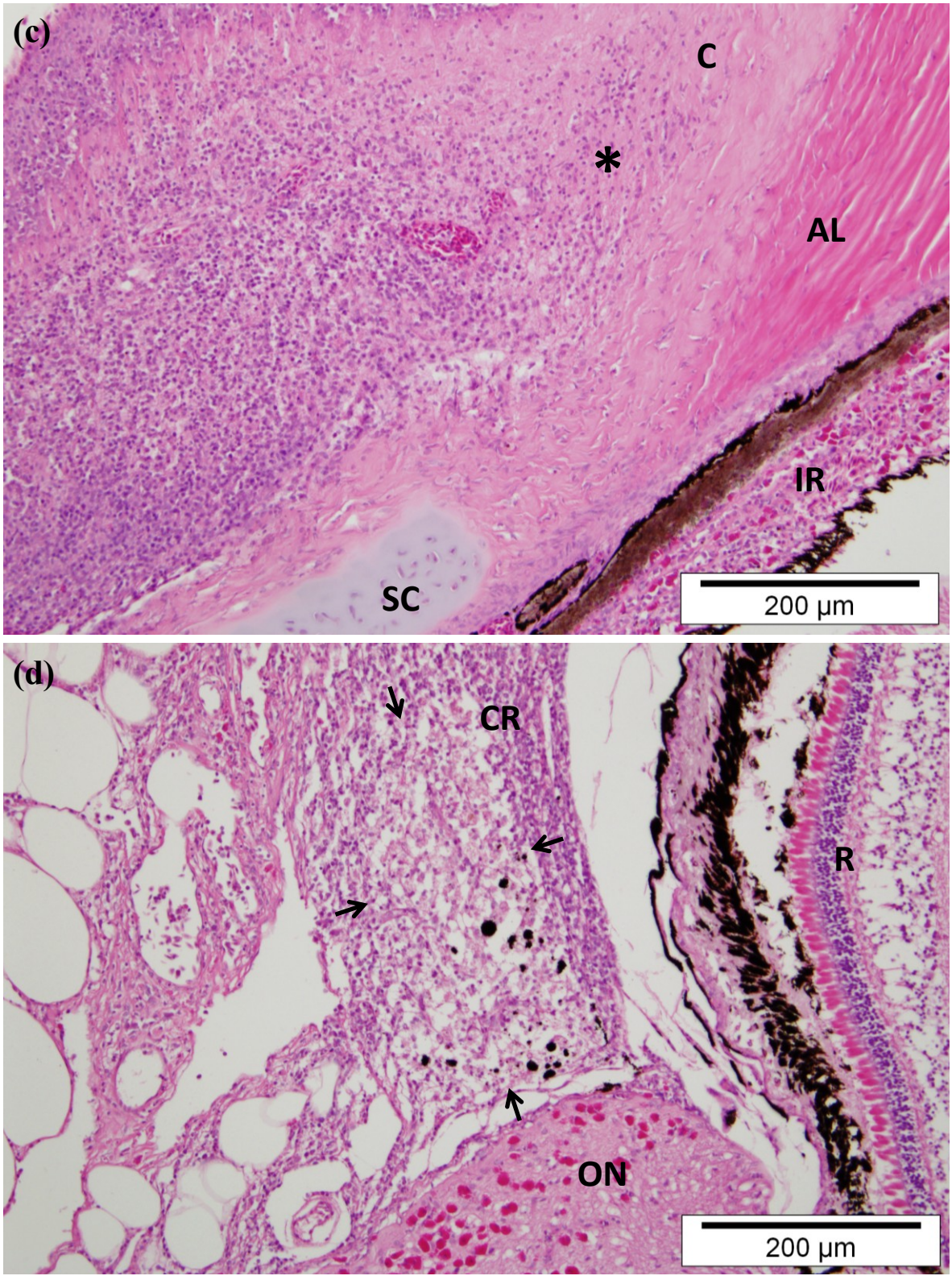


Figure 3.8 (continued). Histopathological changes in eye of fish experimentally infected with *S. agalactiae* ST260. (c) Iridocorneal angle in a fish at day 14 p.i. High infiltration of inflammatory cells is observed surrounding the scleral cartilage and appearing to progress within the cornea (*). This will lead to corneal opacity as observed clinically. Note the presence of severe inflammation involving the iris. **(d)** Fish at day 14 p.i. with severe granulomatous inflammation of the choroid rete. Note the presence of yellow pigmented macrophages in aggregates within the choroid rete (arrows). ON, optic nerve; SC, scleral cartilage; M, oculo-motor muscles; CR, choroid rete; C, cornea; IR, iris; AL, annular ligament.

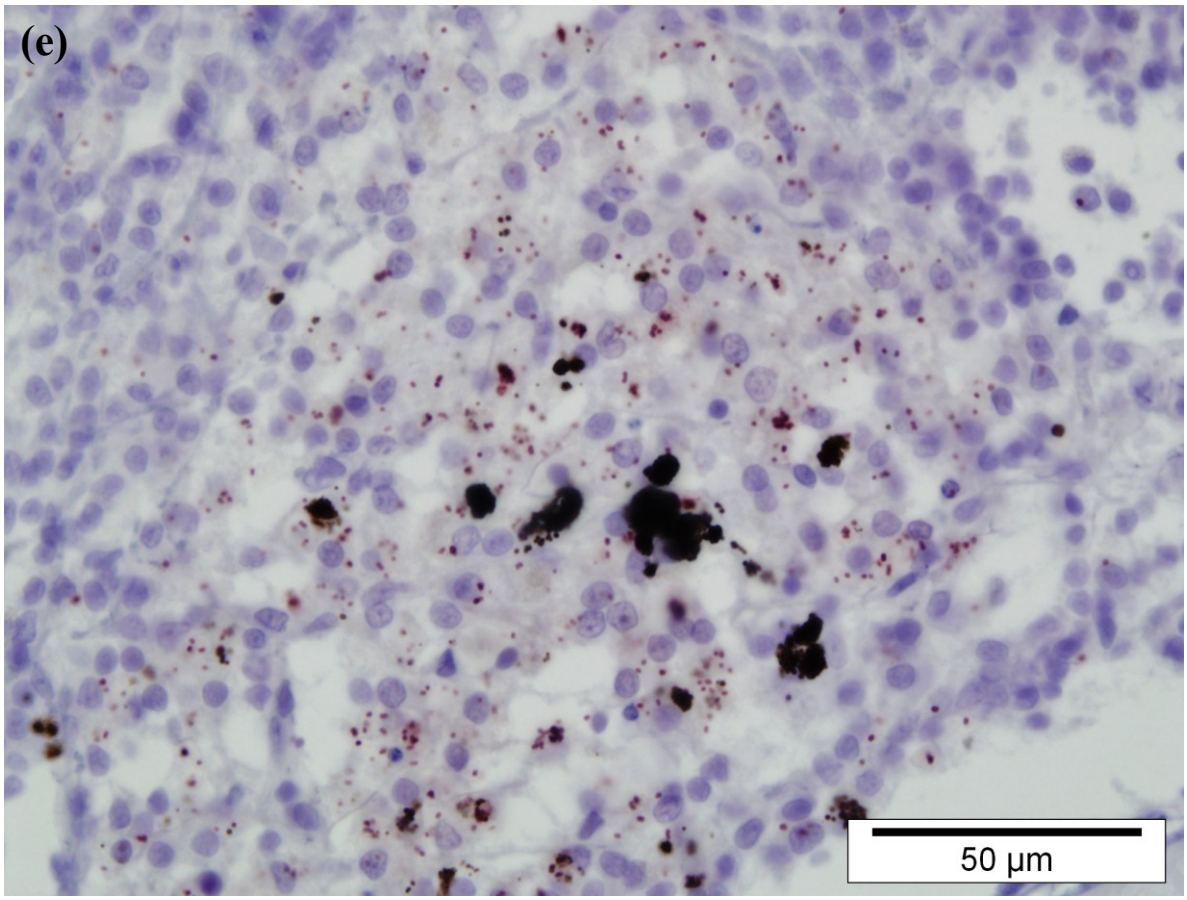
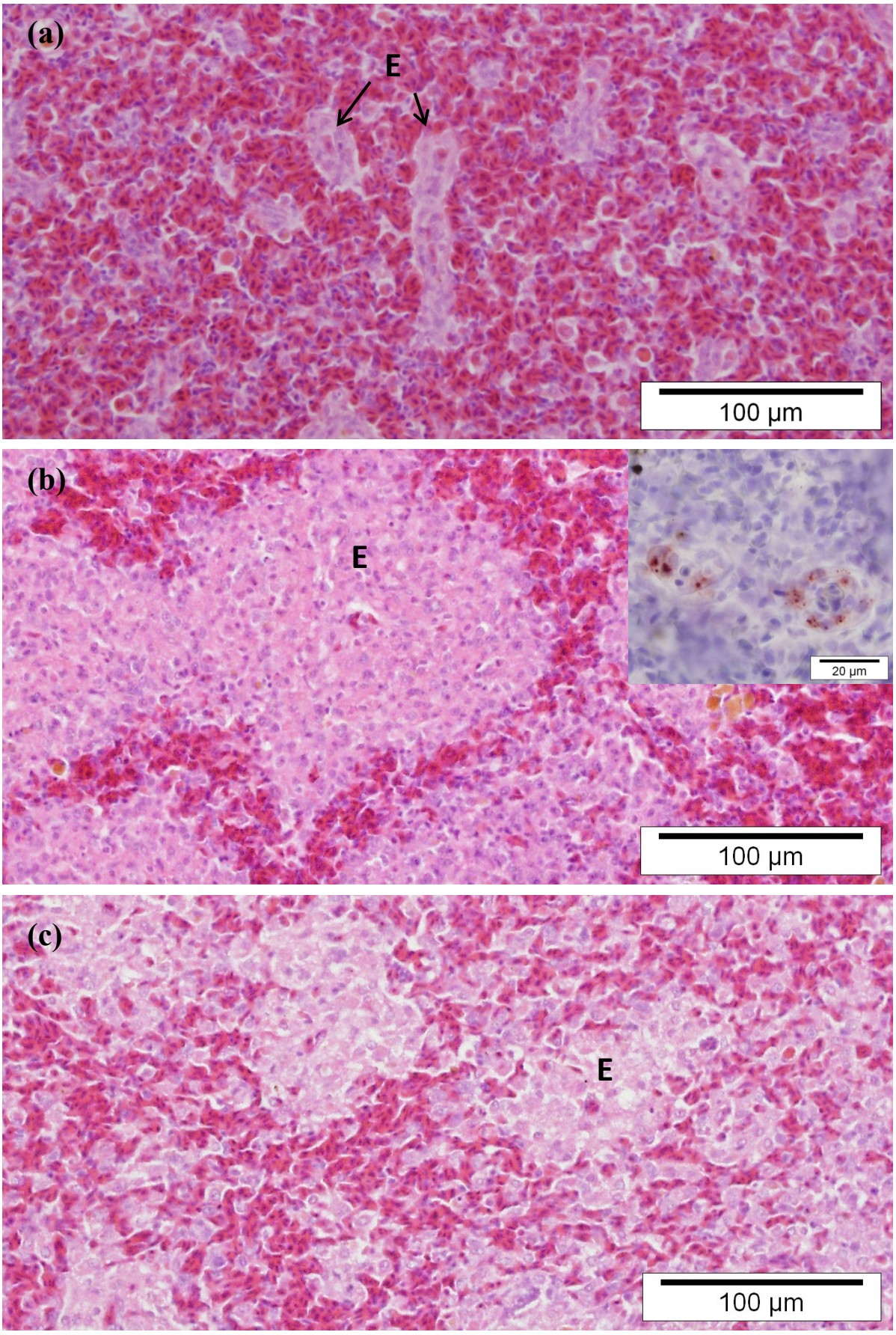


Figure 3.8 (continued). Histopathological changes in eye of fish experimentally infected with *S. agalactiae* ST260. (d) Labelling of bacterial antigens (red) by IHC as seen within the melano-macrophage aggregate presented in (d).


(a)
E
100 µm
(b)
E
20 µm
100 µm
(c)
E
100 µm

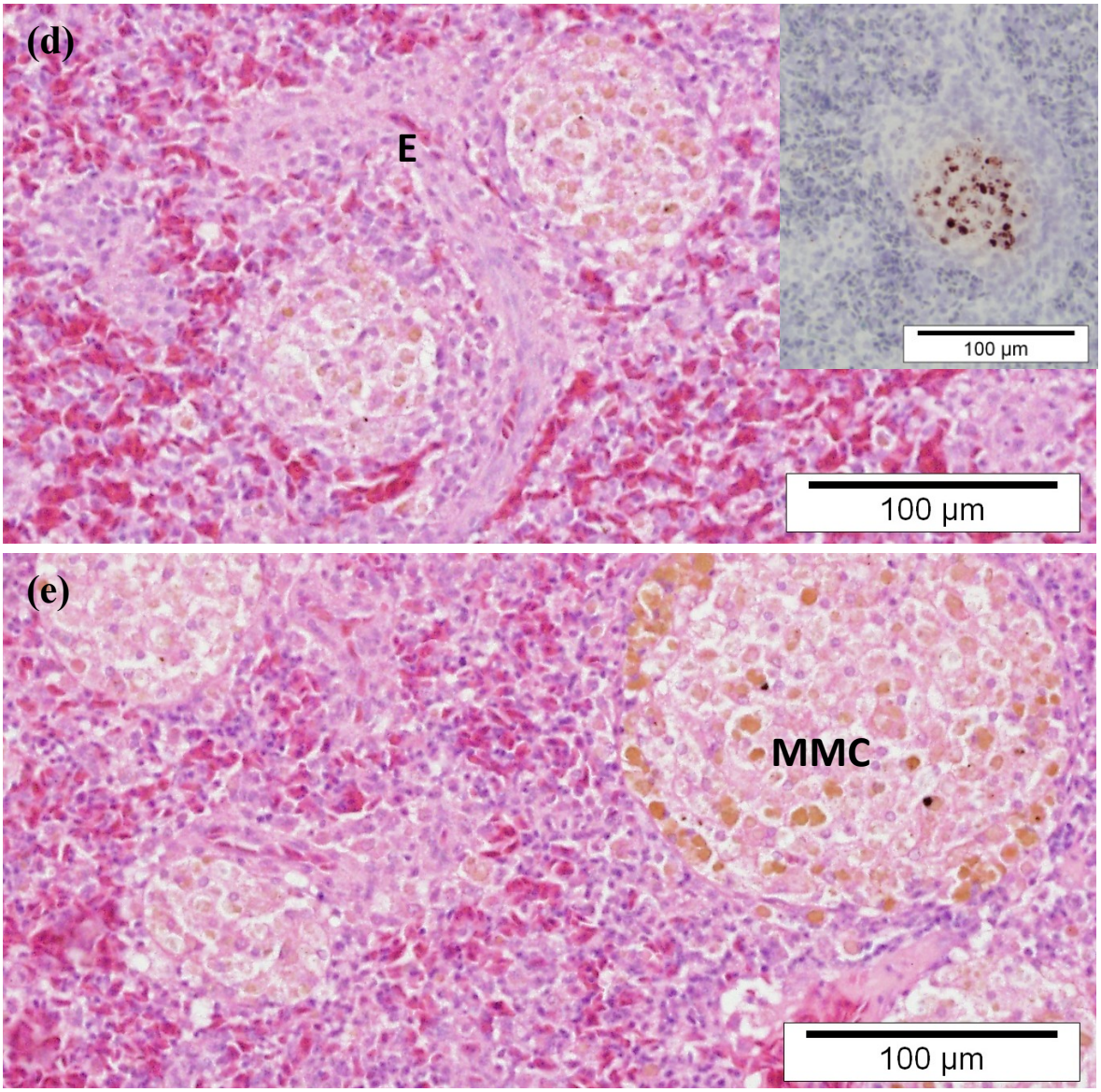


Figure 3.9 Histopathological changes in spleen of fish experimentally infected with *S. agalactiae* ST260. (a) The red pulp, consisting mainly of erythrocytes and thrombocytes within splenic cord and sinusoids, comprises the majority of the splenic parenchyma in negative control fish, whereas the white pulp is present in small clusters. The white pulp consists of lymphoid cells surrounding arterial vessels, melano-macrophage centres (MMCs) and ellipsoids. **(b)** At day 1 p.i. the ellipsoids increased in size due to an accumulation of macrophages within and surrounding the Schweigger-Seidel sheaths of the ellipsoids. Bacteria (red) can be seen trapped within the first layer of the ellipsoids by IHC (inset). **(c)** At day 3, however, ellipsoidal diameter decreased to a size closer to normal, while numerous macrophages appeared yellow pigmented. **(d)** Numerous newly forming melano-macrophage aggregates were observed adjacent to the ellipsoids at day 7 PI. These structures contain lots of bacterial antigen, as seen by IHC (inset). **(e)** At day 14 p.i, numerous large MMCs were observed. These were well organized, encapsulated and surrounded by lymphoid cells. E, ellipsoid; MMC, melano-macrophage centre**.**

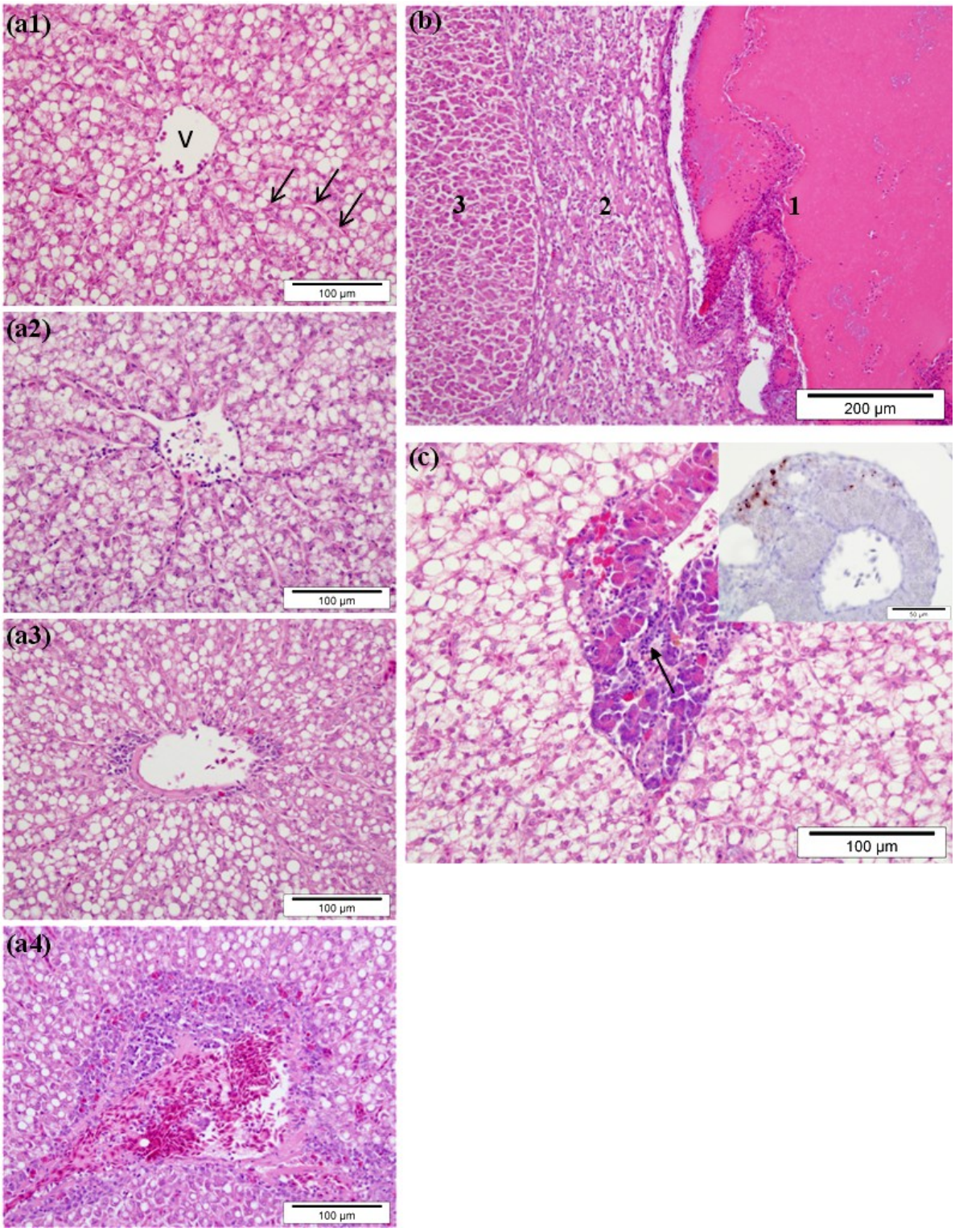


Figure 3.10 Histopathological changes in hepato-pancreas of fish experimentally infected with *S. agalactiae* ST260. **(a)** The hepatic central vein (v) is lined by a thin endothelium and erythrocytes can occasionally been seen in the lumen, whereas sinusoids (arrows) lined by endothelial cells can be observed running into the central vein. The central vein is perfectly normal in negative control fish (a1), but in most fish at day 1 p.i., discrete infiltration of mononuclear cells within the endothelium was observed, indicating the presence of discrete vasculitis (a2). The severity of the vasculitis evolved to mild (a3) and severe (a4) in the fish sampled at day 3 and onwards. **(b)** Hepatic blood vessels from fish still alive at day 14 p.i. contained septic thrombi showing early degree of organization (1) and which were associated with areas of local tissue infarction (2) adjacent to normal hepatic tissue (3). **(c)** Pancreatic tissue of all challenged fish was mildly infiltrated by mononuclear cells (arrow) and to a lesser extent by eosinophilic granular cells. Pancreatic tissue constituted a "hot spot" for the detection of bacterial antigen (inset; IHC).

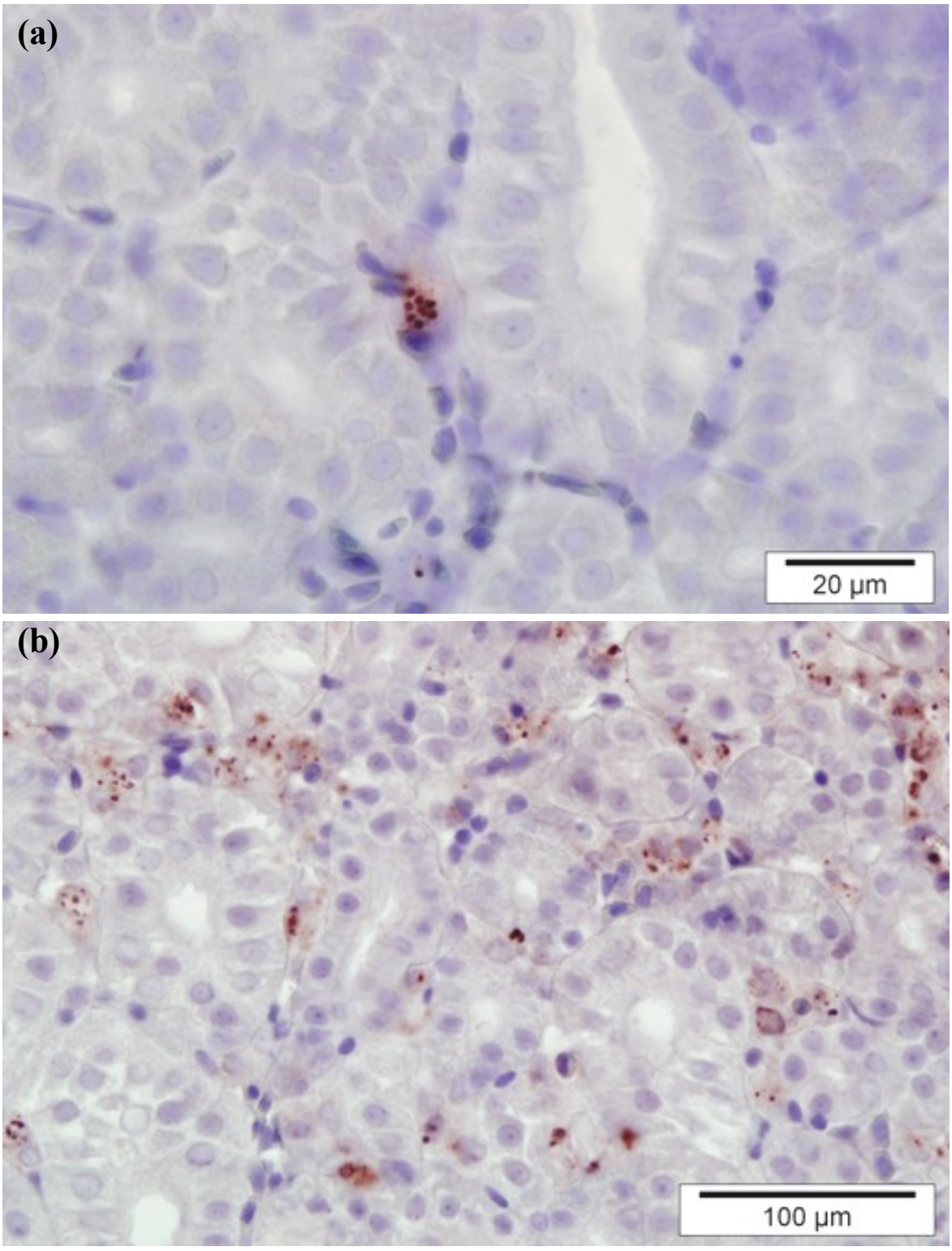


Figure 3.11 IHC of kidney of fish experimentally infected with *S. agalactiae* ST260. (a) Minimal positive labelling can be observed within the peritubular capillaries in a fish at day 1 p.i. In this case bacteria (red) are in clusters suggesting they are endocytosed within cells associated with the endothelial lining (renal portal macrophages). **(b)** Positive labelling is more pronounced in this fish at day 14 p.i.

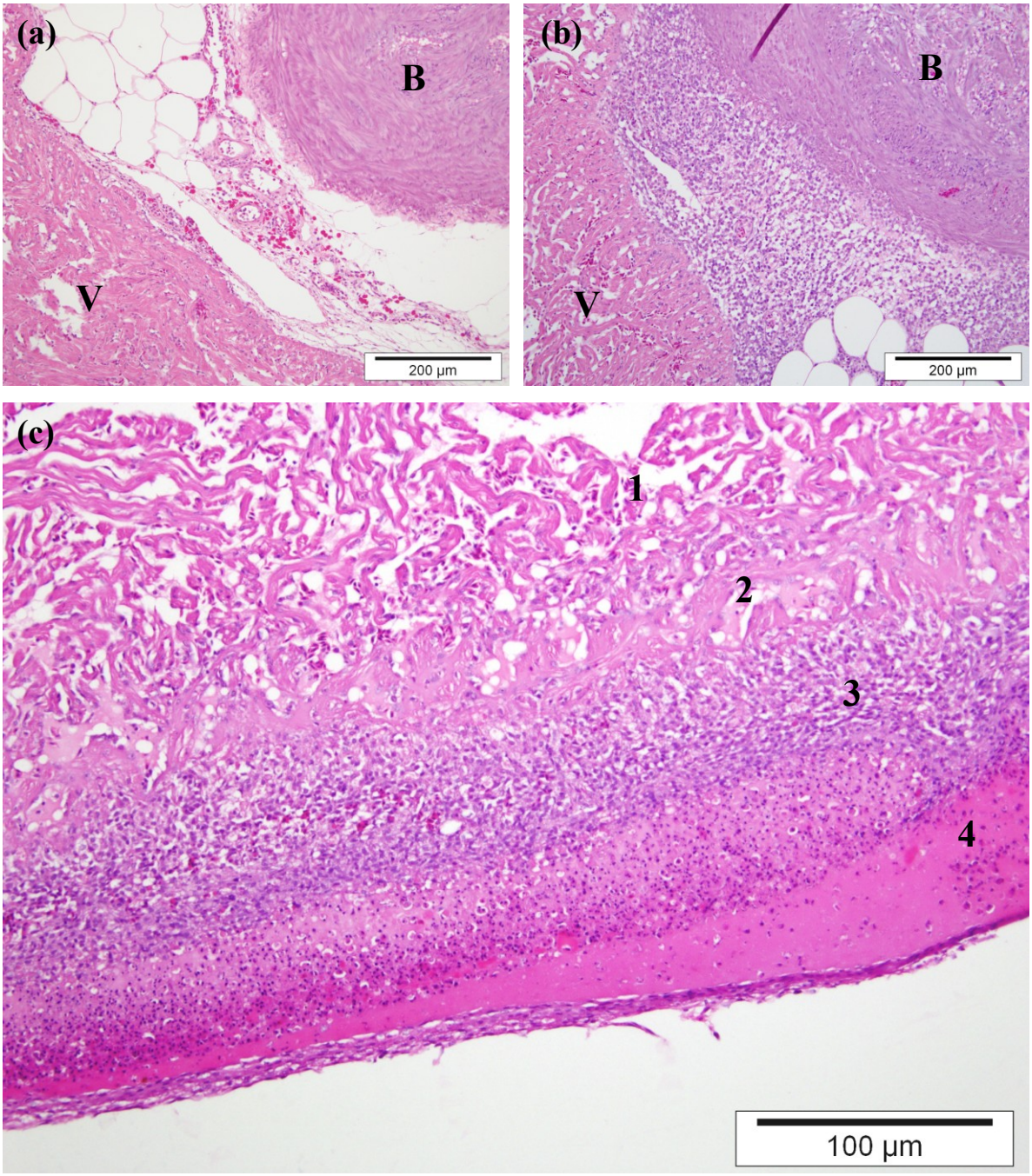


Figure 3.12 Histopathological changes in heart of fish experimentally infected with *S. agalactiae* ST260. (a) Fish at day 1 p.i. showing minimal infiltration of mononuclear cells and eosinophilic granular cells in the connective tissue at the base of the ventricle. **(b)** Severe pericarditis is seen at the base of the ventricle in a fish at day 7 p.i. **(c)** Extreme pericarditis in a fish at day 14 p.i. Adjacent to normal spongy myocardium (1), eosinophilic necrosis of the compact myocardium can be observed (2). The pericardial cavity is filled by a layer of inflammatory cells (3) followed by a layer of proteinaceous exudate (4). B, Bulbus; V, Ventricle.

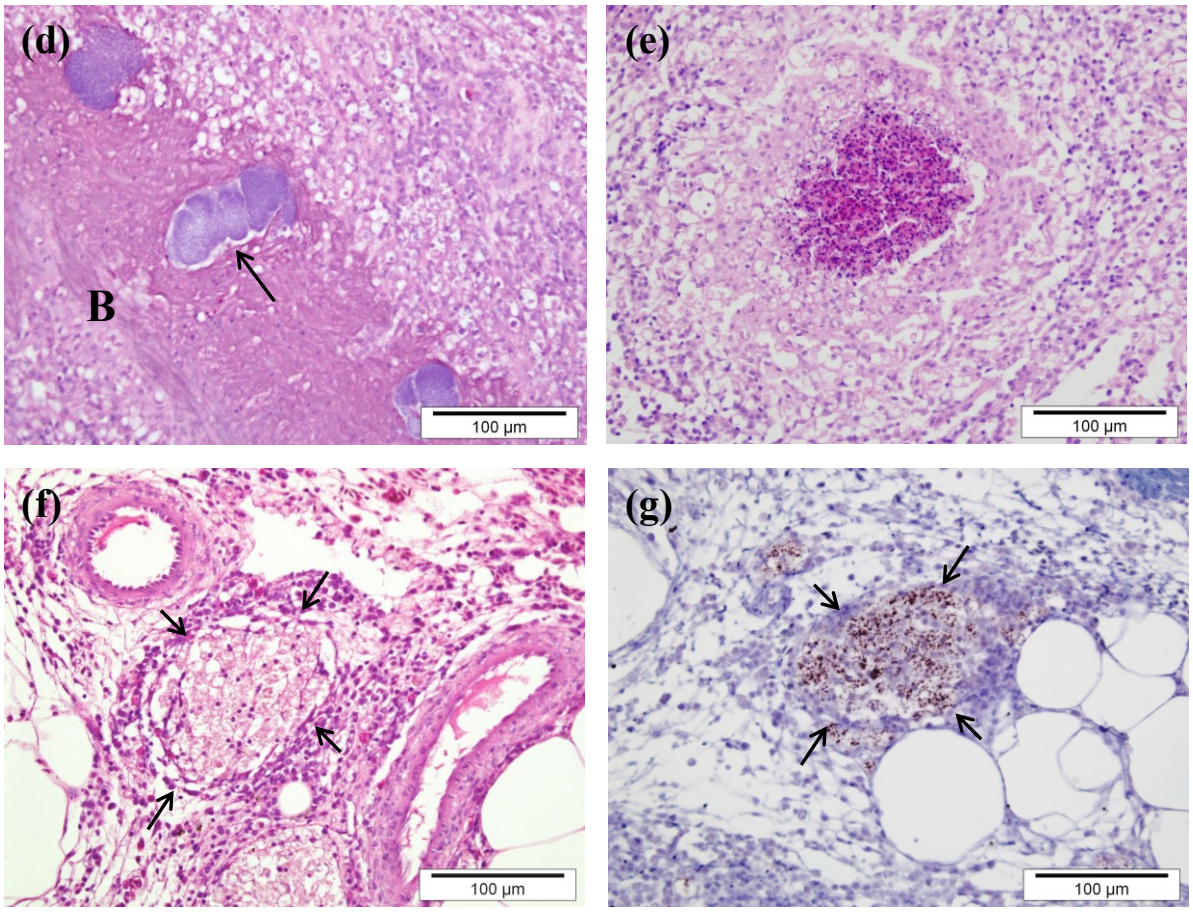


Figure 3.12 (continued). Histopathological changes in heart of fish experimentally infected with *S. agalactiae* ST260. (d) Bacterial mats (arrow) penetrating through bulbus arteriosus in a moribund fish sacrificed between day 7 and day 14 p.i. **(e)** Severe vasculitis of coronary vessels at the atrio-ventricular junction in a fish at 14 p.i. **(f)** At day 14 p.i., some aggregates of yellow pigmented macrophages (arrows delimitation) can be observed within the inflamed pericardial cavity, suggesting the formation of MMC-like structures or granulomata. **(g)** IHC showing large amounts of bacterial antigen (red; arrows delimitation) within the granulomata presented in (f).

3.4.2.4.2 Fish infected with ST23

Brain. No major changes were observed in any fish and at any of the time points, nor were any bacterial antigens seen in IHC.

Eye. A discrete to mild infiltration of mononuclear cells in the choroid rete was observed at day 1 and day 3 (**Figure 3.13a1**). These lesions persisted at day 7 and day 14 with some macrophages appearing yellow pigmented and in aggregates, suggesting the formation of MMC-like structures or granulomata (**Figure 3.13b1**).

By IHC, bacterial antigens were found only within the choroid rete. From day 1 to day 7, this was observed occasionally and confined within macrophages (**Figure 3.13a2**). At day 7 but more predominantly at day 14, positive labelling was significant but observed only within the MMC-like structures (**Figure 3.13b2**).

Spleen. No major changes of splenic tissue were observed microscopically in H&E stained sections at any time point in any fish sampled.

By IHC, a small number of bacteria were observed either in small clusters within macrophages scattered within the interstitium (**Figure 3.14a**), or to a lesser extent within macrophages surrounding the ellipsoids in fish sampled at day 1 and day 3. At day 7, only a few bacteria were observed and localised as in previous time points, but minimal labelling was also observed in the MMCs centres. Finally, at day 14, labelling of bacterial antigens was more pronounced but concentrated within MMCs only (**Figure 3.14b**).

Kidney. No major changes were observed at day 1 and day 3, but at day 7 and more predominantly at day 14, higher numbers of melano-macrophage aggregates and MMCs were seen scattered throughout the renal interstitium.

Using IHC, in all fish at day 1 and day 3, a small number of bacteria were occasionally observed within peritubular vessels, and due to their clustering, presumably within renal

portal macrophages (**Figure 3.15a**). At day 7 p.i., a small number of bacteria were still visible in most fish and as previously described, but a minimal labelling was also observed in melano-macrophage aggregates and MMCs. Finally, at day 14 p.i., labelling of bacterial antigens is more predominant and mostly concentrated within MMCs (**Figure 3.15b**).

Hepato-pancreas. No major changes were observed microscopically within the hepatic tissue at any of the time points using H&E-stained sections or IHC. The pancreatic tissue, however, had numerous changes. At day 1 and day 3, a mild infiltration of mixed inflammatory cells was observed within and surrounding the pancreatic cells in all fish sampled at these time points (**Figure 3.16a**). These lesions often evolved towards the formation of pyogranulomata-like structures at day 7, consisting of a core of macrophages and necrotic tissue surrounded by mixed inflammatory cells (**Figure 3.16b and c**). Well-organised pyogranulomata were observed at day 14 and in lesser extent at day 7, consisting of a core of necrotic material and macrophages surrounded by epithelioid cells, fibroblasts and mixed inflammatory cells (**Figure 3.16d and e**).

No bacteria were detected within the hepatic tissue using IHC. However, a "hot spot" for the presence of bacterial antigens was within and surrounding pancreatic tissue where bacteria mostly appeared within macrophages. Bacteria-laden macrophages were all concentrated within the early pyogranulomata at day 7 and day14 p.i. (**Figure 3.16c and f, insets**).

Heart. No major changes were observed in any of the fish and at any of the time points. No bacterial antigens were observed by IHC.

Other. Abscess-like lesions of the skin as previously described grossly (**Figure 3.6**) were sampled for histopathological examination. These appeared to be large granulomatous lesions (**Figure 3.17a**) containing mixed inflammatory cells (macrophages predominantly) including multinucleated giant cells (**Figure 3.17b**). These structures were well delimited by

epithelioid cells and fibroblasts and surrounded by a lymphoid cuff. Mononuclear inflammatory cells also invaded the muscle tissues adjacent to the granuloma (**Figure 3.17d**).

Positive staining within macrophages was present focally at the periphery of the granulomata through IHC (external and internal to the granuloma) (**Figure 3.17c**).

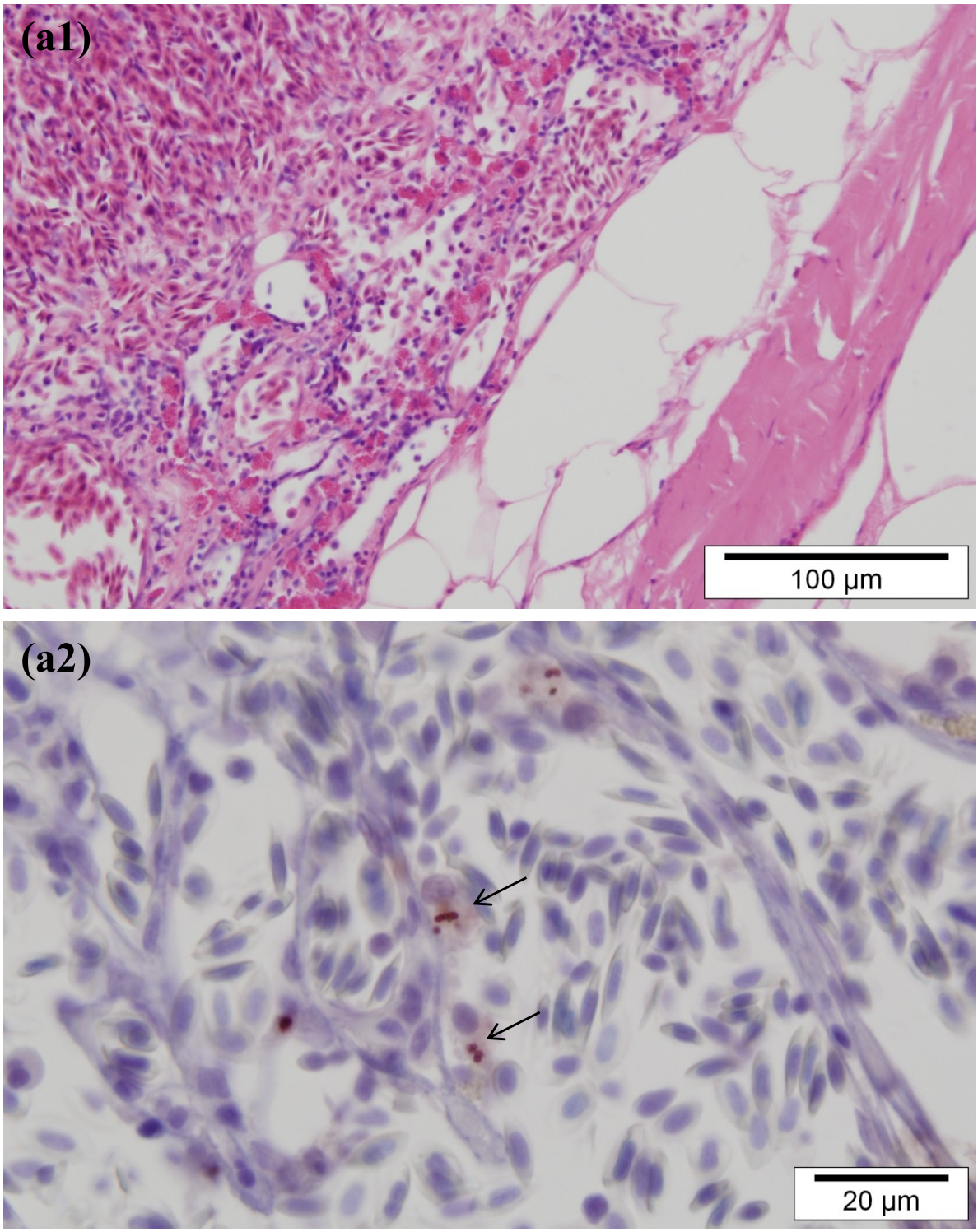


Figure 3.13 Histopathological changes in eye of fish experimentally infected with *S. agalactiae* ST23. (a) Fish at day 1 showing mild infiltration of mononuclear cells and eosinophilic granular cells within the choroid rete. The infiltrated region appears more basophilic (a1). At higher magnification, IHC demonstrates the presence of bacterial antigens (arrows) within macrophages in the choroid rete (a2).

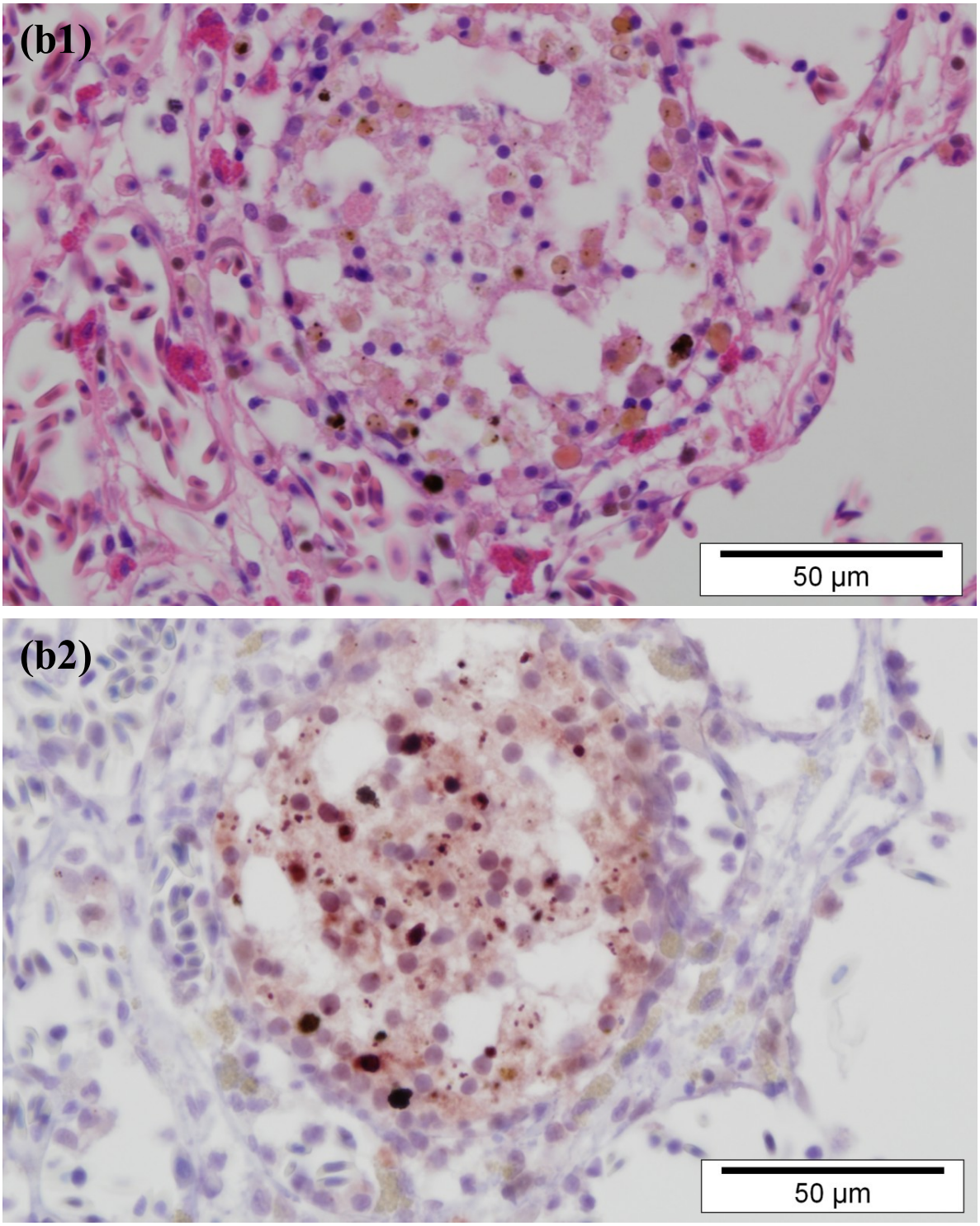


Figure 3.13 (continued). Histopathological changes in eye of fish experimentally infected with *S. agalactiae* ST23. (b) Presence of MMC-like structure in the choroid rete of a fish sampled at day 7. This structure is loosely encapsulated by epithelioid cells and fibroblasts (b1) and the IHC reveals the presence of numerous bacterial antigens within these structures.

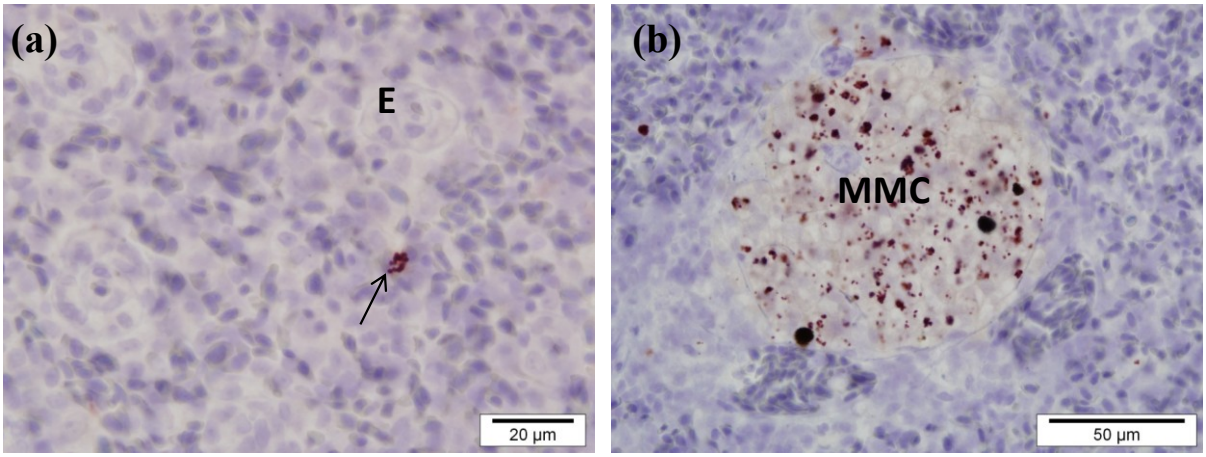


Figure 3.14 IHC of spleen of fish experimentally infected with *S. agalactiae* ST23. (a) Positive labelling of bacterial antigen is an occasional finding in fish at day 1; bacteria are located within a macrophage located in the splenic pulp. **(b)** Positive labelling is more obvious at day 14 and by now is concentrated within MMCs. E, ellipsoid; MMC, melano-macrophage centre.

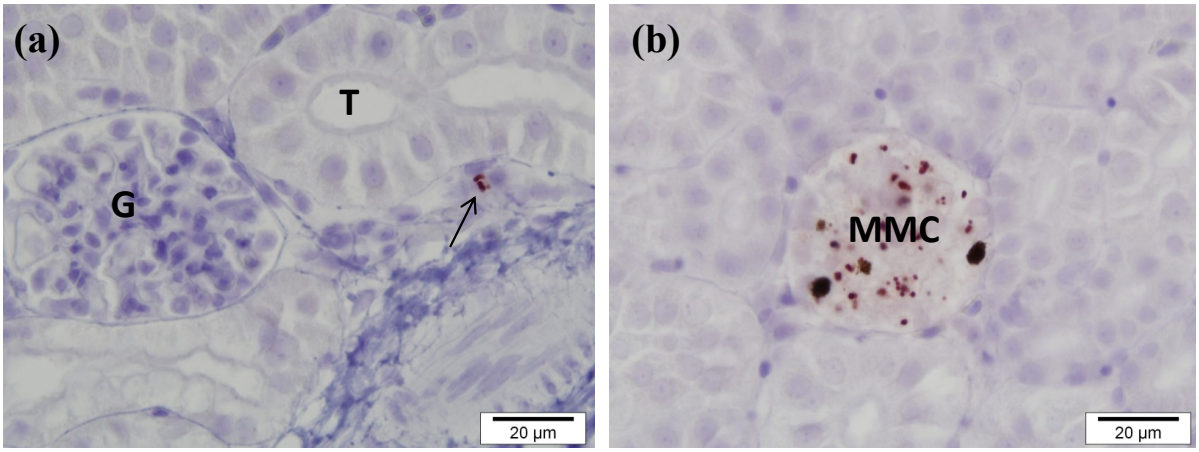


Figure 3.15 IHC of kidney of fish experimentally infected with *S. agalactiae* ST23. (a) Presence of bacterial antigens (arrow) within the peritubular vessel of a fish sampled at day 1. **(b)** Marked labelling of bacterial antigens concentrated within a MMC in a fish sampled at day 14. G., glomerulus; T, tubule; MMC, melano-macrophage centre.

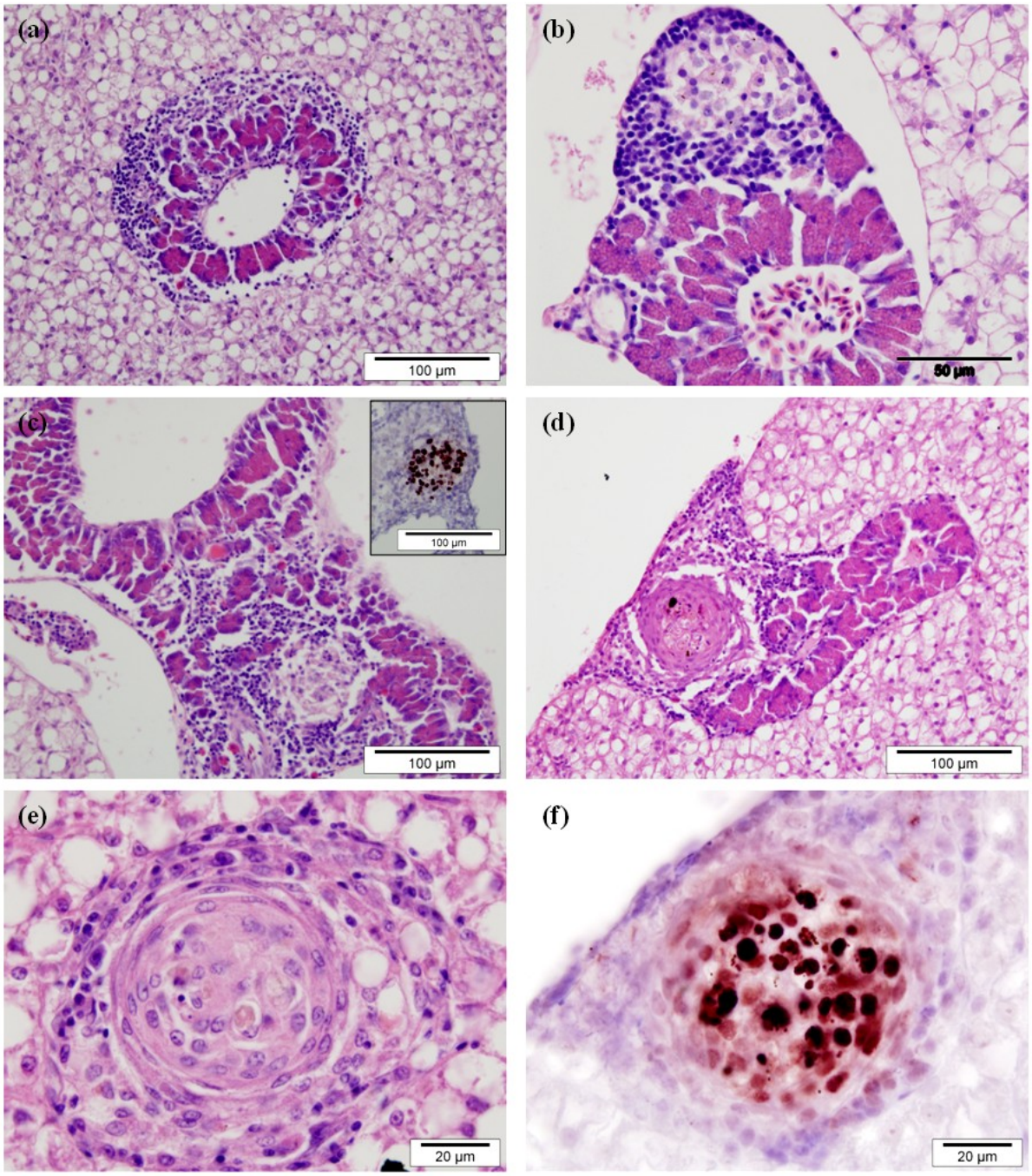


Figure 3.16 Histopathological changes in the hepato-pancreas of fish experimentally infected with *S. agalactiae* ST23. (a) Tissue surrounding the pancreas is infiltrated by mixed inflammatory cells in fish at day 3 p.i. **(b and c)** Inflammatory reaction surrounding the pancreas is more organised in fish at day 7 p.i. It consists of macrophages and necrotic tissue surrounded by mixed inflammatory cells (lymphocytes predominantly). Large number of bacteria (red) can be seen trapped within these structures (c; inset). **(d, e and f)** Well-organised pyogranulomata are found surrounding the pancreas. The core of the structure, highly positive for bacterial antigens (f), is now well encapsulated by epithelioid cells and fibroblasts.

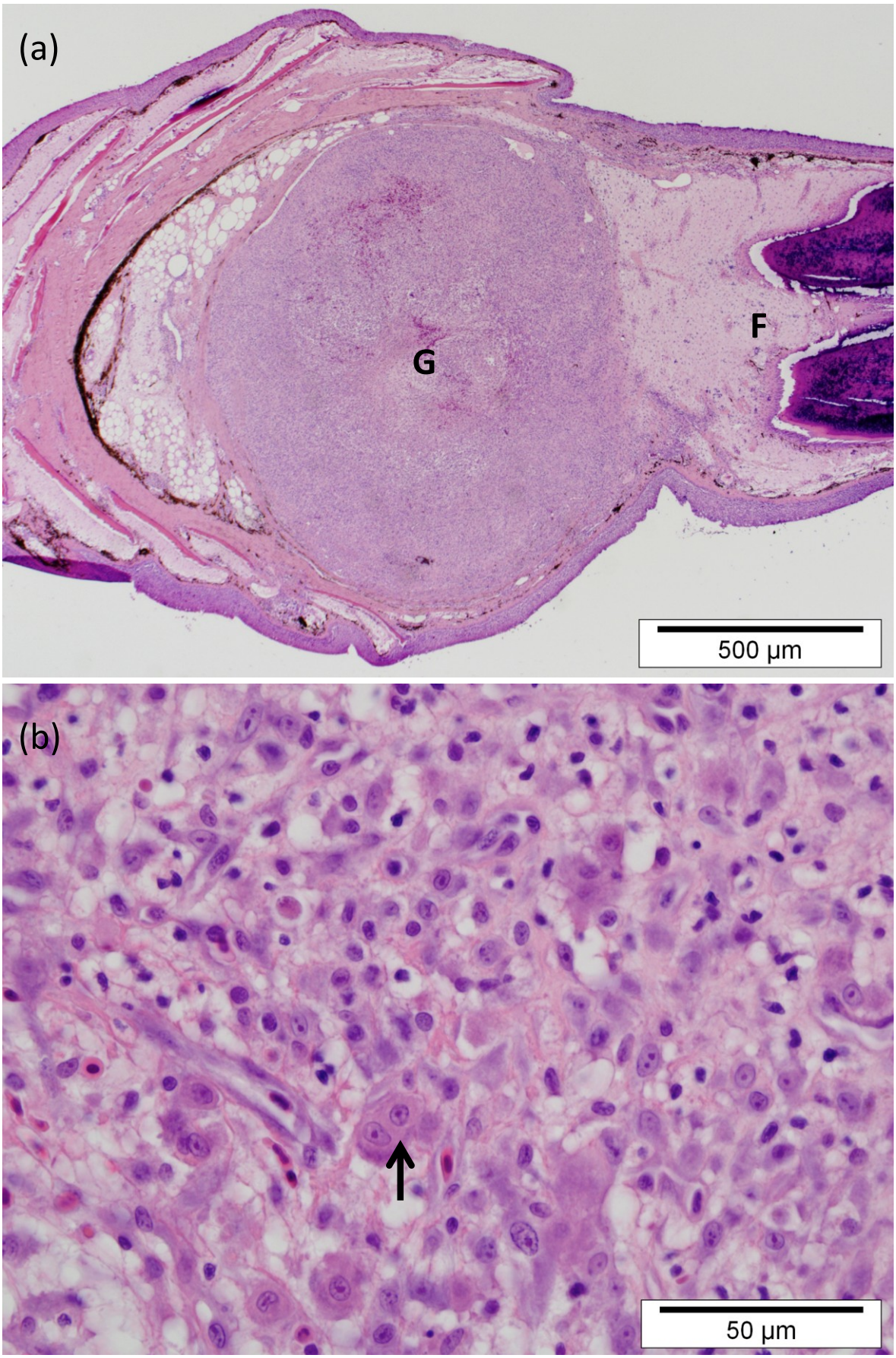


Figure 3.17 Skin lesions of fish experimentally infected with *S. agalactiae* ST23. (a) Well encapsulated granulomatous lesion as observed at the base of the fin of a fish (see Figure 6 for corresponding gross observation). **(b)** High power magnification showing mixed inflammatory cells including polynucleated giant cells (arrow) within the granuloma. G., granuloma; F, fin.

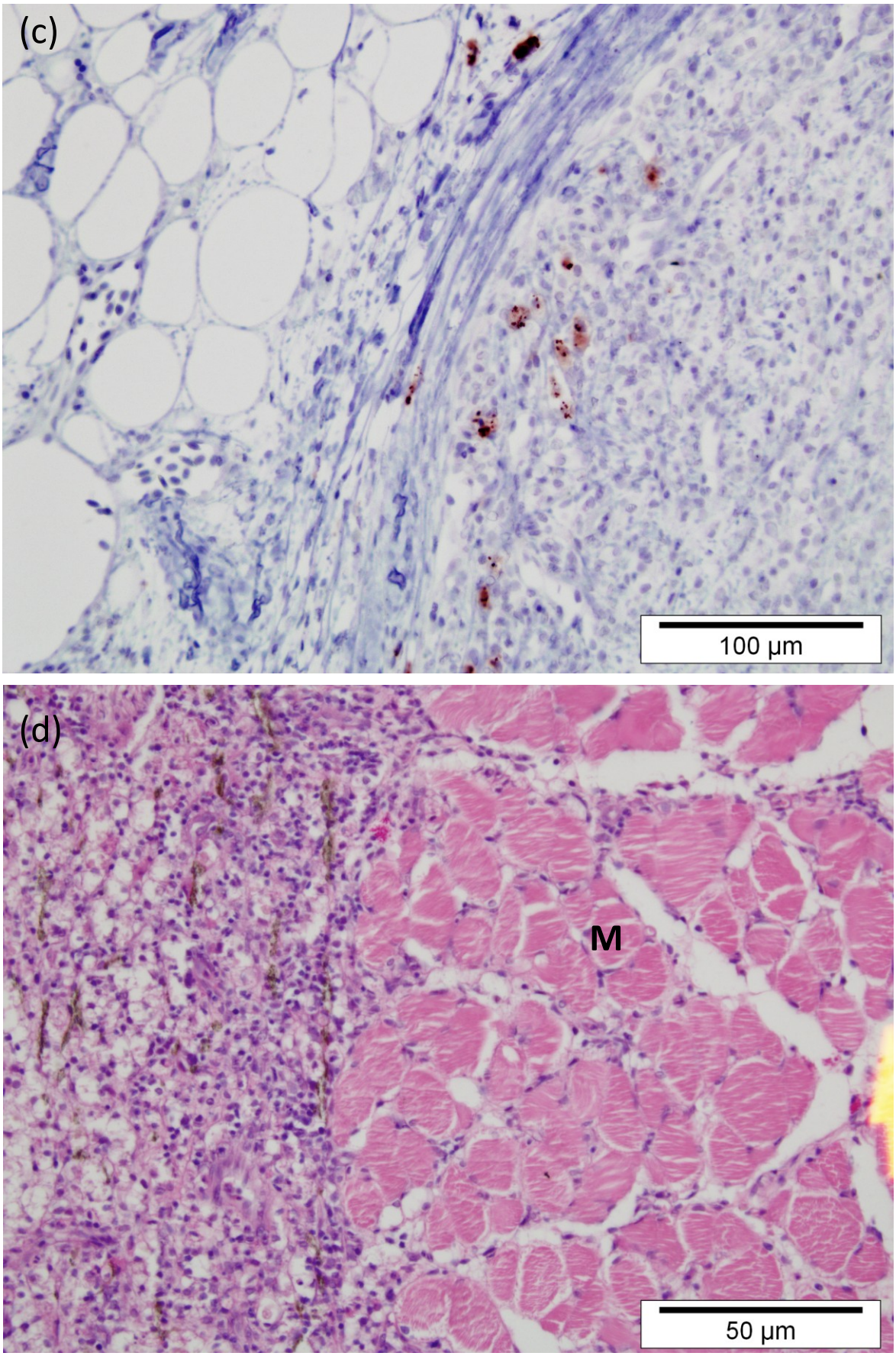


Figure 3.17 (continued). Skin lesions of fish experimentally infected with *S. agalactiae*. (c) IHC showing the presence of positive labelling (red) within and external to the fibrous wall of the granuloma. **(d)** Mixed inflammatory response surrounding a granuloma is infiltrating the adjacent muscular tissue. M., muscle.

3.5 DISCUSSION

Fish were exposed to the pathogen by i.p. route in this study, which by-passed the colonisation step and therefore would possibly not corroborate exactly with the natural progression of the disease. However, this route enabled to develop a robust challenge model and direct comparisons of motalities with other studies, which mostly focused on i.p. route, was made possible. Moreover, lesions observed with the ST260 strain were greatly similar to those reported from the natural outbreak it originated from. For the ST23 strain, however, the i.p. route most probably resulted in a septiceamic state and development of histopathological changes that might not have occurred under natural conditions.

During the passage experiments, bacteria were recovered from the brain and all other investigated organs independently of the strain used, suggesting that both ST260 and ST23 may have the ability to colonise internal organs of fish. The pre-challenges however highlighted major differences in virulence of the 2 strains. Fish infected with ST260 exibited numerous clinical signs that agreed with those observed in other fish studies (Chang et al., 1996; Evans et al., 2002), mortalities were close to 100% even at the lowest bacterial dose given (10^2 cfu/fish), and bacteria were recovered from the spleen of all euthanized fish. In comparison, fish challenged with a high dose of ST23 did not exhibit any clinical signs, no mortality was attributed to the bacterium, and the bacterial recovery from the spleen was 30% of fish after 7 days and 0% after 16 days. These observations show that the ST260 isolate, which belongs by MLST to a putatively fish-adapted subgroup of *S. agalactiae*, is extremely virulent in fish. Other experimental challenge studies performed in tilapia using the same i.p. exposure route found mortalities as high as in our study only once using isolates that were non-haemolytic, similar in that regard to ST260 (Mian et al., 2009). This suggests that piscine non-haemolytic isolates may be more virulent in fish than the β-haemolytic ones. The ST23 isolate, which belongs by MLST to a subgroup of *S. agalactiae* usually associated with humans or bovines and never reported from fish, was found to be non-virulent and cleared below detectable levels within 16 days p.i.

In this study, it was not known whether the ST23 used to challenge tilapia was pathogenic to the seal it was originally isolated from. In another study (Pereira et al., 2010), tilapias were challenged with another ST23, the reference strain NEM316, by i.p. injection of the same bacterial dose than in this study (10^7 cfu/fish). The NEM316 reference strain, known to be pathogenic in murine models of infection (*e.g.* intra-venal infection of mice with 10^6 cfu resulted in 100% mortality after 7 days; Poyart et al., 2001), was found to be non-virulent in tilapia. In agreement with this study, Pereira et al. (2010) successfully recovered the bacterial strain from the brain and kidney of fish after 48h, but no clinical signs or mortalities occurred and the fish were clear of bacteria after 15 days. Other studies evaluated the virulence in tilapia of isolates from distinct hosts of origin and/or STs. Experimental challenges of tilapia using *S. agalactiae* isolated from bovine or human hosts provided evidence that the majority of these isolates are not capable of sustainably infecting tilapia following i.p. injection (Garcia et al., 2008; Pereira et al., 2010), with the exception of some specific clones like the human ST7 (Evans et al., 2009). Even though *S. agalactiae* is a multi-host pathogen, *in vivo* studies therefore indicate niche-dependant variations in the virulence of specific bacterial subpopulations.

Data generated in this study from the challenge experiment enabled for the first time an in-depth description of the pathological changes, the bacterial loads in the organs and the bacterial antigen localisation during the time course of infection, using virulent (ST260) and non-virulent (ST23) isolates of *S. agalactiae*.

The challenge study suggests that following i.p. infection, both isolates gained access to the bloodstream. This was demonstrated by the viable colony counts recorded from the blood samples and from most organs sampled at day 1 p.i., with the spleen and kidney containing the highest concentration of bacteria. Further examination by IHC in this study showed that bacteria were passively trapped within the reticulo-endothelial system of the spleen and kidney, which is in agreement with many other studies that consider these organs as a first

line of defence for the trapping and clearance of blood-borne substances (Press and Evensen, 1999). Following this initial trapping, however, the subsequent progression of the bacterial load and associated pathological changes differed markedly between the two strains.

Concerning the fish infected with ST260, bacterial counts from organs revealed that following day 1 p.i., the bacterium had a particular tropism for the brain and the heart and, accordingly, these organs were the most affected as appeared grossly. Through the investigation of natural clinical outbreaks of tilapia, Hernandez et al. (2009) also identified the brain as the organ of predilection for *S. agalactiae*, but in their study the eye was found to be more frequently infected than the heart. In our study, histopathological changes involved mild to severe granulomatous inflammation and associated necrotic lesions in the pericardial cavity and in the meninges. The pericardial cavity in fish is frequently affected during bacterial septicaemias, which possibly results from bacterial colonisation through lymph drainage (Ferguson, 2006). The presence of bacteria within the brain, however, requires the ability to breach the blood brain barrier. In human strains of *S. agalactiae*, numerous virulence factors putatively involved in the transcellular or paracellular crossing of the blood brain barrier have been suggested (van Sorge et al., 2009), but it is unknown whether similar virulence factors are expressed by ST260. For the crossing of the blood brain barrier, bacteria may also use macrophages as a vehicle (the so-called Trojan horse mechanism) as it has been shown for *S. iniae* (Zlotkin et al., 2003). The ability of *S. agalactiae* to survive within fish macrophages and trigger their apoptosis has not been evaluated. Based on our work and other studies, fish pathogenic *S. agalactiae* always colonise the brain (Suanyuk et al., 2008; Hernandez et al., 2009; Zamri-Saad et al., 2010) but further work is required to understand the mechanism behind it. At the end of the challenge study, the pathological process evolved towards the formation of numerous MMC-like structures in the meninges, pericardium and choroid rete of the eye. Bacteria were concentrated mostly within these MMC-like structures as well as within the MMCs of

the spleen and kidney. However, the viable bacterial load remained high as shown by the colony counts, suggesting that this strain is able to survive within macrophages and/or to evade the immune system of the fish. Moreover, fish infected with ST260 displayed toward the end of the challenge numerous septic thrombi associated with severe local tissue infarction, notably in the liver, but also in the spleen and meninges. Numerous bacteria were found within these lesions, explaining the general increase of bacterial concentration between day 7 and day 14 in these organs. The formation of these thrombi may be due to the degenerative changes of endothelial cells caused by the bacterium (Zamri-Saad et al., 2010), and/or to the initiation of disseminated intravascular coagulation caused by bacterial toxins as suggested in the case of *S. iniae* infected fish (Ferguson et al., 1994). The development of these thrombi will most probably compromise widely the health status of the fish remaining at the end of the challenge period.

Concerning the fish infected with ST23, the concentration of viable bacteria decreased over time, with almost no viable bacteria recovered from fish at day 14 p.i. In histopathology, a mixed inflammatory response associated with the presence of bacteria was observed in the choroid rete and in the pancreatic tissue; lesions evolved towards the formation of MMC-like structures in the rete and pyo-granulomata in the pancreatic tissue. These structures, as well as the MMCs within the spleen and kidney, were strongly positive for bacterial antigens as shown by IHC. Bacteria were restricted to these structures and possibly were non-viable as suggested from the colony counts data. In the case of ST23, the infection therefore appeared to be well-contained, suggesting that the ST23 strain does not successfully evade the immune system of the fish nor does it survive within macrophages, explaining the absence of systemic lesions with this bacterial strain.

Numerous studies investigating the fate of antigenic and non-antigenic particulates within the bloodstream of fish have been published (Ellis, 1980; Secombes and Manning, 1980; Ferguson et al., 1976 and 1982). Some of these studies reported that following the initial

trapping and phagocytosis of particulates in the ellipsoidal sheaths of the spleen, phagocytic cells progressively migrate towards MMCs, leaving ellipsoids clear of any particulate matter. Ferguson (1976) however hypothesised that in the turbot (*Scophthalmus maximus*) spleen MMCs directly arose in and budded out from the ellipsoid sheaths. Interestingly, both processes were observed in the study presented. In the case of fish infected with ST260, a marked increase in the size of ellipsoids due to an accumulation of macrophages within and surrounding the sheaths was observed, phenomena unreported in previous studies but compatible with the high phagocytic activity necessary in the area. Bacteria trapped within ellipsoids at day 1 were phagocytosed and numerous aggregates of bacteria-laden macrophages were then observed adjacent to the ellipsoids; this progressed towards the formation of MMCs. Our observations suggested that numerous MMCs arose from the ellipsoids in the course of infection possibly in an attempt to directly constrain the bacteria within encapsulated structures. In the case of fish infected with ST23, however, observations differed. Low numbers of bacteria were observed by IHC within ellipsoids at day 1, day 3 and day 7, but only at day 14 was there positive labelling of bacterial antigen within MMCs. Melano-macrophage centres were not present in higher numbers from the ST23-infected fish and such "budding" of MMCs as observed with ST260 was not observed. In this case it appeared that macrophages from the spleen and possibly elsewhere migrated to pre-existing MMCs. This might be explained by the fact that ST23 does not appear to survive within macrophages resulting in no urgent need to encapsulate those directly after phagocytosis.

One ST260 infected fish sampled at day 14 p.i. was found to be co-infected by a bacterium other than *S. agalactiae;* this was identified as an *Aeromonas* species. *Aeromonas* species are a common part of intestinal microflora of fish and well recognised opportunistic pathogens responsible for secondary infections (Cahill, 1990; Hu et al., 2012). Among the other bacteria recovered and confirmed as *S. agalactiae*, it cannot be excluded that a distinct *S. agalactiae* was involved. In humans, *S. agalactiae* exists as part of the normal bacterial

flora in the recto-vaginal tract and it also causes neonatal disease through vertical transmission, suggesting adaptive changes to a new environment with up-regulation or activation of specific genes involved in virulence (Yang et al., 2010). S*treptococcus agalactiae* is considered therefore to be a commensal and potentially an opportunistic pathogen in humans. In fish, asymptomatic carriage of *S. agalactiae* has never been reported, but to our knowledge no study has ever evaluated this possibility. For this reason, it was important that bacteria recovered during the challenge not only were confirmed as *S. agalactiae*, but also were subjected to PFGE. The subset of isolates subjected to PFGE allowed a partial exclusion of secondary infection by strains of *S. agalactiae* other than the challenge strains.

This study provided further understanding of the survival of bacteria and progression of associated lesions during the time course of an infection, using two distinct *S. agalactiae* strains. In all cases, the formation of granulomata or other encapsulated structures appeared to be a major component of the fish response. However, the virulence of the two strains differed strongly; unlike the ST23 strain, the ST260 appeared to evade the immune system, possibly survived within macrophages, and caused major lesions in all infected fish. *Streptococcus agalactiae* is a multi-host pathogen, but molecular epidemiological studies demonstrated that specific clones are associated with specific niches, *e.g.* ST260 with poikilothermic hosts, ST67 with cattle, ST23 with bovines and humans (Chapter 2). Furthermore, experimental challenge studies indicate that these associations are not only due to differences in exposure, but that it may correspond to niche-dependant variations in the virulence of strains. Comparative genomic or transcriptomic studies and identification of genes putatively associated with virulence and adaptation in fish are required to provide further understanding into the fitness of specific *S. agalactiae* clones in fish.

3.6 REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Baya, M., Lupiani, B., Hetrick, F.M., Roberson, B.S., Lukacovic, R., May, E., Poukish, C., 1990. Association of *Streptococcus* sp. with fish mortalities in the Chesapeake Bay and its tributaries. J. Fish Dis. 13, 251-253.

Bishop, E.J., Shilton, C., Benedict, S., Kong, F., Gilbert, G.L., Gal, D., Godoy, D., Spratt, B.G., Currie, B.J., 2007. Necrotizing fasciitis in captive juvenile *Crocodylus porosus* caused by *Streptococcus agalactiae*: An outbreak and review of the animal and human literature. Epidemiol. Infect. 135, 1248-1255.

Bowater, R.O., Forbes-Faulkner, J., Anderson, I.G., Condon, K., Robinson, B., Kong, F., Gilbert, G.L., Reynolds, A., Hyland, S., McPherson, G., Brien, J.O., Blyde, D., 2012. Natural outbreak of *Streptococcus agalactiae* (GBS) infection in wild giant Queensland grouper, *Epinephelus lanceolatus* (Bloch), and other wild fish in northern Queensland, Australia. J. Fish Dis. 35, 173-186.

Buller, N.B. 2004. Bacteria from Fish and other Aquatic Animals: A Practical Identification manual, CABI Publishing, Oxfordshire, 361 pp.

Cahill, M.M., 1990. Bacterial flora of fishes: A review. Microb. Ecol. 19, 21-41.

Camus, A.C., Shewmaker, P.L., Mauel, M.J., Wise, D.J., 2008. *Streptococcus ictaluri* arthritis, osteolysis, myositis, and spinal meningitis in channel catfish broodstock. J. Aquat. Anim. Health 20, 54-62.

Chang, P.H. and Plumb, J.A., 1996. Histopathology of experimental Streptococcus sp. infection in tilapia, *Oreochromis niloticus* (L.), and channel catfish, *Ictafurus punctatus* (Ratinesque). J. Fish Dis. 19, 235-241.

Crumlish, M., Thanh, P.C., Koesling, J., Tung, V.T., Gravningen, K., 2010. Experimental challenge studies in Vietnamese catfish, *Pangasianodon hypophthalmus* (Sauvage), exposed to *Edwardsiella ictaluri* and *Aeromonas hydrophila*. J. Fish Dis. 33, 717-722.

Eldar, A., Bejerano, Y., Bercovier, H., 1994. *Streptococcus shiloi* and *Streptococcus difficile*: Two new streptococcal species causing a meningoencephalitis in fish. Curr. Microbiol. 28, 139-143.

Eldar, A., Bejerano, Y., Livoff, A., Horovitcz, A., Bercovier, H., 1995. Experimental streptococcal meningo-encephalitis in cultured fish. Vet. Microbiol. 43, 33-40.

Ellis, A.E., 1980. Antigen-trapping in the spleen and kidney of the plaice *Pleuronectes platessa* L. J. Fish Dis. 3, 413-426.

Evans, J.J., Klesius, P.H., Pasnik, D.J., Bohnsack, J.F., 2009. Human *Streptococcus agalactiae* isolate in Nile tilapia (*Oreochromis niloticus*). Emerg. Infect. Diseases 15, 774.

Evans, J.J., Bohnsack, J.F., Klesius, P.H., Whiting, A.A., Garcia, J.C., Shoemaker, C.A., Takahashi, S., 2008. Phylogenetic relationships among *Streptococcus agalactiae* isolated from piscine, dolphin, bovine and human sources: A dolphin and piscine lineage associated with a fish epidemic in Kuwait is also associated with human neonatal infections in Japan. J. Med. Microbiol. 57, 1369-1376.

Evans, J., Klesius, P., Gilbert, P., Shoemaker, C., Al Sarawi, M., Landsberg, J., Duremdez, R., Al Marzouk, A., Al Zenki, S., 2002. Characterization of β-haemolytic group B

Streptococcus agalactiae in cultured seabream, *Sparus auratus* L., and wild mullet, *Liza klunzingeri* (Day), in Kuwait. J. Fish Dis. 25, 505-513.

Evans, J.J., Pasnik, D.J., Klesius, P.H., Al-Ablani, S., 2006. First report of *Streptococcus agalactiae* and *Lactococccus garviae* from a wild bottlenose dolphin (*Tursiops truncatus*). J. Wildl. Dis. 42, 561-569.

Ferguson, H.W., Claxton, M.J., Moccia, R.D., Wilkie, E.J., 1982. The quantitative clearance of bacteria from the bloodstream of rainbow trout (*Salmo gairdneri*). Veterinary Pathology Online 19, 687-699.

Ferguson, H.W., 1976. The relationship between ellipsoids and melano-macrophage centres in the spleen of turbot (*Scophthalmus maximus*). J. Comp. Pathol. 86, 377-380.

Ferguson, H. W., 2006. Systemic Pathology of Fish: A Text and Atlas of Normal Tissues in Teleosts and Their Responses in Disease, 2nd ed., 366 pp. Scotian Press, London.

Ferguson, H. W., Morales, J., Ostland, V., 1994. Streptococcosis in aquarium fish. Dis. Aquat. Org. 19, 1-6.

Fromm, H.I. and Boor, K.J., 2004. Characterization of pasteurized fluid milk shelf-life attributes. J. Food Sci. 69, M207-M214.

Greisen, K., Loeffelholz, M., Purohit, A., Leong, D., 1994. PCR primers and probes for the 16S rRNA gene of most species of pathogenic bacteria, including bacteria found in cerebrospinal fluid. Journal of Clinical Microbiology 32, 335-351.

Hernández, E., Figueroa, J., Iregui, C., 2009. Streptococcosis on a red tilapia, *Oreochromis* sp., farm: A case study. J. Fish Dis. 32, 247-252.

Hu, M., Wang, N., Pan, Z.H., Lu, C.P., Liu, Y.J., 2012. Identity and virulence properties of aeromonas isolates from diseased fish, healthy controls and water environment in china. Lett. Appl. Microbiol. 55, 224-233.

Kawamura, Y., Itoh, Y., Mishima, N., Ohkusu, K., Kasai, H., Ezaki, T., 2005. High genetic similarity of *Streptococcus agalactiae* and *Streptococcus difficilis*: *S. difficilis* eldar et al. 1995 is a later synonym of S. agalactiae lehmann and neumann 1896 (approved lists 1980). Int. J. Syst. Evol. Microbiol. 55, 961-965.

Lusiastuti, A.M., Textor, M., Seeger, H., Akineden, Ö., Zschöck, M., 2012. The occurrence of *Streptococcus agalactiae* sequence type 261 from fish disease outbreaks of tilapia oreochromis niloticus in Indonesia. Aquacult. Res. In press.

Mian, G., Godoy, D., Leal, C., Yuhara, T., Costa, G., Figueiredo, H., 2009. Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. Vet. Microbiol. 136, 180-183.

Miles, A.A., Misra, S.S., Irwin, J.O., 1938. The estimation of the bactericidal power of the blood. Epidemiology & Infection 38, 732.

Musa, N., Wei, L.S., Musa, N., Hamdan, R.H., Leong, L.K., Wee, W., Amal, M.N., Kutty, B.M., Abdullah, S.Z., 2009. Streptococcosis in red hybrid tilapia (*Oreochromis niloticus*) commercial farms in Malaysia. Aquacult. Res. 40, 630-632.

Nocard, M. and Mollereau, R., 1887. Sur une mammite contagieuse des vaches laitieres. Ann. Inst. Pasteur 1, 109.

Payne, R.W., 2009. Genstat. Wiley Interdisciplinary Reviews: Computational Statistics 1, 255-258.

Pereira, U.P., Mian, G.F., Oliveira, I.C.M., Benchetrit, L.C., Costa, G.M., Figueiredo, H.C.P., 2010. Genotyping of *Streptococcus agalactiae* strains isolated from fish, human and cattle and their virulence potential in Nile tilapia. Vet. Microbiol. 140, 186-192.

Phuektes, P., Mansell, P.D., Browning, G.F., 2001. Multiplex polymerase chain reaction assay for simultaneous detection of *Staphylococcus aureus* and streptococcal causes of bovine mastitis. J. Dairy Sci. 84, 1140-1148.

Poyart, C., Pellegrini, E., Gaillot, O., Boumaila, C., Baptista, M., Trieu-Cuot, P., 2001. Contribution of Mn-cofactored superoxide dismutase (SodA) to the virulence of *Streptococcus agalactiae*. Inf. Immun. 69, 5098-5106.

Press, C.M. and Evensen, Ø., 1999. The morphology of the immune system in teleost fishes. Fish Shellfish Immunol. 9, 309-318.

Ralyea, R.D., Wiedmann, M., Boor, K.J., 1998. Bacterial tracking in a dairy production system using phenotypic and ribotyping methods. J. Food Prot. 61, 1336-1340.

Ramos-Vara, J.A., 2005. Technical aspects of immunohistochemistry. Veterinary Pathology Online 42, 405-426.

Rasheed, V. and Plumb, J.A., 1984. Pathogenicity of a non-haemolytic group B streptococcus sp. in gulf killifish (*Fundulus grandis* Baird and Girard). Aquaculture 37, 97-105.

Robinson, J.A. and Meyer, F.P., 1966. Streptococcal fish pathogen. J. Bacteriol. 92, 512-512.

Rothman, R.E., Majmudar, M.D., Kelen, G.D., Madico, G., Gaydos, C.A., Walker, T., Quinn, T.C., 2002. Detection of bacteremia in emergency department patients at risk for infective endocarditis using universal 16S rRNA primers in a decontaminated polymerase chain reaction assay. J. Infect. Dis. 186, 1677-1681.

Salloum, M., van der Mee-Marquet, N., Domelier, A., Arnault, L., Quentin, R., 2010. Molecular characterization and prophage DNA contents of *Streptococcus agalactiae* strains isolated from adult skin and osteoarticular infections. J. Clin. Microbiol. 48, 1261-1269.

Secombes, C.J. and Manning, M.J., 1980. Comparative studies on the immune system of fishes and amphibians: Antigen localization in the carp *Cyprinus carpio* L. J. Fish Dis. 3, 399-412.

Suanyuk, N., Kong, F., Ko, D., Gilbert, G.L., Supamattaya, K., 2008. Occurrence of rare genotypes of *Streptococcus agalactiae* in cultured red tilapia *Oreochromis* sp. and Nile tilapia *O. niloticus* in Thailand-Relationship to human isolates? Aquaculture 284, 35-40.

van Sorge, N.M., Quach, D., Gurney, M.A., Sullam, P.M., Nizet, V., Doran, K.S., 2009. The group B streptococcal serine-rich repeat 1 glycoprotein mediates penetration of the blood-brain barrier. J. Infect. Dis. 199, 1479-1487.

Yang, Q., Zhang, M., Harrington, D.J., Black, G.W., Sutcliffe, I.C., 2010. A proteomic investigation of *Streptococcus agalactiae* grown under conditions associated with neonatal exposure reveals the upregulation of the putative virulence factor C protein β antigen. Int. J. Med. Microbiol. 300, 331-337.

Ye, X., Li, J., Lu, M., Deng, G., Jiang, X., Tian, Y., Quan, Y., Jian, Q., 2011. Identification and molecular typing of *Streptococcus agalactiae* isolated from pond-cultured tilapia in China. Fish. Sc. 77, 623-632.

Zamri-Saad, M., Amal, M.N.A., Siti-Zahrah, A., 2010. Pathological changes in red tilapias (*Oreochromis* spp.) naturally infected by *Streptococcus agalactiae*. J. Comp. Pathol. 143, 227-229.

Zappulli, V., Mazzariol, S., Cavicchioli, L., Petterino, C., Bargelloni, L., Castagnaro, M., 2005. Fatal necrotizing fasciitis and myositis in a captive common bottlenose dolphin (*Tursiops truncatus*) associated with *Streptococcus agalactiae*. J. Vet. Diagn. Invest. 17, 617-622.

Zlotkin, A., Chilmonczyk, S., Eyngor, M., Hurvitz, A., Ghittino, C., Eldar, A., 2003. Trojan horse effect: Phagocyte-mediated *Streptococcus iniae* infection of fish. Infect. Immun. 71, 2318-2325.

CHAPTER 4

Comparative genomic analysis of *Streptococcus agalactiae*: evidence for niche restriction in a piscine lineage.

Christian M.J. Delannoy, Ruth N. Zadoks, Margaret Crumlish, Jimmy Turnbull, Michael C. Fontaine.

Status: This manuscript is in preparation for submission to a relevant peer-reviewed journal, whilst part of it has already been published as a Genome Announcement (Delannoy, C.M.J., Zadoks, R.N., Lainson, F.A., Ferguson, H.W., Crumlish, M., Turnbull J.F., Fontaine, M.C. 2012 Draft genome sequence of a non-haemolytic fish-pathogenic *Streptococcus agalactiae* strain. Journal of Bacteriology, 194:6341-6342). Some of this work has also been orally presented at the 3rd Research Conference of the Institute of Aquaculture (Oct. 2012; Stirling, UK).

Contribution: The candidate designed the study and performed the laboratory and analytical work, with the exception of the PCR screening (section 4.2.6.2.4.) which was undertaken by an MSc student (David Rodgers) under the supervision of the candidate. Some bioinformatics and phylogenetic advice were provided by Dr A. Lainson (MRI, UK) and Frank Wright (The James Hutton Institute, UK) respectively. The candidate wrote the present manuscript with corrections and editing from listed co-authors.

4.1 ABSTRACT

In addition to causing local or systemic infections in humans and mastitis in bovines, *S. agalactiae* is also a major cause of systemic disease in fish. Here we provide the first description for the genome sequence of a non-haemolytic *S. agalactiae* isolated from tilapia (strain STIR-CD-17) and belonging by multi-locus sequence typing (MLST) to clonal complex (CC) 552, which corresponds to a presumptive fish-adapted subgroup of *S. agalactiae*. The genome was compared to 13 *S. agalactiae* genomes from human (n=7), bovine (n=2), fish (n=3) and unknown (n=1) origins. Phylogenetic analysis based on the core genome identified isolates of CC552 as the most diverged of all *S. agalactiae*. Conversely, genomes from β-haemolytic isolates of CC7 recovered from fish were found to cluster with human isolates of CC7, further supporting the possibility that some strains may represent a zoonotic or anthroponotic hazard. Comparative analysis of the accessory genome enabled the identification of a cluster of genes uniquely shared between CC7 and CC552, and encoding for proteins possibly providing enhanced fitness in specific niches. Other genes identified were specific to STIR-CD-17 or to CC552 based on genomic comparisons; however the extension of this analysis through the PCR screening of a larger population of *S. agalactiae* suggested that some of these genes may occasionally be present in isolates belonging to CC7. Some of these genes, occurring in clusters, exhibited typical signatures of mobile genetic elements, suggesting their acquisition through horizontal gene transfer. It is not possible to date to determine whether genes were acquired through intraspecies transfer or through interspecies transfer from the aquatic environment. Finally, general features of STIR-CD-17 highlighted a distinctive genome characterised by an absence of well-conserved insertion sequences, an abundance of pseudogenes, a smaller genomic size than observed among human or bovine *S. agalactiae*, and an apparent loss of metabolic functions considered conserved within the bacterial species, indicating that the fish-adapted subgroup of isolates (CC552) has undergone niche restriction.

Keywords: Genome, *Streptococcus agalactiae*, tilapia, ST260

4.2 INTRODUCTION

Streptococcus agalactiae has been the subject of numerous typing studies over the past decades, with multi-locus sequence typing (MLST) generally accepted as the standard typing method for global population studies (Urwin and Maiden, 2003). Isolates were found to cluster into well-defined clonal complexes (CCs), with some of those corresponding to sub-populations associated with specific host-species or clinical manifestation. Among the different clonal complexes identified in humans, CC17 corresponds to a very homogenous group exclusively adapted to humans and that is over-represented among isolates responsible for invasive infections in neonates (Brochet et al., 2006). Other isolates found in humans mostly belong to CC1, CC19 and CC23, but no strong correlation between these isolates and human carriage or invasion was found (Jones et al., 2006). Moreover, isolates belonging to CC23 are occasionally identified in other host species, including bovines, and an expanded MLST scheme provided evidence for two distinct clusters of CC23 exclusively associated with either humans or bovines (Sørensen et al., 2010). In bovines, however, the vast majority of mastitis-causing isolates belong to CC67, which is considered to constitute the bovine-adapted subpopulation of *S. agalactiae* (Bisharat et al., 2004), even though more recent work shows that other strains may be involved (Sørensen et al., 2010; Zadoks et al., 2011). Finally, a putatively fish-adapted subpopulation (CC552) has recently been described (Chapter 2), whereas other isolates identified in fish belong to the human-associated CC7 and CC283 (Chapter 2; Evans et al., 2008).

In the past decade, seven *S. agalactiae* isolates from human origin and one isolate of uncertain origin have had their whole genome sequenced. These isolates belong to the major representative CCs encountered in humans (CC1, CC17, CC19 and CC23), in addition to the less common CC7. Comparative genomic analyses of these strains revealed a composite organisation with around 80% of genes shared by all strains - the core or backbone genome – and 20% of genes not present in all strains – the dispensable or accessory genome (Tettelin et al., 2005). A total of 358 genes specific to single strains were uncovered, and

prediction analyses further suggested that each additional *S. agalactiae* genome sequenced would reveal approximately 33 previously uncharacterised genes (Tettelin et al., 2005 and 2008). Many of the strain-specific genes cluster within 11 to 14 large variable regions called genomic islands (Glaser et al., 2002; Tettelin et al., 2005; Brochet et al., 2008a), most of which correspond to putative integrative conjugative elements and related elements (Brochet et al., 2008b). Through their plasticity and potential for gene acquisition and transfer, these genomic islands have literally shaped the evolution of *S. agalactiae* through acquisition of genes that may enhance fitness and virulence in specific environments (Brochet et al., 2008a; Tettelin et al., 2008). Furthermore, some genes may well have been acquired from other bacterial species (Davies et al., 2009; Tettelin et al., 2005). For example, a region unique to one of the seven *S. agalactiae* genomes was found to be shared with *S. pneumoniae*, *i.e.* a region that contains an α-galactosidase and a system for transport and metabolism of sugars which may be of importance for use of the host α-galactosides (Tettelin et al., 2005).

More recently, two isolates of *S. agalactiae* belonging to the bovine adapted sub-group (CC67) have had their genome sequenced. One of those genomes has been the subject of a comparative study against the human *S. agalactiae* genomes (Richards et al., 2011), which unveiled distinctive features with 183 genes believed to be specific to the bovine strain. PCR screening of a global collection of bovine and human *S. agalactiae* further supported that most of the specific genes are strongly correlated with a bovine origin. Some of those genes are suggested to contribute to virulence in bovines (*e.g.* LPXTG anchored proteins) or to provide metabolic advantages (*e.g.* lactose operon). As for the human *S. agalactiae*, these genes mostly clustered into major genomic islands, with strong support for them having been acquired horizontally (Richards et al., 2011). Furthermore, the authors provided evidence that some of these genes have been acquired by horizontal gene transfer between bovine *S. agalactiae* and other mastitis-causing *Streptococcus* species that are commonly found in cattle. Finally, the bovine *S. agalactiae* genome was found to contain a high

frequency of insertion sequences, which supports the hypothesis that an evolutionary bottleneck resulted in the isolation of this population (Richards et al., 2011).

Last year, a number of genome announcements describing *S. agalactiae* from fish were published (Wang et al., 2012; Liu et al., 2012). However, the selection of these isolates often appeared to be arbitrary, and no in-depth genomic analyses of these isolates have been published so far. In the present study, for whole genome sequencing, we rationally selected a fish isolate based on previous *in vitro* (Chapter 2) and *in vivo* (Chapter 3) work. The selected strain belongs to a fish-adapted subgroup of *S. agalactiae* (CC552) and was observed to be highly virulent in Nile tilapia following experimental intra-peritoneal infection. In addition to describing general features of this genome, we compared it to existing *S. agalactiae* genome sequences at the level of the core and the accessory genome, with a particular focus on the identification of genes specific to the fish strain. We subsequently screened additional isolates from distinct origins for the presence/absence of these genes to provide preliminary evidence for their putative importance for virulence or adaptation in fish.

4.3 MATERIALS & METHODS

4.3.1 Bacterial isolate

The *S. agalactiae* strain STIR-CD-17 used in this study was isolated from the heart of a moribund fish during a disease outbreak affecting farmed tilapia (*Oreochromis* sp.) in Honduras in 2008. It was selected among our collection of *S. agalactiae* from aquatic hosts based on previous epidemiological, phenotypic and genotypic characterisation (Chapter 2). The isolate is non-haemolytic, a trait that is relatively common among isolates from fish and cattle but rare among human isolates (Finch & Martin, 1984; Evans et al., 2008).The isolate belongs to sequence type (ST) 260 and CC552, corresponding to a cluster of strains that has been associated exclusively with disease in aquatic poikilotherms (fish and frogs) (Chapter 2; Lopez Sanchez et al., 2012). The strain was also subjected to standardized 3-set

genotyping developed for human *S. agalactiae* which is based on the determination of serotype and detection of 7 surface protein encoding genes (*bca*, *bac*, *alp1*, *alp2*, *alp3*, *alp4* and *rib*) and 7 mobile genetic elements (IS*1381*, IS*861*, IS*1548*, IS*Sa4*, IS*Sag1*, IS*Sag2* and GBSi1) using PCR (Chapter 2). STIR-CD-17 was found to be of serotype Ib but none of the surface protein genes or mobile genetic elements were present, further supporting the fact that it does not appear to be closely related to *S. agalactiae* of human origin. Finally, experimental challenge by intra-peritoneal injection showed that the isolate caused high morbidity and mortality in Nile tilapia (Chapter 3).

4.3.2 Growth conditions and DNA extraction

The bacterial isolate, which had been stored as a glycerol stock at -80˚C, was streaked onto 5% (v/v) sheep blood agar plates (SBA; E&O Laboratories, Bonnybridge, UK) and grown aerobically at 28°C for 72h in order to assess purity and haemolytic properties. The isolate was Gram-stained and confirmed to belong to Lancefield Group B by a group B-specific latex agglutination test (Slidex Strepto Plus B; bioMérieux, Marcy L'Étoile, France). A single colony was then selected and used to inoculate 5ml of Brain Heart Infusion broth (BHI; Oxoid, Basingstoke, United Kingdom). After overnight static incubation at 28˚C in an aerobic environment, genomic DNA was extracted from cells harvested from 1ml of BHI culture using an Epicentre MasterPure Gram-positive DNA purification kit (Epicentre, Madison, WI) with modifications for optimal recovery of *S. agalactiae* DNA. Briefly, bacterial culture was centrifuged, supernatant was removed and cells were re-suspended in 150µl of TE Buffer (10 mM Tris-hydrochloride buffer, pH 7.5, containing 1.0 mM EDTA), 6µl of mutanolysin (5U/µl; Sigma-Aldrich, Irvine, UK) and 1µl of ready to use lysozyme (as provided by the manufacturer), and with incubation at 37°C for 1 hour. The remainder of the protocol was performed according to manufacturer's instructions, with the exception of the Proteinase K/Gram Positive Lysis Solution incubation time (30 min instead of 15 min) and the RNase incubation time (1 hr instead of 30 min). DNA concentration was quantified using a NanoDrop 1000 (Thermo Scientific) and the product (0.5µg) visually examined on a

Gel Red (Cambridge Bioscience, Cambridge, UK) stained 0.8% agarose gel (running conditions: 100V/cm for 1 hr) in order to assess the absence of DNA shearing.

4.3.3 Genome sequencing, assembly and annotation

Genome sequencing (100bp paired-end reads) was performed using an Illumina Solexa Genome Analyzer by the Genepool sequencing core facility (University of Edinburgh, UK). *De novo* assembly of Solexa reads was achieved by Genepool using VELVET 0.6 and yielded 208 contiguous sequences (Contigs). Because of time constraints, no attempt was made to close the genome or to organise the Contigs into a meaningful order. Following instructions from the National Centre for Biotechnology Information (NCBI) Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP; http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline), all Contigs of less than 200 nucleotides were removed and the remaining 96 Contigs were submitted for annotation through the PGAAP. The draft genome sequence of *S. agalactiae* STIR-CD-17 has been deposited in GenBank under the accession number ALXB00000000.

4.3.4 Genomic analysis for strain identity confirmation

The molecular typing results obtained by PCR analysis of STIR-CD-17 (described in Chapter 2 and summarized in section 4.3.1.) were all evaluated by *in-silico* analysis of the genome.

Firstly, the species identity was assessed by identification of the *cfb* gene and genes involved in group B carbohydrate biosynthesis within the STIR-CD-17 genome. For MLST, the sequence of the seven housekeeping genes described by Jones et al. (2003) were retrieved from the genome and submitted to the MLST database (http://pubmlst.org/) for locus type and ST assignment.

The presence/absence of the 7 surface protein-encoding genes and 7 mobile genetic elements assessed as part of the 3-set genotyping was evaluated using BioEdit software

(Hall, 1999; http://www.mbio.ncsu.edu/bioedit/bioedit.html). A local nucleotide database file for STIR-CD-17 was created within BioEdit, thereby enabling the identification of the presence of known nucleotide sequences within STIR-CD-17 through the Basic Local Alignment Search Tool (BLASTN). Reference nucleotide sequences for the 7 surface protein-encoding genes and mobile elements were extracted from NCBI (GenBank accession numbers as provided in Chapter 2, **Table 2.1**) and BLAST searched against the STIR-CD-17 genome using BioEdit.

The 3-set genotyping system also involves identification of the capsular serotype, for which the genetic basis lies within the *cps* genes. The *cps* genes from STIR-CD-17 were investigated through BLASTN search analysis (http://www.ncbi.n-lm.nih.gov/BLAST) to compare them with sequences from the GenBank database. The complete *cps* locus of STIR-CD-17 was also compared to that of H36B (serotype Ib) and A909 (serotype Ia) using the Artemis Comparison Tool (ACT; Carver et al., 2008) provided by the Sanger Centre (http://www.sanger.ac.uk/Software/ACT). This software requires genome or Contig information to be provided as Fasta and GeneBank format files, in addition to a comparison file. To generate comparison files, the DOUBLE ACT v2 web interface (http://www.hpa-bioinfotools.org.uk/pise/double_act.html) was used using BlastN and default settings.

Finally, the *cyl* operon which is the genetic basis for beta-hemolysin/cytolysin production (Pritzlaff et al., 2001) was analysed in order to investigate whether it would provide an explanation for the non-haemolytic phenotype. For this, nucleotide sequence from STIR-CD-17 was compared to that of the non-haemolytic isolate ATCC13813 (www.atcc.org/) and that of the beta-haemolytic isolate A909 using ACT.

4.3.5 General genomic analysis for prediction of sub-cellular localization and function of putative proteins

In silico prediction of sub-cellular localization of proteins encoded by the genome was performed using the PSORTb program version 3.0.2 (http://psort.org; Yu et al., 2010),

employing the model built for Gram positive bacteria. PSORTb is able to categorize predicted proteins in one of the following localization sites: cytoplasm, cytoplasmic membrane (inner membrane), outer membrane, and extracellular space. However, the program does not necessarily lead to a specific prediction and the localization site can be designated as unknown.

Functional categories of proteins were provided by the PGAAP based on the analysis of Clusters of Orthologous Genes (COGs) using the COGnitor program (Tatusov et al. 2001) (http://www.ncbi.nlm.nih.gov/COG/xognitor.html). In order to assess the putative function of hypothetical proteins with no associated COG, nucleotide sequences were screened for conserved Pfam domains (http://pfam.sanger.ac.uk/search).

4.3.6 Comparative genomic analysis

The complete list of annotated *S. agalactiae* genomes publicly available from GenBank (Feb. 2012) and used in this study for comparative genomic analyses is reported in **Table 4.1**. It consists of seven *S. agalactiae* of human origin, two *S. agalactiae* of bovine origin and one *S. agalactiae* of unknown origin. Moreover, the genomes of three *S. agalactiae* of piscine origin that have been made available more recently (after May 2012) have been added to the table and included in some of the analyses.

Table 4.1 Origin, clonal complex (CC), sequence type (ST), serotype and accession number for genomes used in the comparative genomic studies.

Host of origin	Strain ID	Source	Country	CC	ST	Serotype	Accession Number	Key references
Unknown	**NEM316**	NA	NA	23	23	III	NC_004368	Glaser et al., 2002; Sørensen et al., 2010
Human	COH1	Blood	NA	17	17	III	NZ_AAJR00000000	Tettelin et al., 2005
	CJB111	Blood	NA	1	1	V	NZ_AAJQ00000000	Tettelin et al., 2005
	2603V/R	NA	NA	19	110	V	NC_004116	Tettelin et al., 2002
	18RS21	Umbilicus	USA	19	19	II	NZ_AAJO00000000	Tettelin et al., 2005; Lancefield et al., 1975
	515	Cerebrospinal fluid	NA	23	23	Ia	NZ_AAJP00000000	Tettelin et al., 2005
	A909	Umbilicus	NA	7	7	Ia	NC_007432	Tettelin et al., 2005; Lancefield et al., 1975
	H36B	Umbilicus	USA	7	6	Ib	NZ_AAJS00000000	Tettelin et al., 2005; Lancefield et al., 1975
Bovine	FSL S3-026	Milk	USA	67	67	III	AEXT00000000	Richards et al., 2011
	ATCC 13813	Milk	NA	67	61	II	AEQQ00000000	NA
Fish (Nile tilapia)	ZQ0910	NA	China	7	7	NA	AKAP01000000	Wang et al., 2012
	GD201008-001	NA	China	7	7	Ia	CP003810.	Liu et al., 2012
	SA20-06	Brain	Brazil	552	553	Ib	NC_019048	Mian et al., 2009; Pereira et al., 2010
	STIR-CD-17	Heart	Honduras	552	260	Ib	ALXB00000000	Delannoy et al., 2012

Strains in bold indicate those for which the complete genome is available. The STs and corresponding CCs are presented as provided in the key references. For strains with no ST report (ZQ0910, SA20-06 and ATCC13813), MLST was carried out in silico by extracting from the genomes the seven housekeeping genes described by Jones et al. (2003) and submission of those into the MLST database (http://pubmlst.org/) for locus type and ST assignment. NA, Not Available.

4.3.6.1 Comparative genomic analysis of the core genome

In order to automatically extract, concatenate and align nucleotide sequences from the core genome of *S. agalactiae,* we used the Panseq v.2.0. Web server (http://lfz.corefacility.ca/panseq/; Laing et al., 2010) which is a multifunctional Web-based tool designed to analyse closely-related genomes. For this, the "Core/Accessory Genome Analysis Function" was selected and genomic sequences loaded within the interface. The program was configured using default settings except that the core genome threshold was set to 10 so that any region not found in all 10 genomes was removed. The Web server only allows the analysis of 10 genomic sequences which were selected from those listed in **Table 4.1** all *S. agalactiae* genomes from fish (n=4) and bovines (n=2), and four genomes from human origin (515, COH1, H36B and A909). The resulting nexus file for the core genome was then loaded into SEAVIEW (Gouy et al., 2010). The Gblocks program (Castresana, 2000) implemented within SEAVIEW was used to discard ambiguous regions and gaps, followed by a degree of manual refining. The final alignment of 803,614 bp-long sequences was then subjected to phylogenetic analysis using the TOPALi v2.5 package (Milne et al., 2004). For this, the model was optimised using the model selection feature in TOPALi which ensured the selection of the best model parameters for generating trees with our specific dataset. The selected model (general time reversible model [GTR] combined with the models of gamma distribution [G] and invariable sites [I], referred to as GTR+G+I; Rodríguez et al., 1990) was then used to estimate a Bayesian phylogenetic tree using the MrBayes program (Ronquist et al., 2003) launched from TOPALi. The MrBayes settings were 2 runs of 625,000 generations and a burn-in period of 125,000 generations, with trees sampled every 10 generations and inclusion of invariable sites. The consensus tree was then imported into DENDROSCOPE v3.2.1 (Huson et al., 2007), where visualization and editing were performed as required.

A second tree was constructed in which the number of genomes representing the same CC was reduced in order to include other genomes; 2 *S. agalactiae* genomes from fish (STIR-

CD-17 and GD201008-001), 1 genome from bovine (FSL S3-026), 4 genomes from humans (COH1, 18RS21, 2603V/R, 515, CJB111 and H36B) and 1 genome of unknown origin (NEM316) were used. The tree was generated identically as previously described, except that it was built on a core genome alignment of 921,707 bp-long sequences.

4.3.6.2 Comparative genomic analysis of the accessory genome

4.3.6.2.1 GeneRator analysis

A comparative study was undertaken using an in-house software, developed at the Moredun Research Institute and called GeneRator. GeneRator is a graphical tool for comparing predicted proteins in annotated genomes (**Figure 4.1**). As an input it requires a list of predicted protein sequences for each genome (in Fasta format), including one reference genome and numerous comparison genomes. GeneRator performs a reciprocal BLAST comparison between each predicted protein of the reference genome against those of the other genomes. Moreover, for each position it then calculates if protein pairs are reciprocal best hits to provide an estimate of the homology between amino-acid sequences. The output is graphical with the predicted proteins from the reference genome displayed on the top line and their corresponding best BLAST hit within the comparison genomes shown below. Proteins are represented in blocks that are colour coded. Block colours within the reference genome indicate the putative general function of the predicted proteins based on COG classification. Block colours within the comparison genomes indicate the level of homology (BLASTP score) between the predicted reference and comparison proteins, with the border colour indicating whether or not these sequences are reciprocal best hits. The reference genome used in this analysis was STIR-CD-17 and the comparison genomes were the 10 *S. agalactiae* genomes publicly available from GenBank at the time of analysis (March 2012), consisting of seven isolates of human origin, two isolates of bovine origin and one isolate of unknown origin (**Table 4.1**). Based on this analysis, predicted protein sequences from STIR-CD-17 that did not find a reciprocal best hit with a BLAST score > 80 in the other genomes were identified and their corresponding genes considered putatively strain-specific

or fish-specific. Predicted amino acid sequences from STIR-CD-17 that did find a reciprocal best hit with a BLAST score > 80 in genomes belonging to CC7 (A909 and H36B) only were also identified, and their corresponding genes considered putatively fish-associated because isolates belonging to CC7 occur in fish (Chapter 2).

4.3.6.2.2 Normalised BLAST score ratio analysis

To assess the homology between predicted proteins whilst correcting for query length to avoid skewed results (as would be obtained with E-values), a normalised BLAST score ratio (BSR) analysis was used. The BSR analysis performed here was a modification of the technique described by Rasko et al. (2005), which enables assessment of relatedness of predicted proteins using BLAST. Predicted proteins of STIR-CD-17 selected as putatively fish-specific or fish-associated based on GeneRator analysis (section 4.3.6.2.1) were BLASTP searched against the NCBI reference proteins database (http://www.ncbi.n-lm.nih.gov/BLAST; last accessed 13 Jan. 2013). For each protein, the BLASTP bit-score for the alignment against itself (REF_SCORE) and for the most similar proteins within the database (QUE_SCORE) was obtained and normalized by dividing the QUE_SCORE by the REF_SCORE. Amino acid sequences with a normalized bit-score ≥ 0.8 were considered as homologous. Normalized bit-scores < 0.8 were an indication of divergence ($0.4 < BSR < 0.8$) or uniqueness ($BSR, \leq 0.4$) (Rasko et al., 2005).The BSR analysis enabled to confirm results obtained through GeneRator analysis. Moreover, the BSR analysis was extended to the whole NCBI database, including other *Streptococcus* species as well as the new *S. agalactiae* genomes from fish (**Table 4.1**) which had been lodged with GenBank after completion of the GeneRator analysis.

4.3.6.2.3 ACT comparisons

Using GeneRator and BSR analyses, genes of interest were selected based on comparative studies of their corresponding proteins. However it was recognised that such analyses might be biased by annotation discrepancies between genomes and it was therefore necessary to

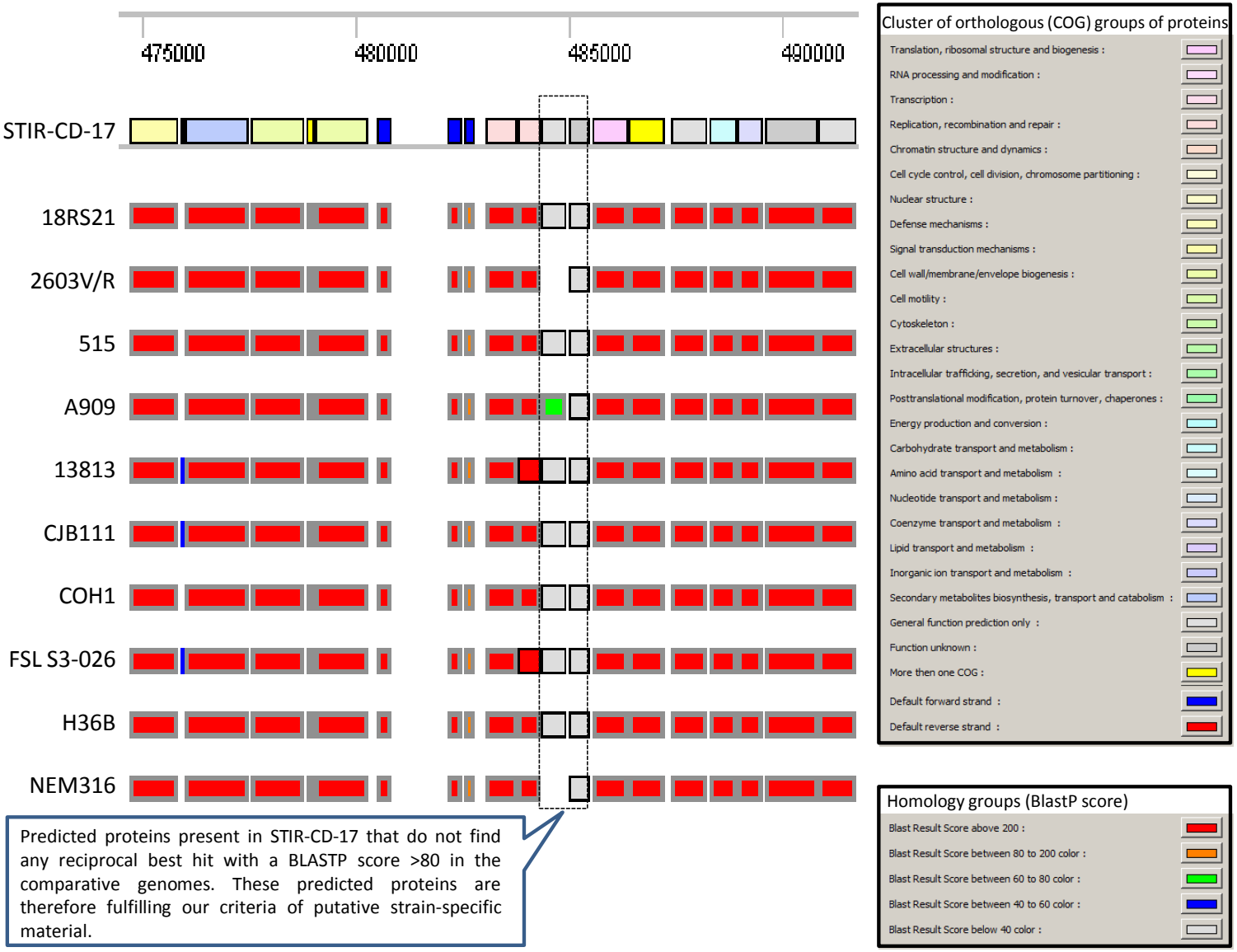
confirm observations at the nucleotide level. For this, genome sequences were compared pairwise using the Artemis Comparison Tool (ACT) as previously described (section 4.3.4). The ACT comparison not only allowed for confirmation of presence/absence of genes within genomes, but it also provided insights into the genetic organisation and conservation of sequences flanking regions of interest. Due to time constraints, ACT comparisons were performed against a limited number of other genomes (minimum of 3) which were selected based on individual findings.

4.3.7 Distribution of selected genes across *S. agalactiae* from different host species

Our 3-step comparative analysis (GeneRator/BSR/ACT) led to the identification of candidate strain-specific genes (unique to STIR-CD-17), fish-specific genes (unique to isolates belonging to CC552; STIR-CD-17 and SA20-06) and fish-associated genes (unique to isolates belonging to CC7 and CC552; A909, H36B, ZQ0910, GD201008-001, STIR-CD-17 and SA20-06).

Figure 4.1 (next page) Example of the GeneRator software output from comparison of *S. agalactiae* genomes. Predicted proteins from the reference genome (STIR-CD-17) are displayed as blocks on the top line and the position of their corresponding gene within the genome is indicated above these blocks. These blocks are colour coded based on putative general function provided by the Cluster of Orthologous Groups (COG) classification. Results of the reciprocal BLASTP comparison between predicted proteins from the reference genome against proteins from the comparison genomes are provided as block in the following 10 lines. Block colours in the comparison genomes indicate the level of homology (BLASTP score) between the reference and comparison sequences whilst the border colour indicates whether these sequences are reciprocal best hits. The border colour can be grey or black; the reciprocal best hits are grey only. Based on this analysis, protein sequences from STIR-CD-17 that did not find reciprocal best hits with a BLAST score > 80 in the other genomes were retained and considered as putatively strain-specific or fish-specific. An example is shown for predicted proteins of unknown function.

Comparative genomic analysis



Conservation of genes of interest was assessed among a panel of *S. agalactiae* of bovine and aquatic origin. Internal primers were designed for 1 strain-specific gene, 1 fish-specific gene and 3 fish-associated genes using the Primer Select module of the DNAStar computer program from Lasergene (DNASTAR Inc., Madison, USA). A description of selected genes and primers is provided in **Table 4.2**. Although false negatives could be obtained due to nucleotide polymorphism at the priming sites, the relatively low level of nucleotide divergence observed at the intra-specific level of *Streptococcus* species should render this a rare event (Richards et al., 2011). The aquatic *S. agalactiae* selected for PCR screening corresponded to the 26 fish isolates and 6 sea mammal isolates from the collection described in Chapter 2 (section 2.3.1). The bovine isolates were rationally selected from a collection of 71 isolates based on whole genome typing by PFGE and core genome typing by MLST (unpublished data, R.N. Zadoks). A final panel of 19 isolates with distinct PFGE patterns and including 4 different STs were assessed in this study; these were all obtained from quarter milk samples collected from dairy cows in Denmark (Isolates provided by Y. Mahmod, Faculty of Health and Medical Sciences, Copenhagen University, Denmark).

The identity of all bacterial isolates was confirmed and chromosomal DNA prepeared as previously described (section 2.3.1). PCR reactions were performed in a final 25µl volume containing 12.5µl of GoTaq Green Master Mix (Promega, Madison, USA), 0.25 µM of each primer and 2 µl of DNA template. The thermal cycling was conducted as follows: denaturation of DNA at 94°C for 5 min followed by 35 cycles of 94°C for 1 min, 49 ^{o}C to 61°C for 45 sec and 72°C for 30 sec, with a final step at 72°C for 7 min. The annealing temperature varied between 49^{o}C and 61^{o}C depending on the primers used (melting temperatures supplied by the manufacturer - Eurofins MWG Operon, Munich, Germany). The PCR products for all reactions were visualized following electrophoresis through a 1.5% (w/v) agarose gel containing Gel Red (Cambridge Bioscience, Cambridge, UK).

Table 4.2. Putative *S. agalactiae* strain-specific, fish-specific and fish-associated gene targets and the PCR primers used for their detection.

Target	Locus tag	Primer sequence	Tm
Strain-specific gene			
Beta-hydroxyacyl dehydratase	M3M_04280	71 AAATAATCCGATTGTTCCTG 91	51.2
FabA/FabZ		346 ATATACTATAAATTTCCCTTCTAA 321	50.8
Fish-specific gene			
Putative RecB familly exonuclease	M3M_01062	5 CTATGCCGAAGATGATTATTTG 28	54.7
(cas4)		491 CTTCTTGGCGTAGTTCCTCAGTA 467	60.6
Fish-associated genes			
Galactokinase	M3M_01167	466 AAATCGGCAAGCAGACAGAAAATGAAT 494	60.4
		1041 GCAATAGCACAACCGCCAAAACC 1017	62.4
Alpha galactosidase	M3M_01172	663 AAGGGTGCTAGTAGTGCCGAACATAAT 691	63.4
		1113 AACCAGCCATCATCCATAACAAAAAGT 1085	60.4
Sugar ABC transporter permease	M3M_01182	489 ATTGGTATTTGGAGCACTGTAGG 513	58.9
		740 TCTTATTATAGGCCGGACTTGTA 716	57.1

Melting temperatures (Tm) were provided by the manufacturer (Eurofins MWG).

4.4 RESULTS AND DISCUSSION

4.4.1 Genome structure and general features

The draft genome was composed of 96 Contigs and contained a total of 1,805,303 nucleotide residues. Complete and draft genomes of *S. agalactiae* from human origin range from 2.09 Mb (*S. agalactiae* 515) to 2.3 Mb (*S. agalactiae* A909), whereas those from bovine origin range from 2.11 Mb (*S. agalactiae* ATCC13813) to 2.46 Mb (*S. agalactiae* FSL S3-026). The STIR-CD-17 genome is therefore of reduced size in comparison to other genomes, an observation which might either be real, or instead indicate its incomplete nature. New evidence based on complete *S. agalactiae* genome sequences from fish suggests that strains belonging to the putatively fish-adapted subgroup of *S. agalactiae* (CC552; Chapter 2), like STIR-CD-17, have a genome of reduced size. Indeed, the complete SA20-06 strain (ST553) is approximately 1.82 Mb, whereas the 2 other *S. agalactiae* genomes from fish, that belong to a human-related subgroup (CC7; Chapter 2) have a size larger than 2Mb. In STIR-CD-17, the average G+C content (35%) and the number of rRNA and tRNA genes (21 and 80 respectively) are lower than previous reports from other *S. agalactiae* genomes. Such observations are not supported by the complete SA20-06

genome, suggesting that the G+C content might be biaised by the incompleteness of the STIR-CD-17 genome.

Altogether a total of 1,697 protein-encoding genes were predicted automatically, with 506 (29.8%) of the predicted products being annotated as hypothetical proteins. In addition, 102 pseudogenes, in which frameshift and nonsense mutations introduced multiple stop codons throughout the gene, were identified. The 1,697 predicted genes of *S. agalactiae* were sorted with respect to COG classification (**Table 4.3**), which enables the allocation of general functions to predicted proteins, including some of those annotated as hypothetical. A total of 352 (20.73%) genes were determined to be associated with information storage and processing, 268 (15.8%) with cellular processes and signaling, and 548 (32.3%) with metabolism. Finally, a total of 529 (31.2%) residual genes, which could not be categorized into COG classes, had poorly characterized functions and features.

In addition, a report from the PGAAP revealed that genes with functions considered as conserved amongst the *S. agalactiae* species (core genome) are missing from the STIR-CD-17 genome. This concerns genes with the following functions: 1,4 dihydroxy-2-naphthoate octaprenyltransferase, peptide deformylase, glutamate dehydrogenase, acid phosphatase/phosphotransferase, preprotein translocase subunit SecY, and fructose-6-phosphate aldolase. Whether or not this information corresponds to a true absence would require further assessment (see Chaper 4 section 4.4.6 and Chapter 5 section 5.4.1.3).

4.4.2 Strain identity confirmation

Using genomic data, STIR-CD-17 was confirmed as *S. agalactiae* and belonging to Lancefield group B by identification of the conserved group B carbohydrate biosynthesis locus and the *cfb* gene, both of which are considered to be unique to the species *S. agalactiae* (Podbielski et al., 1994; Sutcliffe et al., 2008).

Table 4.3 Distribution of predicted proteins among functional categories based on COG classification.

Category	Code	Functional classification	Annotated genes	
			No.	%
Information storage and processing				**20.7**
	J	Translation, ribosomal structure, and biogenesis	146	8.6
	A	RNA processing and modification	0	0.0
	K	Transcription	110	6.5
	L	Replication, recombinaison and repair	96	5.7
	B	Chromatin structure and dynamic	0	0.0
Cellular processes and signaling				**15.8**
	D	Cell cycle control, cell division and chromosome partitioning	17	1.0
	Y	Nuclear structure	0	0.0
	V	Defense mechanisms	43	2.5
	T	Signal transduction mechanisms	46	2.7
	M	Cellwall/membrane/envelope biogenesis	89	5.2
	N	Cell motility	5	0.3
	Z	Cytoskeleton	0	0.0
	W	Extracellular structures	0	0.0
	U	Intracellular trafficking, secretion, and vesicular transport	17	1.0
	O	Posttranslational modification, protein turnover, and chaperones	51	3.0
Metabolism				**32.3**
	C	Energy production and conversion	45	2.7
	G	Carbohydrate transport and metabolism	140	8.2
	E	Amino acid transport and metabolism	131	7.7
	F	Nucleotide transport and metabolism	65	3.8
	H	Coenzyme transport and metabolism	40	2.4
	I	Lipid transport and metabolism	40	2.4
	P	Inorganic ion transport and metabolism	70	4.1
	Q	Secondary metabolite biosynthesis, transport and catabolism	17	1.0
Poorly characterized				**31.2**
	R	General functional prediction only	178	10.5
	S	Function unknown	141	8.3
	-	Not annotated	211	12.4

The strain was found to be of ST260 by PCRs and sequencing of PCR products; this result was further assessed using genomic data. All locus types obtained from *in silico* analysis corresponded to those of ST260 with the exception of the *tkt* gene. The ST260 has allele 19 at the *tkt* locus, but the sequence extracted from the genome differed from allele 19 by one nucleotide. According to the genome sequence, STIR CD-17 is therefore not of ST260 but would be a single locus variant of ST260. It is unknown if this inconsistency is due to a sequencing error or to a point mutation that might have occurred after repeated sub-culture. Similar discrepancies between isolate locus type based on PCR and amplicon sequencing versus extraction of sequence from whole genome sequence data have also been observed for *Streptococcus canis* (R.N. Zadoks, personal communication).

Previous PCR-based typing of STIR-CD-17 using a standardised 3-set genotyping approach revealed the absence of all surface protein-encoding genes (SPGs) and mobile genetic elements (MGEs) analysed. BLAST analysis of those sequences against the STIR-CD-17 genome further indicated their absence from this strain, with one exception; sequences homologous to the insertion sequence IS*Sag1* were found fragmented within multiple small Contigs, suggesting that this MGE or remnants of it may be present. Sequences containing the primer annealing sites were identified (Contig14 and Contig55). Alignment of those Contigs with the reference sequence AF329276 (Suanyuk et al., 2008; **Table 2.1**) revealed some level of divergence, including at the priming site for one of the primers (9 nucleotide differences out of 40) which might explain the negative results obtained by PCR. IS*Sag1* spans from base positions 318 to 1633 within the reference sequence; Contig55 covers bases 382 to 586, Contig14 covers bases 568 to 1010, Contig15 covers bases 960 to 1228 and Contig20 covers bases 1374 to 1632 of the reference sequence. Therefore, parts of the IS are found suggesting the presence of a single or multiple copies of ISSag1 within STIR-CD-17. In order to support these observation, the presence of the IS was evaluated within SA20-06 and a single copy of the IS (SaSA20_1656) with 90% identity match with the reference sequence (Table 2.1 in Chapter 2) was found.

The 3-set genotyping system also included the identification of the capsular serotype. For this, the *cps* locus was investigated by BLAST search and compared pairwise with the *cps* locus from A909 (Ia) and H36B (Ib) using ACT. Specific *cps* genes organisation and homology as previously described (Cieslewicz et al., 2005) confirmed our strain to be of serotype Ib, even though previously unreported gene polymorphism was observed in the *cps*K gene (see Chapter 5, section 5.4.3.1).

Finally, the *cyl* operon was analysed in order to assess whether it would provide an explanation for the absence of haemolysis as observed when growing the bacterium on blood agar plates. Genomic analysis of the only non-haemolytic strain available (the bovine

strain 13813) revealed that it completely lacks the *cyl* operon. In STIR-CD-17, genomic analysis showed that several of the *cyl* genes are either absent or partially present within the operon, which might explain the phenotypic observation (see Chapter 5, section 5.4.2.1).

In summary, results of *in silico* MLST and 3-set genotyping are largely consistent with those obtained by PCR, confirming the isolate's identity and providing new information on the putative presence of a divergent insertion sequence.

4.4.3 Comparative genomic analysis of the core genome

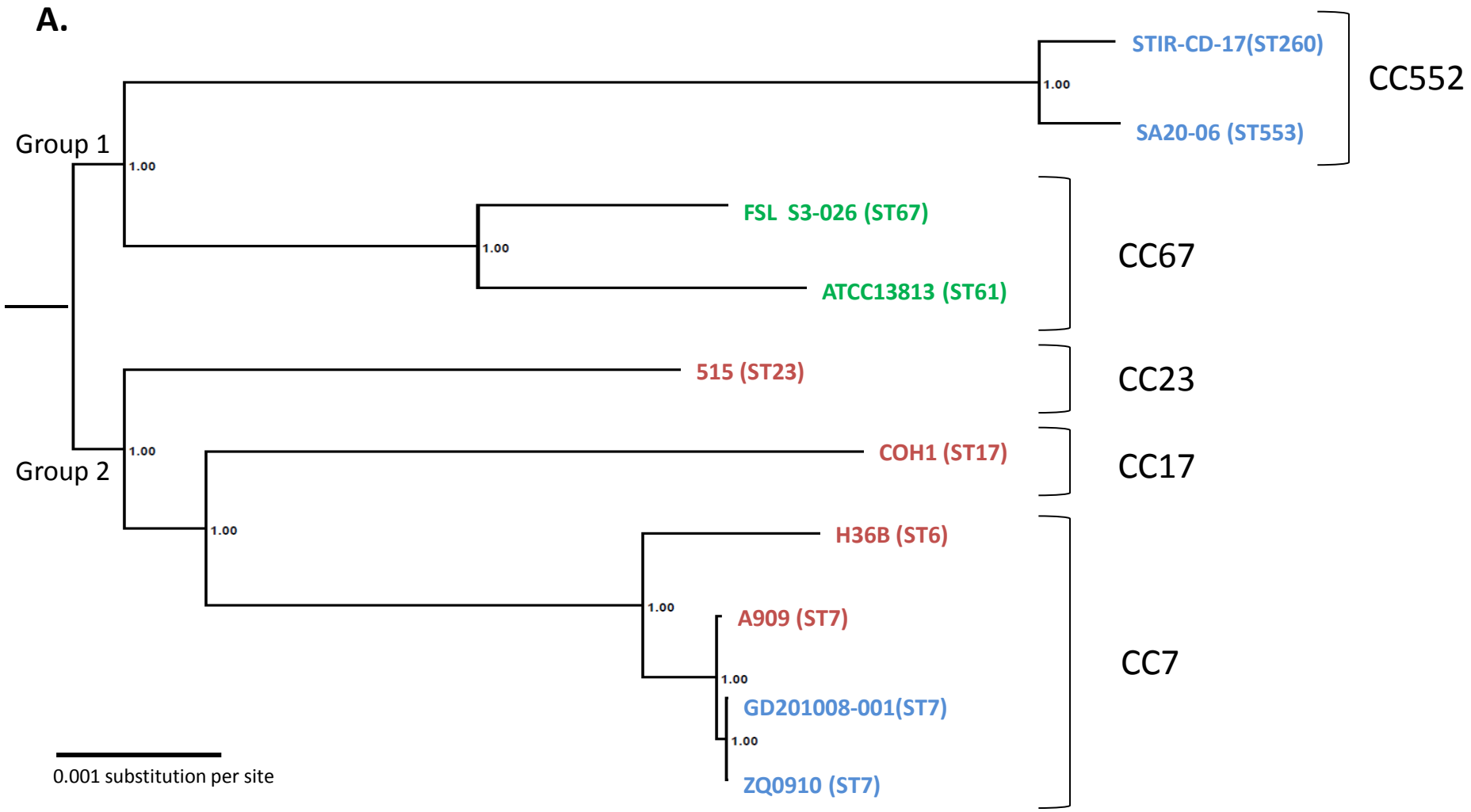
Bayesian phylogenetic trees were constructed based on the core genome of 10 *S. agalactiae* isolates from human, bovine and fish origins. The first tree (**Figure 4.2, A**) showed two primary isolate groups. Within each primary group, the isolates formed sub-groups which strongly agreed with clonal complex groupings based on MLST, with two sub-groups (CC552 and CC67) in group 1, and 3 sub-groups (CC23, CC17 and CC7) within group 2. A second tree was constructed in which the number of genomes representing CC7, CC67 and CC552 was reduced in order to include other genomes representing other CCs (CC1, CC19) or, in the case of CC23, intra-CC diversity (**Figure 4.2, B**). Clustering based on core genome continued to be consistent with grouping based on MLST, even though the latter only represents a small subset of the core genome. However, phylogenetic relationships between CCs are not always accurately reflected by MLST (Sørensen et al., 2010). For example, phylogenetic analysis based solely on MLST suggests a close relationship between the ST17 responsible of neonatal invasive disease and the bovine-adapted CC67, and authors have hypothesised that the hyper-invasive neonatal clone has arisen from a bovine ancestor (Bisharat et al., 2004). However, subsequent studies considering a larger part of the genome did not support this close relationship between these ST17 and CC67 (Brochet et al., 2006; Sørensen et al., 2010). It has been demonstrated that the un-even distribution of the 7 housekeeping genes used for MLST, which are located primarily in one half of the genome, is responsible for phylogenetic inaccuracies (Sørensen et al., 2010). Our tree based

on alignment of the core genome (> 0.8 Mb) does not support any close relationship between isolates of ST17 and CC67, further supporting the observation of others (Brochet et al., 2006; Sørensen et al., 2010).

Isolates belonging to CC552 have only been recovered from fish (and a frog) and are believed to be fish-adapted (Chapter 2). Sequence types belonging to that CC do not share any single locus variants, double locus variants or triple locus variants with STs of isolates recovered from mammals (Chapter 2), indicating that these isolates form a largely distinct sub-population. In our tree, the fish isolates from CC552 formed the most divergent isolate group, which is in agreement with our MLST observations and results published elsewhere (Chapter 2; Brochet et al., 2006). However, the tree published by Brochet et al. (2006), which is based on comparative genomic hybridization, suggests that all mammal isolates form a monophyletic group separated from the fish isolates, whereas our tree suggests that the fish and bovine-adapted strains share a more recent common ancestor than the bovine-adapted and human isolates.

Isolates of ST7 have been identified in humans as a cause of invasive disease in neonates and adults (Jones et al., 2003) or as part of the natural vaginal microbiota (Haguenoer et al., 2011). Despite extensive studies of *S. agalactiae* in cattle, ST7 has been reported only once as a cause of mastitis in this species (Haguenoer et al., 2011). Isolates belonging to ST7 are therefore considered to be mostly human-associated. However, piscine isolates have been subjected to MLST in recent years and isolates belonging to CC7 (ST7 and its SLV ST500) have been reported from outbreaks in Kuwait, China and Thailand (Evans et al., 2008; Ye et al., 2011; Chapter 2). Moreover, a human neonatal meningitis isolate of ST7 was found to be pathogenic in fish after intra-peritoneal injection (Evans et al., 2009). Comparing the mortalities as obtained in challenge studies by intra-peritoneal infection of tilapia, piscine and human strains belonging to CC7 appear to be less virulent in fish than those belonging to CC552 (Chapter 3; Evans et al., 2002). Our phylogenetic tree supports the observation

that piscine and human *S. agalactiae* of CC7 are closely related based on analysis of the core genome. This observation suggests that strains belonging to CC7 may have retained genes for virulence and adaptation in fish or have horizontally acquired those genes from isolates of CC552 or elsewhere. Comparative analysis of the accessory genome for identification of genes unique to isolates of CC552 or shared with CC7 may provide further insights into the molecular basis of virulence and adaptation of *S. agalactiae* in fish, which will be covered in the next section.


A.
STIR-CD-17(ST260)
SA20-06 (ST553)
CC552
Group 1
1.00
FSL S3-026 (ST67)
ATCC13813 (ST61)
CC67
515 (ST23)
CC23
Group 2
COH1 (ST17)
CC17
H36B (ST6)
A909 (ST7)
GD201008-001(ST7)
ZQ0910 (ST7)
CC7
0.001 substitution per site

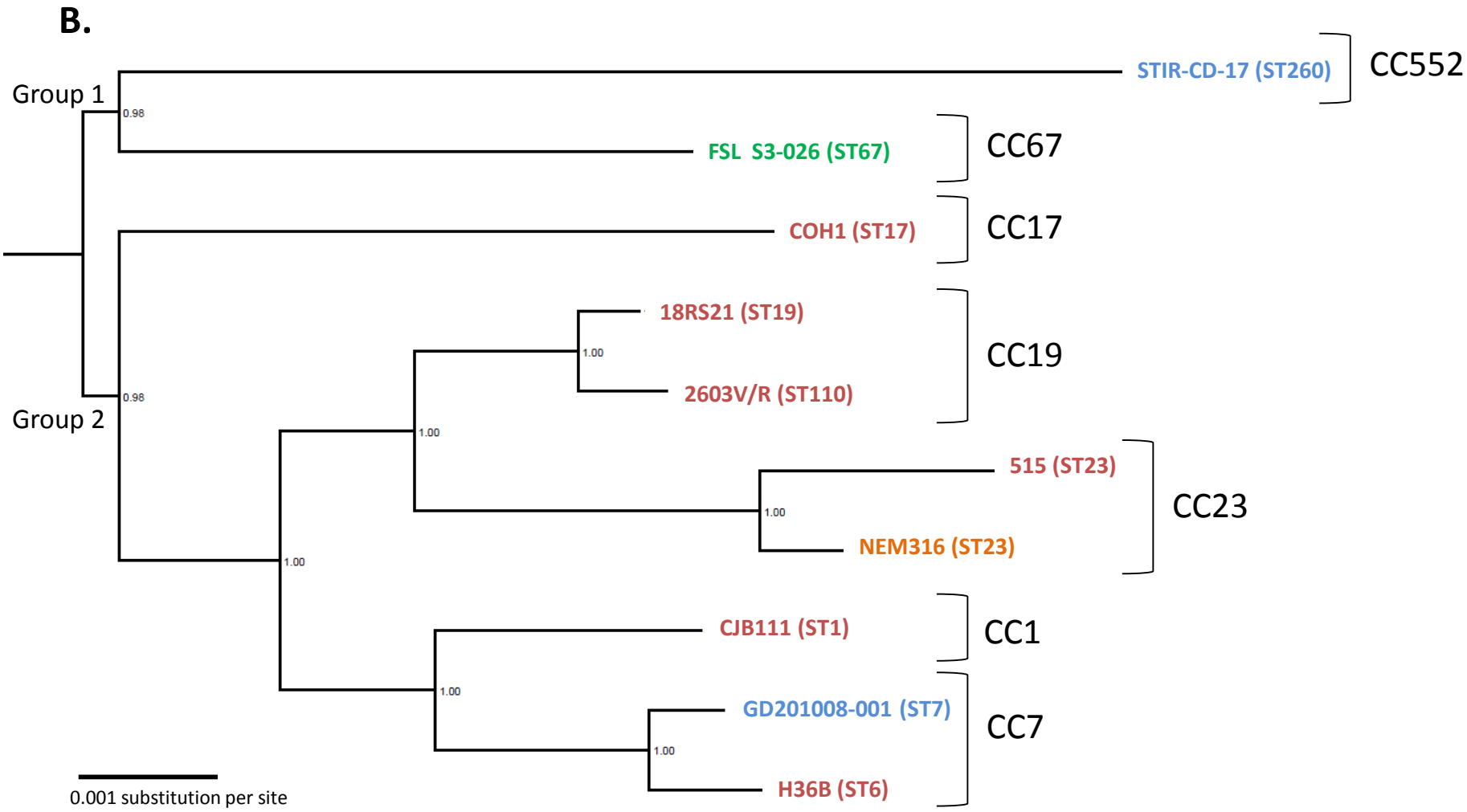


Figure 4.2 Phylogenetic evidence for host-adaptation of *S. agalactiae*. The bayesian phylogenetic trees A and B were constructed based on the core genome of 10 isolates of *S. agalactiae*. The host of origin for each isolate is indicated by the colour of the isolate name: blue, fish; green, bovine; red, human; orange, unknown. Sequence types (ST) and clonal complex (CC) of isolates included in the analysis are indicated. Posterior probabilities are shown at each node and the scale bar represents the substitutions per site.

4.4.4 Comparative genomic analysis of the accessory genome

A 3-step comparative genomic approach enabled the identification of putative strain-specific, fish-specific and fish-associated genes. Some of these genes may have been lost by other genomes that share a common ancestor with STIR-CD-17, or they may have been acquired through horizontal transfer of genetic material. These genes are present in small clusters, called loci in this study. A total of 8 loci were identified (**Table 4.4**). In the subsequent sections, details for each locus will be presented and discussed.

Table 4.4 List of candidate strain-specific, fish-specific and fish-associated genes as obtained by GeneRator analysis and normalized BLAST score ratio (BSR).

NCBI annotation			PSORBTb	Generator analysis	BlastP results based on NCBI search		
Locus tag	Bp	Putative function and COG			Species and strains	Product	BSR
LOCUS 1 (Contig23)							
M3M_05402	801	hypothetical protein	C	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0045	1
					Eremococcus coleocola ACS-139-V-Col8	oxidoreductase, NAD-binding domain protein	0.07
					Tribolium castaneum	hypothetical protein TcasGA2_TC008151	0.07
M3M_05407	591	hypothetical protein	C	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_004	1
					S. oralis SK304	hypothetical protein HMPREF1125_1048	0.17
					Campylobacter jejuni subsp. *jejuni BH-01-0142*	hypothetical protein CJBH_0152c	0.1
LOCUS 2 (Contig753)							
M3M_01252	222	hypothetical protein	C	-	***S. agalactiae SA20-06***	hypothetical protein gbs0229	1
					S. agalactiae 515/NEM316	Unknown	0.62
					S. agalactiae ATCC 13813	hypothetical protein HMPREF9171_2185	0.62
LOCUS 3 (Contig753)							
M3M_01147	195	aldose 1-epimerase, interruption-C (*gal*M); COG2017 Galactose mutarotase and related enzymes	U	**A909, H36B**	***S. agalactiae SA20-06***	aldose 1-epimerase	1
					S. agalactiae A909	hypothetical protein SAK_0542	0.99
					S. agalactiae H36B	aldose 1-epimerase	0.99
					S. agalactiae GD201008-001/ZQ0910	hypothetical protein	0.99
					Streptococcus suis ST3	galactose mutarotase-like protein	0.7
M3M_01152	564	aldose 1-epimerase, interruption-N (*gal*M); COG2017 Galactose mutarotase and related enzymes	C	**A909, H36B**	***S. agalactiae SA20-06***	aldose 1-epimerase	1
					S. agalactiae A909	hypothetical protein SAK_0539	0.97
					S. agalactiae H36B	aldose 1-epimerase, interruption-N	0.97
					S. agalactiae GD201008-001/ZQ0910	hypothetical protein	0.96
					Streptococcus suis ST3	galactose mutarotase-like protein	0.61
M3M_01157	996	UDP-glucose 4-epimerase (*gal*E1); COG1087 UDP-glucose 4-epimerase	U	**A909, H36B**	***S. agalactiae SA20-06***	UDP-glucose 4-epimerase	1
					S. agalactiae A909/H36B/GD201008-001/ZQ0910	UDP-glucose 4-epimerase	0.99
					S. gallolyticus UCN34	UDP-glucose 4-epimerase	0.83
M3M_01162	1482	galactose-1-phosphate uridylyltransferase (galT); COG4468 Galactose-1-phosphate uridyltransferase	C	**A909, H36B**	***S. agalactiae SA20-06***	galactose-1-phosphate uridylyltransferase	1
					S. agalactiae H36B	galactose-1-phosphate uridylyltransferase	0.99
					S. agalactiae GD201008-001	galactose-1-phosphate uridylyltransferase	0.99
					S. agalactiae A909/ZQ0910	galactose-1-phosphate uridylyltransferase	0.99
					S. sanguinis SK355	UTP-hexose-1-phosphate uridylyltransferase	0.69
M3M_01167	1173	galactokinase (*gal*K); COG0153 Galactokinase	C		***S. agalactiae H36B***	galactokinase	0.91
				H36B	***S. agalactiae SA20-06***	galactokinase	0.91
					S. gallolyticus subsp. gallolyticus TX20005	galactokinase	0.77
M3M_01172	2202	alpha-galactosidase (*gal*A); COG3345 Alpha-galactosidase	C	**A909, H36B**	***S. agalactiae SA20-06***	alpha-galactosidase	1
					S. agalactiae A909/GD201008-001/ZQ0910/H36B	alpha-galactosidase	0.99
					S. canis FSL Z3-227	alpha-galactosidase	0.69
M3M_01177	828	ABC transporter permease; COG0395 ABC-type sugar transport system, permease component	CM	**A909, H36B**	***S. agalactiae SA20-06***	ABC transporter permease	1
					S. agalactiae A909/GD201008-001/ZQ0910/H36B	ABC transporter permease	1
					S. canis FSL Z3-227	ABC transporter permease	0.87
M3M_01182	903	sugar ABC transporter permease; COG1175 ABC-type sugar transport systems, permease components	CM	**A909, H36B**	***S. agalactiae SA20-06***	binding-protein-dependent transport system	1
					S. agalactiae A909/GD201008-001/ZQ0910/H36B	sugar ABC transporter permease	0.93
					S. canis FSL Z3-227	ABC-type sugar transport permease component	0.82
					S. iniae 9117	ABC superfamily ATP binding cassette transporter	0.79
M3M_01187	1035	sugar ABC transporter sugar-binding protein; COG1653 ABC-type sugar transport system, periplasmic component	U	**A909, H36B**	***Streptococcus agalactiae SA20-06***	sugar ABC transporter substrate-binding protein	0.99
					S. agalactiae A909/GD201008-001/ZQ0910/H36B	sugar ABC transporter sugar-binding protein	0.74
					Streptococcus porcinus str. Jelinkova 176	ABC transporter, solute-binding protein	0.57
M3M_01192	831	AraC family transcriptional regulator; COG2207 AraC-type DNA-binding domain-containing proteins	C	**A909, H36B**	***S. agalactiae A909/GD201008-001/ZQ0910/H36B/SA20-06***	AraC family transcriptional regulator	1
					S. canis FSL Z3-227	transcriptional regulator	0.66
					S. suis 05ZYH33	transcriptional regulator	0.65
M3M_01197	285	phosphotransferase system, galactitol-specific IIB component; COG3414 Phosphotransferase system, galactitol-specific IIB component	CM	**A909, H36B**	***S. agalactiae SA20-06***	PTS system lactose/cellobiose specific transporter	1
					S. agalactiae A909/GD201008-001/ZQ0910/H36B	PTS system galactitol-specific transporter subunit IIB	0.99
					S. ictaluri 707-05	PTS system, Lactose/Cellobiose specific IIB subunit	0.83
					S. pseudoporcinus SPIN 20026	PTS system, galactitol-specific IIB component	0.82
M3M_01202	1332	PTS system, galactitol-specific IIC component; COG3775 Phosphotransferase system, galactitol-specific IIC component	CM	**A909, H36B**	***S. agalactiae A909/GD201008-001/ZQ0910/H36B***	PTS system galactitol-specific transporter subunit IIC	1
					S. agalactiae SA20-06	PTS system component	0.92
					Granulicatella adiacens ATCC 49175	PTS system, galactitol-specific IIC component	0.83
					S. iniae 9117	PTS family galactitol (gat) porter component IIC	0.81
M3M_01207	465	PTS system, galactitol-specific IIA component; COG1762 Phosphotransferase system mannitol/fructose-specific IIA domain	C	**A909, H36B**	***S. agalactiae SA20-06***	PTS system galactitol-specific transporter subunit IIA	0.99
					S. agalactiae A909/GD201008-001/ZQ0910/H36B	PTS system, galactitol-specific IIA component	0.98
					Granulicatella elegans ATCC 700633	PTS system IIA component	0.45
M3M_01212	831	rhamnulose-1-phosphate aldolase; COG0235 Ribulose-5-phosphate 4-epimerase and related epimerases and aldolases	C	**A909, H36B**	***S. agalactiae A909/GD201008-001/ZQ0910/H36B***	rhamnulose-1-phosphate aldolase	1
					S. agalactiae SA20-06	rhamnulose-1-phosphate aldolase	0.99
					S. anginosus subsp. whileyi CCUG 39159	putative rhamnulose-1-phosphate aldolase	0.7
M3M_01217	1338	PTS system, galactitol-specific IIC component; COG3775 Phosphotransferase system, galactitol-specific IIC component	CM	**A909, H36B**	***S. agalactiae SA20-06***	PTS system sugar-specific transporter permease	1
					S. agalactiae A909/GD201008-001/ZQ0910	PTS system galactitol-specific transporter subunit IIC	1
					S. anginosus subsp. whileyi CCUG 39159	PTS system sugar-specific permease protein	0.9
M3M_01222	279	PTS system galactitol-specific enzyme IIB component; COG3414 Phosphotransferase system, galactitol-specific IIB component	U	**A909, H36B**	***S. agalactiae SA20-06***	PTS system lactose/cellobiose specific transporter	1
					S. agalactiae A909/GD201008-001/ZQ0910/H36B	PTS system galactitol-specific transporter subunit IIB	0.98
					S. anginosus subsp. whileyi CCUG 39159	PTS system, lactose/cellobiose-specific IIB subunit	0.89
					Granulicatella elegans ATCC 700634	PTS system galactitol-specific transporter subunit IIB	0.86
					Streptococcus ictaluri 707-05	PTS system, Lactose/Cellobiose specific IIB subunit	0.84
M3M_01227	450	PTS system, galactitol-specific IIA component; COG1762 Phosphotransferase system mannitol/fructose-specific IIA domain	C	**A909, H36B**	***S. agalactiae A909/GD201008-001/ZQ0910/H36B/SA20-06***	PTS system galactitol-specific transporter subunit IIA	1
					S. porcinus str. Jelinkova 176	phosphoenolpyruvate-dependent sugar family porter	0.22
					S. anginosus subsp. whileyi CCUG 39159	phosphoenolpyruvate-dependent sugar family porter	0.22
M3M_01232	2043	PTS system IIA domain-containing protein; COG3711 Transcriptional antiterminator	C	**A909, H36B**	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0403	1
					S. agalactiae ZQ0910	PTS system IIA domain-containing protein	1
					S. agalactiae A909/GD201008-001	PTS system IIA domain-containing protein	1
					S. agalactiae H36B	MW0309, putative	0.99
					Coprobacillus sp. 29_1	hypothetical protein HMPREF9488_00517	0.25
LOCUS 4 (Contig753)							
M3M_01047	294	hypothetical protein; COG1343 Uncharacterized protein predicted to be involved in DNA repair	C	-	***S. agalactiae SA20-06***	CRISPR-associated endoribonuclease Cas2	1
					S. constellatus subsp. constellatus SK53	CRISPR-associated endoribonuclease Cas2	0.95
					S. mutans UA159	hypothetical protein SMU_1753c	0.94
		csd1 pseudogene					
M3M_01062	672	putative RecB family exonuclease; COG1468 RecB family exonuclease	C	-	***S. agalactiae SA20-06***	CRISPR-associated protein Cas4	1
					S. mutans LJ23	CRISPR-associated protein cas4	0.88
					S. mutans UA159	hypothetical protein	0.86
M3M_01068	850	hypothetical protein; COG3649 Uncharacterized protein predicted to be involved in DNA repair	C	-	***S. agalactiae SA20-06***	Csd2 family CRISPR-associated protein	1
					S. dysgalactiae subsp. equisimilis RE378	putative cytoplasmic protein	0.98
					S. dysgalactiae subsp. equisimilis SK1251	hypothetical protein HMPREF9963_1905	0.98
M3M_01097	729	hypothetical protein	C	**515**	***S. agalactiae SA20-06***	CRISPR-associated protein Cas5	1
					S. dysgalactiae subsp. equisimilis AC-2713	hypothetical protein SDSE_1670	0.96
					S. canis FSL Z3-227	hypothetical protein SCAZ3_08370	0.96

Table 4.4 (Continued)

NCBI annotation			PSORBTb	Generator analysis	BlastP results based on NCBI search		
Locus tag	Bp	Putative function and COG			Species and strains	Product	BSR
M3M_01102	2424	ATP-dependent RNA helicase; COG1203 Predicted helicases	C	515, COH1, FSL-S3-026	***S. agalactiae SA20-06***	CRISPR-associated helicase Cas3	1
					S. dysgalactiae subsp. equisimilis SK1250	CRISPR-associated helicase Cas3	0.96
					S. dysgalactiae subsp. equisimilis AC-2713	Pre-mRNA-processing ATP-dependent RNA helicase	0.95
M3M_01107	222	Fic protein family; COG2184 Protein involved in cell division	U	515	***S. agalactiae 515***	Fic protein family family	0.97
					S. suis 05ZYH33	hypothetical protein SSU05_0462	0.61
					S. suis ST1	hypothetical protein SSUST1_0463	0.61
LOCUS 5 Contig 751)							
misc_feature		CHAP domain protein					
misc_feature		resolvase family site-specific recombinase					
misc_feature		resolvase family site-specific recombinase					
M3M_00445	204	bacteriocin	E	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0545	1
					S. equi subsp. zooepidemicus MGCS10565	bacteriocin BlpN-like	0.58
					S. equi subsp. equi	Streptococcus equi subsp. equi 404	0.57
M3M_00450	231	hypothetical protein	U	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0544	1
					S. pneumoniae SP6-BS73	bacteriocin BlpM	0.69
					S. equi subsp. equi 4047	bacteriocin	0.65
misc_feature		abortive infection protein AbiGII					
M3M_00465	582	hypothetical protein; COG1672 Predicted ATPase (AAA+ superfamily)	C	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0542	1
					S. macedonicus ACA-DC 198	Abortive infection protein AbiGI	0.69
					S. dysgalactiae subsp. dysgalactiae ATCC 27957	hypothetical protein SDD27957_04365	0.69
M3M_00470	138	Tn5252 Orf28; COG3942 Surface antigen	U	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0541	0.97
					S. intermedius F0395	hypothetical protein HMPREF9682_00655	0.87
					S. suis D12	hypothetical protein SSUD12_0897	0.78
M3M_00475	327	hypothetical protein; COG0270 Site-specific DNA methylase	C	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0540	1
					S. agalactiae NEM316	hypothetical protein	0.37
					S. pneumoniae NorthCarolina6A-23	modification methylase HpaII	0.36
LOCUS 6 (Contig371)							
M3M_04250	357	integrase; COG0582 Integrase	U	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0928	0.99
					S. pneumoniae 70585	Integrase	0.63
					S. pneumoniae 2061617	phage integrase family protein	0.63
M3M_04255	189	hypothetical protein	U	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0927	1
					S. pneumoniae GA47461	hypothetical protein SPAR97_1602	0.68
					S. pneumoniae GA17484	hypothetical protein SPAR47_088	0.68
M3M_04260	126	Cro/CI family transcriptional regulator	U	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0926	1
					S. pneumoniae CCRI 1974	hypothetical protein SpneC1_02124	0.68
					S. pneumoniae CCRI 1974M2	hypothetical protein SpneC19_10413	0.68
M3M_04265	339	hypothetical protein	U	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0925	0.69
					Oenococcus oeni AWRIB429	hypothetical protein AWRIB429_1949	0.16
					Oenococcus oeni AWRIB548	phage terminase large subunit	0.16
M3M_04270	792	hypothetical protein	U	-	*Bacillus amyloliquefaciens DC-12*	hypothetical protein BamyaD_16251	0.14
					Bacillus sp. 5B6	hypothetical protein MY7_0533	0.13
					Bacillus amyloliquefaciens AS43.3	hypothetical protein B938_03325	0.13
M3M_04275	639	hypothetical protein; COG0477 Permeases of the major facilitator superfamily	CM	-	*Clostridiales bacterium OBRC5-5*	hypothetical protein HMPREF1135_01905	0.4
					Lachnospiraceae oral taxon 107 str. F0167	hypothetical protein HMPREF0491_01439	0.37
					Staphylococcus hominis SK119	multidrug resistance protein 1	0.11
M3M_04280	369	beta-hydroxyacyl-(acyl-carrier-protein) dehydratase FabA/FabZ; COG0764 3-hydroxymyristoyl/3-hydroxydecanoyl-	U	-	*Geobacillus thermoleovorans CCB_US3_UF5*	putative thioester dehydrase	0.31
					Coraliomargarita akajimensis DSM 45221	beta-hydroxyacyl- dehydratase FabA/FabZ	0.29
					Pirellula staleyi DSM 6068	beta-hydroxyacyl- dehydratase FabA/FabZ	0.29
M3M_04285	540	PadR family transcriptional regulator; COG1695 Predicted transcriptional regulators	C	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_0923	1
					Lachnospiraceae oral taxon 107 str. F0167	hypothetical protein HMPREF0491_01453	0.26
					Clostridiales bacterium OBRC5-5	hypothetical protein HMPREF1135_01904	0.25
LOCUS 7 (Contig381)							
M3M_01921	618	DJ-1/PfpI family protein; COG0693 Putative intracellular protease/amidase	U	-	***S. agalactiae SA20-06***	hypothetical protein SaSA20_1488	1
					S. criceti HS-6	hypothetical protein STRCR_0670	0.76
					S. sanguinis SK1087	ThiJ/PfpI family intracellular protease	0.61
M3M_01926	531	hypothetical protein; COG3797 Uncharacterized protein conserved in bacteria	C	-	***S. agalactiae SA20-06***	Streptococcus agalactiae SA20-06	1
					S. suis ST1	hypothetical protein SSUST1_1897	0.73
					S. suis R61	hypothetical protein SSUR61_0033	0.71
LOCUS 8 Contig 751)							
M3M_00390	270	hypothetical protein	U	-	*Ashbya gossypii*	AAR142Cp	0.19
					delta proteobacterium	Response regulator receiver	0.19
					Photorhabdus asymbiotica	Gramicidin S synthetase 2	0.19
M3M_00395	897	hypothetical protein	C	-	*Krokinobacter sp.*	hypothetical protein	0.06
					-	-	
					-	-	
M3M_00400	1017	hypothetical protein; COG0457 FOG: TPR repeat	C	-	*Gloeobacter violaceus*	hypothetical protein	0.06
					Helicobacter pylori F32	hypothetical protein HPF32_0454	0.05
					-	-	
M3M_00405	1263	serine hydroxymethyltransferase; COG0112 Glycine/serine hydroxymethyltransferase	C	-	*Fusobacterium sp.*	conserved hypothetical protein	0.32
					Coprococcus eutactus ATCC 27759	hypothetical protein COPEUT_02118	0.29
					Erysipelotrichaceae bacterium 3_1_53	hypothetical protein HMPREF0983_03234	0.25
M3M_00410	318	integrase; COG4974 Site-specific recombinase XerD	U	-	*S. porcinus str. Jelinkova 176*	phage integrase, N-terminal SAM domain protein	0.73
					S. anginosus subsp. whileyi CCUG 39159	site-specific recombinase, phage integrase family	0.71
					S. mitis bv. 2 str. SK95	phage integrase, N-terminal SAM domain protein	0.66

The 8 clusters of selected genes are referred to here as loci. For each gene are provided: the locus tag within the STIR-CD-17 genome, the size of the gene (bp), the putative function and Cluster of Orthologous Groups classification of the corresponding protein as provided by the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP), and the predicted sub-cellular localization of the corresponding protein using the PSORTb program version 3.0.2 (http://psort.org; Yu et al., 2010) (C, Cytoplasm; CM, Cytoplasmic Membrane; E, Extracellular; U, Unknown). The presence of predicted protein sequences with a reciprocal best hit with a BLAST score > 80 in other *S. agalactiae* genomes based on GeneRator analysis is shown in the table with the name of the strain. The first BLASTP results for each protein are provided as well as the calculated BLAST score ratio (BSR). Pseudogenes are highlighted in blue, and genes later selected for PCR screening are highlighted in orange.

4.4.4.1 Accessory genome – Locus 1

Locus 1 comprises two protein-encoding genes that were present in the fish-adapted subgroup of *S. agalactiae* but not in other genomes included in the analysis. Their corresponding proteins did not contain any known domain from which a putative function could be deduced and they have been predicted to have a cytoplasmic localization. The 2 genes lay between genes that are well conserved among other streptococcal genomes, namely the genes *purK* (ATPase subunit) and *purB* (adenylosuccinate lyase). Conserved flanking regions, as identified here, can act as substrates for homologous recombination between strains, representing a so-called "minimal mobile element" (MME; Saunders and Snyder, 2002). Accordingly, the region delimited by *purK* and *purB* is a well-recognised MME containing variable inserts within the genome sequence of other pathogenic streptococci, *e.g. S. pneumoniae* and *S. pyogenes*, as demonstrated by Herbert et al. (2005). Selected genomes were compared using ACT in our study, which revealed that the region delimited by *purK* and *purB* is occupied by distinct genes between isolates of different lineages, *e.g.* STIR-CD-17 (CC552), GD201008-001 (CC7) and NEM316 (CC23), whereas the region is identical between isolates belonging to the same CCs. For example, the same putative membrane protein was identified among all *S. agalactiae* belonging to ST7 with no distinction between isolates from human or from fish origin (**Figure 4.3**).

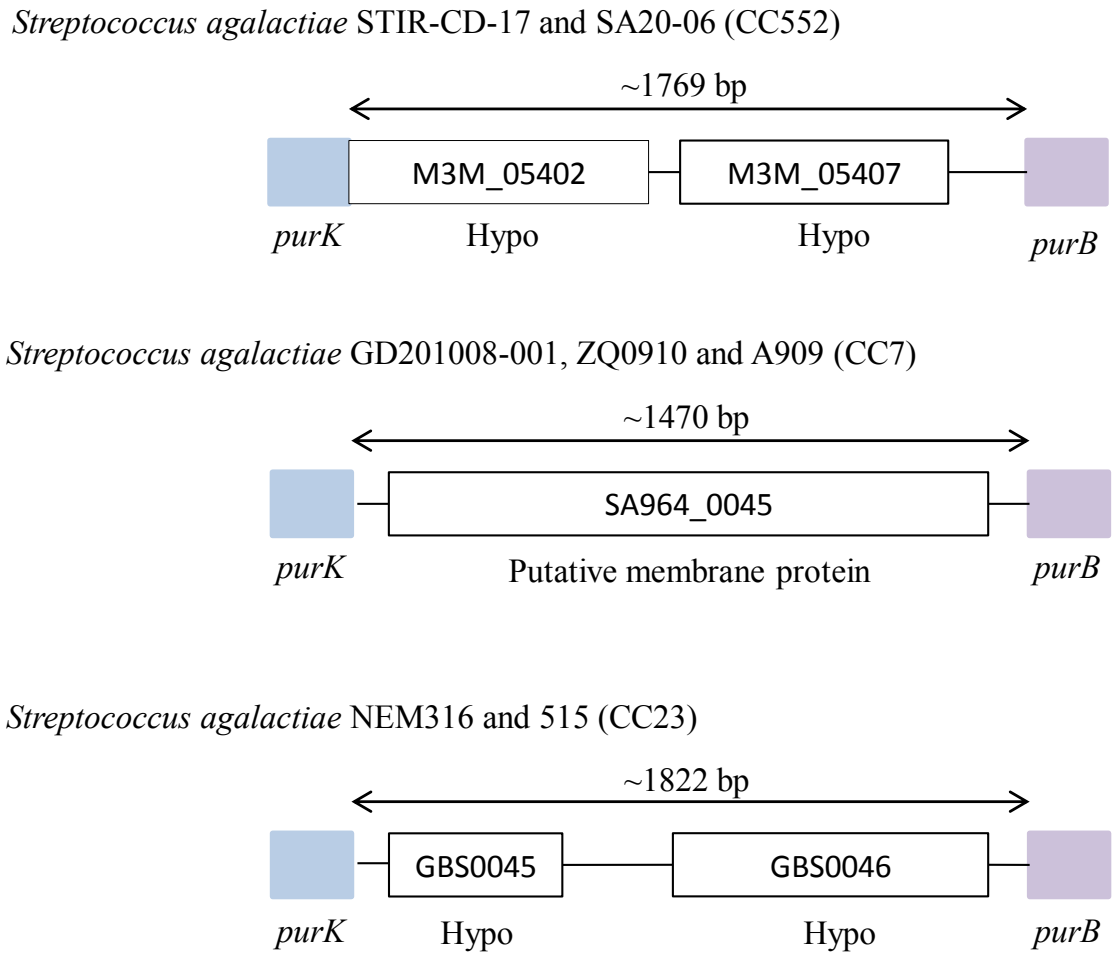


Figure 4.3 Schematic representation of a minimal mobile element (MME) in *S. agalactiae* of distinct lineages. Different intergenic regions are depicted between *purK* (pale blue) and *purB* (pale purple) for *S. agalactiae* isolates of CC552, CC7 and CC23. Hypo, Hypothetical protein.

4.4.4.2 Accessory genome – Locus 2

Locus 2 is located on the putative pathogenicity island (PAI) IV that has been described in human *S. agalactiae* (Glaser et al., 2002; Herbert et al., 2005). It contains one single gene encoding for a protein uniquely found present in the fish-adapted subgroup of *S. agalactiae*. This putative protein is of unknown function and predicted to be localised in the cytoplasm. In this study, ACT comparison showed that this locus is absent in other *S. agalactiae* genomes from human or bovine origin, and the corresponding region occupied by a cluster of genes that include either the virulence genes *rib* or *bca* (**Figure 4.4**). The presence of *rib* and *bca* is mutually exclusive as only one of those is found in any isolate, which is in agreement with 3-set genotyping results (Kong et al., 2002; Chapter 2).

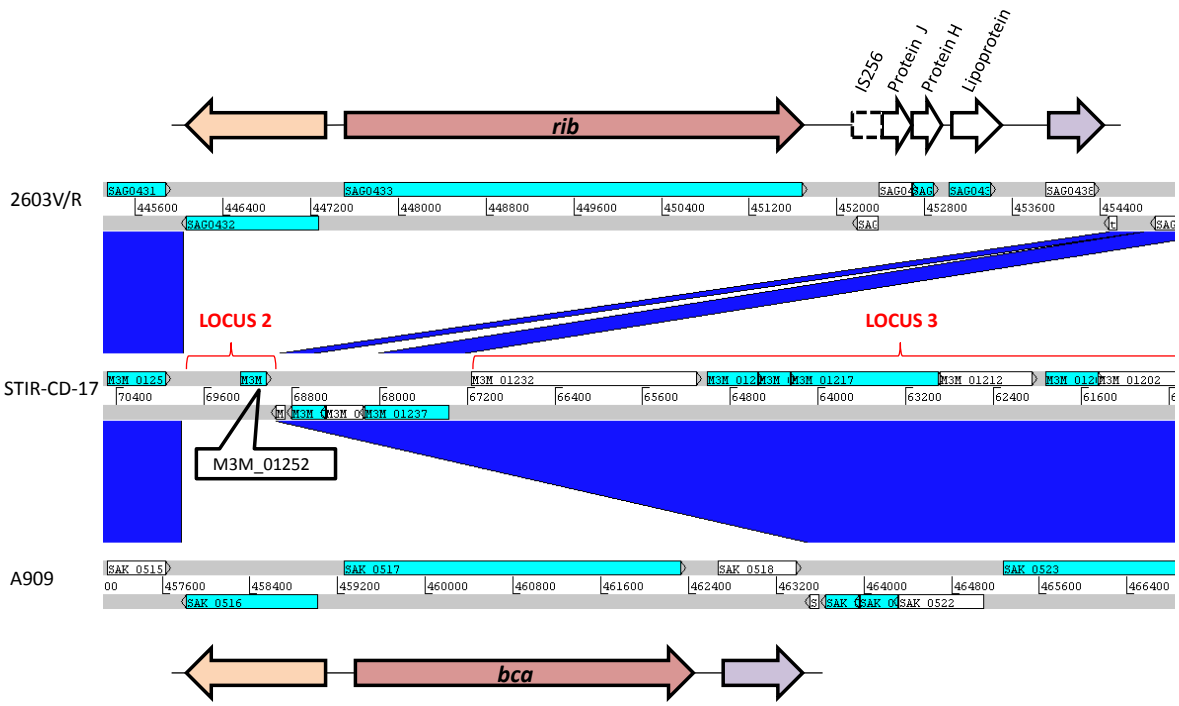


Figure 4.4 Analysis of the genomic region comprising one of the surface protein genes *rib* or *bca* in *S. agalactiae* 2603V/R (top sequence) and A909 (bottom sequence) using ACT. The strain 2603V/R contains the *rib* gene whereas A909 contains the *bca* gene. These genes are found to belong to a genomic region delimited by an AraC family transcriptional regulator (yellow arrow) and a site specific recombinase (purple arrow). By ACT comparison, that genomic region is found to be absent from STIR-CD-17 (central sequence) and replaced by ~1Kbase sequence containing a putative strain-specific gene (M3M_01252). The blue bars separating each genome represent similarity matches identified by reciprocal NBLAST analysis, with a score cutoff of 100.

4.4.4.3 Accessory genome – Locus 3

Locus 3, like locus 2, is located on the PAI IV from human *S. agalactiae* (Glaser et al., 2002; Herbert et al., 2005), lying 4 genes downstream of locus 2 (**Figure 4.4**) and 5 genes upstream of locus 4. Locus 3 comprises 18 genes. The corresponding proteins are found in all *S. agalactiae* from fish, including those belonging to CC7, as well as the human *S. agalactiae* belonging to CC7. These genes are therefore considered to be "fish-associated" in *S. agalactiae*, even though some of the corresponding proteins are also found to be well conserved in other *Streptococcus* spp, *e.g. S. ictaluri* and *S. canis*, or in other bacterial species, *e.g. Granulicatella elegans*. All those genes encode for proteins involved in carbohydrate transport and metabolism with the exception of 1 transcriptional regulator. Among these genes, one codes for an α-galactosidase (*galA*), an enzyme that catalyzes the hydrolysis of galactose-containing oligosaccharides. It also contains genes for all the

enzymes of the Leloir pathway (*galK*, *galE*, *galM* and *galT*) involved in the transport and degradation of galactose. These genes have been well characterised in lactic acid bacteria (Grossiord et al., 1998). Galactose is present in dairy products but also in fish tissues like the brain, where it forms part of glycolipids and glycoproteins (Tocher, 2003). The presence of these genes in meningo-encephalitis causing bacteria may therefore provide some metabolic advantages. In *S. thermophilus*, it is strongly suggested that the primary role for Leloir pathway enzymes is to produce precursor sugars for assembly of exopolysaccharides (EPS; Levander and Rådström, 2001). EPS are secreted externally, and differ from the capsular polysaccharides (CPS) that are tightly associated with the cell surface (Levander and Rådström, 2001). The production of EPS in bacteria leads to a loose "fluffy" pellet phenotype following centrifugation (Forde and Fitzgerald, 2002); interestingly, a similar phenotypic observation was also made with STIR-CD-17 (Chapter 3, section 3.3.1.3). Yet, to our knowledge no EPS has been described in *S. agalactiae*. However, despite being responsible for distinct end products, some of the *eps* genes responsible for EPS production in *S. thermophilus* are known to be organised in the same manner and to share very high homology with *cps* genes responsible for capsule polysaccharide production in *S. agalactiae* (Stingele et al., 1996), suggesting a common origin of these genes. Consequently, whether some strains of *S. agalactiae* have the ability to produce EPS and if so the mechanism behind it remains to be investigated further. The *S. agalactiae* capsule is composed of numerous polysaccharides that include glucose, galactose and rhamnose (Cieslewicz et al., 2005); the enzymes encoded by Locus 8 may therefore also play a role in the production of precursors involved in capsule formation.

4.4.4.4 Accessory genome – Locus 4

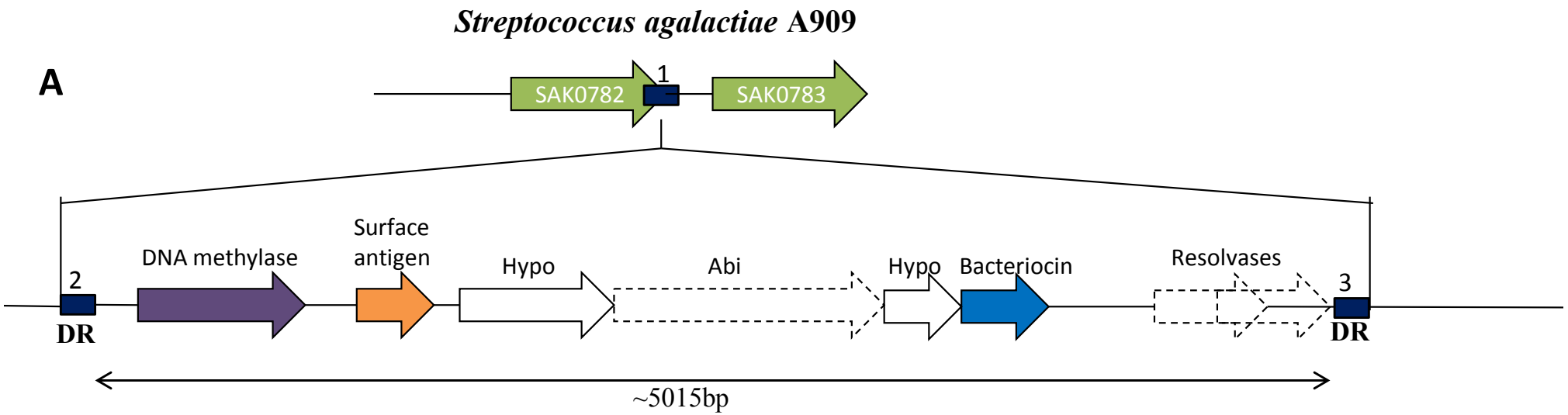
The genetic composition and organisation of locus 4 corresponds to a CRISPR-cas module (clustered regularly interspaced short palindromic repeats-CRISPR-associated proteins) of subtype IC (Makarova et al., 2011). CRISPR modules are carried within the genomes of many bacteria and constitute an adaptative immunity system that acts against mobile genetic

elements (Horvath et al., 2010). Based on this comparative analysis, this locus is unique to the fish-adapted isolates but remnants of it are present in 3 other *S. agalactiae* genomes (515, COH1, FST S3-026) and in other *Streptococcus* spp. (**Table 4.4**). This locus, with the very same sequence as in STIR-CD-17, was also recently reported from an ST260 isolate recovered from a frog (Lopez-Sanchez et al., 2012). Two of the genes in this locus are mutated (pseudogenes) and the system is therefore probably inactive in our strain, as also suggested for the frog isolate (Lopez-Sanchez et al., 2012).

4.4.4.5 Accessory genome – Locus 5

Locus 5 is composed of 5 genes and 4 pseudogenes that were only detected in the fish-adapted subgroup of *S. agalactiae*. This locus is located on the PAI VI reported from human *S. agalactiae* (Herbert et al., 2005) and contains three protein-encoding genes and one pseudogene with defence mechanism functions. One of the genes encodes a hypothetical protein that has a conserved cytosine-DNA methylase domain (PF00145) and might therefore play a role in protection against bacteriophages (Kumar et al., 1994). Indeed, it has been shown that restriction endonucleases discriminate between endogenous and foreign DNA by their methylation patterns and therefore lead to the cleavage of DNA not protected by methylation (Briggs et al., 1994). Two other genes that are adjacent to each other encode proteins with domains corresponding to those found in bacteriocins. Bacteriocins are proteinaceous toxins produced by bacteria to inhibit the growth of similar or closely-related strains. Typically, production of bacteriocin is co-ordinated by an operon carrying up to 6 genes including those encoding a two-component regulatory system, a small peptide pheromone, and a dedicated ABC transporter, as well as an immunity protein (Dawid et al., 2007). Such an operon was not found in our strain and it is therefore unlikely that the proteins would be expressed and localised extracellularly in STIR-CD-17. Finally, the identified pseudogene would, if expressed, encode an abortive infection protein AbiGII which is considered to confer protection against bacteriophages by blocking their multiplication (Chopin et al., 2005).

Locus 5 also contains a gene that encodes a protein suspected to be a surface antigen (COG classification) and containing a CHAP domain (pfam05257). The CHAP domain is associated with amidase functions, and is usually found in proteins involved in bacterial cell wall metabolism of bacteria (Bateman and Rawlings, 2003). Sub-cellular localisation prediction further supports that this protein, if produced, is extracellularly located. The exact role of this protein in piscine *S. agalactiae* deserves further exploration as it is possibly involved in interaction with the host or in the formation of capsule. Finally, locus 5 also contains 1 gene coding for a hypothetical protein containing a domain of unknown function (InteProScann; DUF4095) and 2 pseudogenes encoding resolvases. In this study, ACT comparison analyses revealed that the locus is delimited by direct repeats of 159bp and that the corresponding region for Locus 5 in other genomes is occupied by the same, single 159 bp sequence (**Figure 4.5**). This observation suggests that it might have been a mobile element (Sadowski, 1986) that has lost its mobility through inactivation (by mutation) of the resolvases.



B

```
          *        20        *        40        *        60        *        80        *       100       *       120       *       140       *       160
1 GTAGGGTTACTACAACGAAACAAAGAGTAAAAATCCGTTATTTTAAGCATTCTTCAAGCATTTTGTCTTTGTTGAAGAAGGTGA-TTTGACACCAAAAAGGTATTAAAAAACATATTGACGTGACCGTTTGTTTTGAAGTGGCTTGCGTAGACAAAAAA
2 GTAGCGTTGCTACAACGAAGCAAAGGGTAAAAATCCTTTATTTTAAGCACTTTTTCAAGCATTTTGTCTTTATTGAAAAGAGTGATTTTAACATAAAAAAGGTATTAAAAAACATATTGACGTGACCGTTTGTTTTGAAGTGGCTTGCGTAGACAAAAAA
3 GTAGGGTTACTACAACGAAACAAAGAG-AAAAATCCTGCATTTTAAGTAGTTTTACAAGCATTTTGTCTTTGTTGAAGAAGGTGAATTTGACACCAAAAAGGTATCAAAAAACATATTGAAGTGACCGTTTGTTTAGATGTTGATTGCGTAGACAAAAAG
```

Figure 4.5 (A) Genetic structure of the Locus 5 which carries genes (full arrows) and pseudogenes (dashed arrows). Locus 5 integrated into and disrupted a conserved gene encoding a putative outer-membrane protein. Direct repeats (DR) flank the extremities of the locus and a single example of the corresponding region is also found in other genomes not containing Locus 2. Hypo, Hypothetical protein. **(B) Alignment of the DR and equivalent region in A909, with annotation numbering referring to (A)**. For the alignment, sequences were extracted from the ACT comparison file, aligned using the MAFFT multiple sequence alignment web service implemented in JALVIEW 2.4 (Clamp *et al*., 2004). Shading levels are set using Genedoc software with the grey colour highlighting single nucleotide differences**.**

4.4.4.6 Accessory genome – Locus 6

Locus 6 is composed of 8 protein encoding genes, from which 3 are unique to STIR-CD-17 while the remaining 5 are unique to the fish-adapted subgroup of *S. agalactiae* (homologues in SA20-06 only). ACT comparative analysis revealed that a 6-phospho-beta-glucosidase coding gene present in other genomes is disrupted in STIR-CD-17 by the insertion of Locus 6, suggesting that genes from that locus are part of a mobile genetic element (**Figure 4.6**). Among the genes present are a putative integrase, two transcriptional regulators and three hypothetical proteins with no domain of known function. Also found are one beta-hydroxyacyl dehydratase involved in fatty acid biosynthesis (M3M_04280; PF07977) and one major facilitator family transporter (M3M_04275; PF07690). The latter protein, predicted to be localised in the cytoplasmic membrane, is known to transport small solutes in response to chemiosmotic ion gradient (Pao et al., 1998).

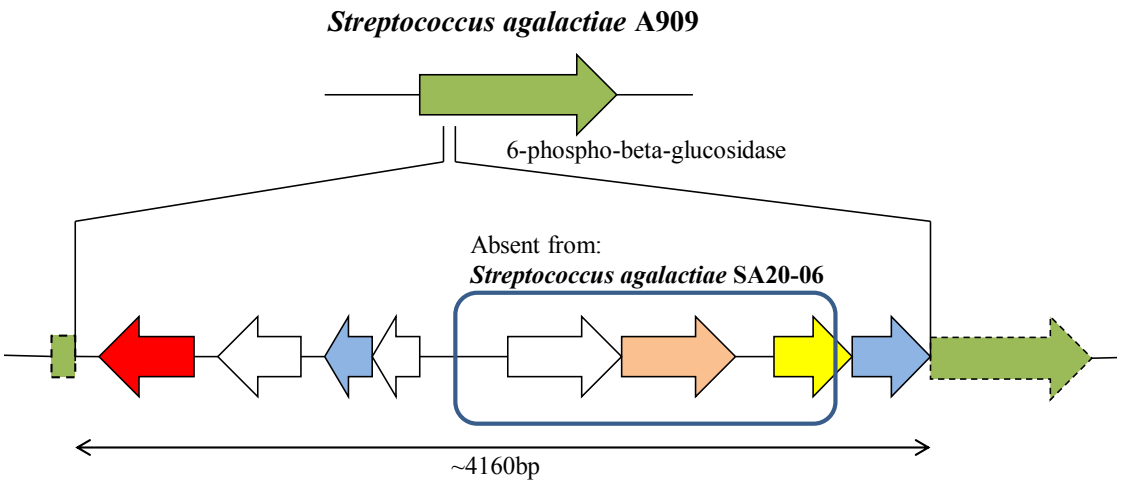


Figure 4.6 Genetic structure of Locus 6, which carries genes encoding an integrase (red), a beta-hydroxyacyl dehydratase (yellow), a putative permease (orange), hypothetical proteins (white), and transcriptional regulators (blue). Locus 5 integrated into and disrupted the gene encoding a 6-phospho-beta-glucosidase.

4.4.4.7 Accessory genome – Locus 7

Locus 7 comprises two genes that thus far have only been found in the fish-adapted subgroup of *S. agalactiae* (homologues identified within SA20-06 only). One of the genes encodes a hypothetical protein that contains a conserved domain of unknown function

(PF08002) and is predicted to be cytoplasmic. The other gene encodes a DJ-1/PfpI family protein (COG0693/PF01965), which includes proteins with intracellular protease function or transcriptional regulators (Halio et al., 1996; Ohnishi et al., 2005). These genes are inserted between 2 genes that are well conserved among streptococci: *ksgA* (16S ribosomal RNA methyltransferase KsgA/Dim1 family protein) and *rnmV* (primase-like protein). In other *S. agalactiae*, a distinct insert between these flanking genes has been identified by ACT comparisons in this study, hence fulfilling the criteria for an MME as described in section 4.4.4.1.1. Comparative analysis using ACT was performed, as for locus 1, using isolates belonging to CC552, CC7 and CC23. Unlike Locus 1, less diversity was observed as the isolates belonging to CC7 and CC23 all contained a homologous region between *rnmV* and *ksgA* (**Figure 4.7**).

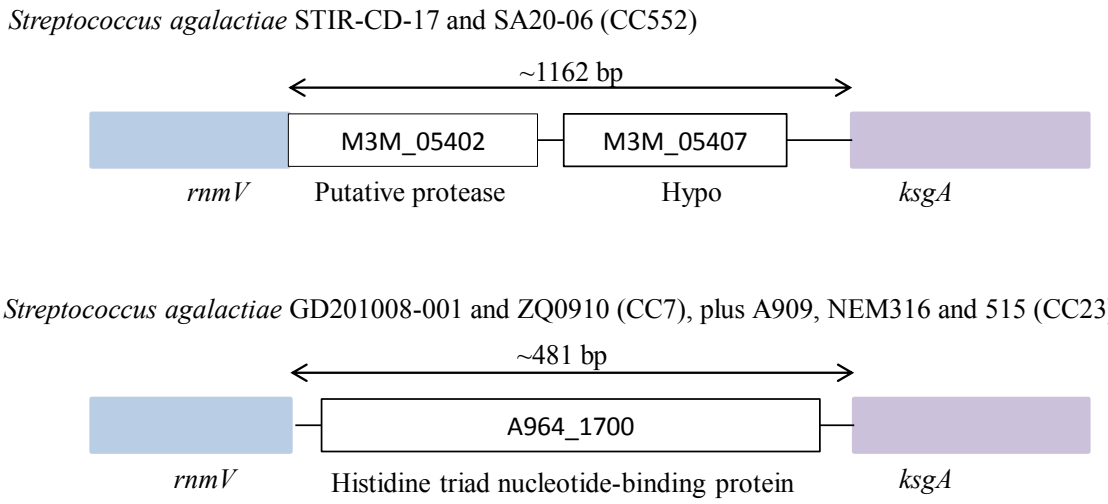


Figure 4.7 Schematic representation of a minimal mobile element (MME) in *S. agalactiae* of distinct lineages. Different intergenic regions are depicted between *rnmV* (pale blue) and *ksgA* (pale purple) between *S. agalactiae* isolates of CC552 and isolates of CC7/CC23. Hypo, Hypothetical protein**.**

4.4.4.8 Accessory genome – Locus 8

Locus 8 is part of a putative PAI (PAI VI; Herbert et al., 2005) that has been described in *S. agalactiae* NEM316 and 2603V/R. This PAI also contains the *cyl* operon, located 4 genes downstream from Locus 8. In our dataset, Locus 8 is however unique to STIR-CD-17 and contains genes encoding 3 hypothetical proteins with no domains of known function, 1

serine hydroxymethyltransferase (SHMT) and 1 integrase. The SHMT is an enzyme that catalyzes the reversible cleavage of serine to form glycine and monocarbonic groups, essential in several biosynthetic pathways. In a recent study, it was shown that the SHMT of halotolerant bacteria is up-regulated under conditions of high salinity, resulting in an increased salinity tolerance due to an accumulation of glycine betaine within the cell (Waditee-Sirisattha et al., 2012). It is unknown whether this enzyme plays a similar role in our strain, but such a mechanism may be important considering that *S. agalactiae* can infect fresh water and salt water host species (Bowater et al., 2012). The presence of an integrase encoding gene suggests that this locus may be a mobile genetic element, a hypothesis further supported by the ACT comparison (**Figure 4.8**). Indeed, the ACT comparison reveals that the locus was integrated into and disrupted a gene conserved among other *S. agalactiae* (as confirmed by BLASTN search).

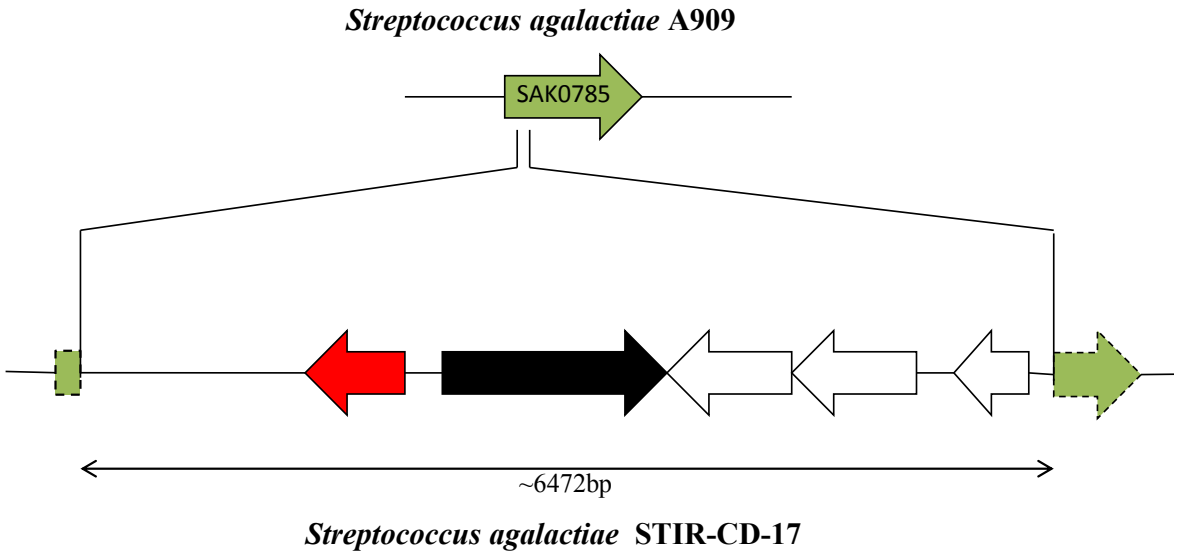


Figure 4.8 Genetic structure of Locus 8 which carries genes encoding an integrase (red), a serine hydroxymethyltransferase (black) and three hypothetical proteins (white). Locus 8 was integrated into and disrupted a conserved gene encoding a hypothetical protein.

4.4.5 Distribution of selected genes across *S. agalactiae* from different host species

The 3-step comparative analysis (GeneRator/BSR/ACT) of a limited number of available genomes enabled the identification of candidate strain-specific genes (SSG; unique to STIR-

CD-17), fish-specific genes (FSG; unique to isolates belonging to the fish adapted CC552; STIR-CD-17 and SA20-06) and fish-associated genes (FAG; unique to isolates belonging to CC7 and CC552; A909, H36B, ZQ0910, GD201008-001, STIR-CD-17 and SA20-06). As a proof of concept, a collection of bovine, sea mammal and piscine *S. agalactiae* isolates was screened by PCR for identification of presence/absence of 1 SSG (M3M_04280), 1 FSG (M3M_01062) and 3 FAGs (M3M_01167, M3M_01172 and M3M_01182), belonging to the previously described Loci 6, 4 and 3 respectively (**Table 4.4.**). These results, combined with results obtained from *in silico* analyses (**Table 4.4.**) of genomes of bovine, human, fish and unknown origins are provided in **Table 4.5**.

Six different profiles of presence/absence were obtained. Piscine isolates of ST260 were positive for the 5 genes (profile 6), whereas the genes were completely absent from 100% of the seal isolates, 95% of the bovine isolates and 72% of the human isolates (profile 1). Piscine isolates of ST261 all had profile 5, which differs from profile 6 in the absence of M3M_01062, whereas isolates from CC7 showed a range of profiles. All piscine isolates were positive for a minimum of 3 genes (profiles 2, 3, 4 and 5). There was therefore a strong correlation between the carriage of the genes and the host of origin being fish, with however some exceptions.

The 3 FAGs were expected to be present in isolates of CC552 and CC7. Accordingly, all fish isolates of CC552 and most isolates of CC7 were positive for these genes, including 1 dolphin isolate and 2 human isolates. However, 2 fish isolates of CC7 (STIR-CD-23 and STIR-CD-24) only had 2 of the 3 FAGs, whereas 1 bovine isolate of CC1 (MRI Z1-198) had all 3 FAGs.

The FSG was expected to be present in all fish isolates of CC552, however such association was not fully supported by the PCRs as it was not found within the fish isolates of ST261 (STIR-CD-29 to 32). Moreover the FSG was found within fish isolates belonging to human-

associated CCs (CC283 and CC7) and the dolphin isolate (profile 4), showing that they are FAG rather than FSG.

Finally, the SSG, which was expected to be present within fish isolates of ST260 only, was found within other fish isolates, including isolates of ST7 and ST261, and should therefore also be classified as FAG.

In conclusion, a strong correlation was found between the carriage of these genes and the host of origin being fish. However the characterisation of SSG, FSG and FAG was not fully supported by the PCRs. All of these genes will therefore be considered "putatively fish-associated" in the remainder of this chapter.

Table 4.5 (next page) Distribution of putative strain-specific, fish-specific and fish-associated genes among a population of *S. agalactiae* of bovine (green), human (red), sea mammal (black), fish (blue) and unknown (orange) origin. The 6 profiles obtained are the result of PCR screening (*in vitro* analysis; IV) or bioinformatics analysis based on available genomes (*in silico* analysis; IS). *In silico* analysis is based on complete genomes (isolate name in bold) as well as incomplete genomes for which the absence of a gene cannot be guaranteed. Strain name, origin, clonal complex (CC) and sequence type (ST) are provided. The multiple piscine isolates originating from a single country in Kuwait, Honduras, or Colombia are epidemiologically related.

Isolate ID	Host	Country	CC	ST	Source of data	M3M_04280	M3M_01062	M3M_01167	M3M_01172	M3M_01182	PROFILE
NEM316	NA	NA	23	23	IS	-	-	-	-	-	
515	Human	NA	23	23	IS	-	-	-	-	-	
2603V/R	Human	NA	19	110	IS	-	-	-	-	-	
18RS21	Human	USA	19	19	IS	-	-	-	-	-	
COH1	Human	NA	17	17	IS	-	-	-	-	-	
CJB111	Human	NA	1	1	IS	-	-	-	-	-	
MRI Z1-654	Bovine	Denmark	1	1	IV	-	-	-	-	-	
MRI Z1-658	Bovine	Denmark	1	1	IV	-	-	-	-	-	
MRI Z1-660	Bovine	Denmark	1	1	IV	-	-	-	-	-	
MRI Z1-651	Bovine	Denmark	1	458	IV	-	-	-	-	-	
FSL S3-026	Bovine	USA	67	67	IS	-	-	-	-	-	
ATCC 13813	Bovine	NA	67	61	IS	-	-	-	-	-	
MRI Z1-618	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-620	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-621	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-624	Bovine	Denmark	19	19	IV	-	-	-	-	-	1
MRI Z1-625	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-629	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-630	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-633	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-635	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-638	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-639	Bovine	Denmark	19	19	IV	-	-	-	-	-	
MRI Z1-619	Bovine	Denmark	19	44	IV	-	-	-	-	-	
MRI Z1-656	Bovine	Denmark	23	199	IV	-	-	-	-	-	
MRI Z1-643	Bovine	Denmark	23	23	IV	-	-	-	-	-	
MRI Z1-199	**Phocine**	UK	23	23	IV	-	-	-	-	-	
MRI Z1-200	**Phocine**	UK	23	23	IV	-	-	-	-	-	
MRI Z1-201	**Phocine**	UK	23	23	IV	-	-	-	-	-	
MRI Z1-202	**Phocine**	UK	23	23	IV	-	-	-	-	-	
MRI Z1-208	**Phocine**	UK	23	23	IV	-	-	-	-	-	
STIR-CD-23	Piscine	Thailand	7	7	IV	-	+	+	-	+	2
STIR-CD-24	Piscine	Thailand	7	7	IV	-	+	+	-	+	
MRI Z1-684	Bovine	Denmark	1	1	IV	-	-	+	+	+	
H36B	Human	USA	7	6	IS	-	-	+	+	+	
A909	Human	NA	7	7	IS	-	-	+	+	+	
ZQ0910	Piscine	China	7	7	IS	-	-	+	+	+	3
GD201008-001	Piscine	China	7	7	IS	-	-	+	+	+	
STIR-CD-26	Piscine	Thailand	7	500	IV	-	-	+	+	+	
STIR-CD-28	Piscine	Thailand	7	500	IV	-	-	+	+	+	
STIR-CD-27	Piscine	Thailand	7	7	IV	-	-	+	+	+	
MRI Z1-198	**Delphine**	UK	7	399	IV	-	+	+	+	+	
STIR-CD-21	Piscine	Thailand	7	7	IV	-	+	+	+	+	
STIR-CD-14	Piscine	Vietnam	283	491	IV	-	+	+	+	+	4
STIR-CD-25	Piscine	Thailand	283	283	IV	-	+	+	+	+	
SA20-06	Piscine	Brazil	552	553	IS	-	+	+	+	+	
STIR-CD-01	Piscine	Kuwait	7	7	IV	+	-	+	+	+	
STIR-CD-02	Piscine	Kuwait	7	7	IV	+	-	+	+	+	
STIR-CD-03	Piscine	Kuwait	7	7	IV	+	-	+	+	+	
STIR-CD-04	Piscine	Kuwait	7	7	IV	+	-	+	+	+	
STIR-CD-05	Piscine	Kuwait	7	7	IV	+	-	+	+	+	5
STIR-CD-29	Piscine	Belgium	552	261	IV	+	-	+	+	+	
STIR-CD-30	Piscine	Australia	552	261	IV	+	-	+	+	+	
STIR-CD-31	Piscine	Australia	552	261	IV	+	-	+	+	+	
STIR-CD-32	Piscine	Australia	552	261	IV	+	-	+	+	+	
STIR-CD-09	Piscine	Colombia	552	260	IV	+	+	+	+	+	
STIR-CD-10	Piscine	Colombia	552	260	IV	+	+	+	+	+	
STIR-CD-11	Piscine	Colombia	552	260	IV	+	+	+	+	+	
STIR-CD-12	Piscine	Colombia	552	260	IV	+	+	+	+	+	
STIR-CD-13	Piscine	Costa Rica	552	260	IV	+	+	+	+	+	6
STIR-CD-17	Piscine	Honduras	552	260	IV and IS	+	+	+	+	+	
STIR-CD-18	Piscine	Honduras	552	260	IV	+	+	+	+	+	
STIR-CD-19	Piscine	Honduras	552	260	IV	+	+	+	+	+	
STIR-CD-07	Piscine	Honduras	552	260	IV	+	+	+	+	+	

4.4.6 Evidence for niche restriction

The two piscine isolates belonging to the fish-adapted sub-group of *S. agalactiae* (STIR-CD-17 and SA20-06) were found to be very similar based on their core genome as well as their accessory genome. Because the STIR-CD-17 genome is incomplete, the genome of SA20-06 was used as a representative of the fish-adapted sub-group for whole genome comparison using ACT (**Figure 4.9.**). This analysis showed that the SA20-06 genome is of reduced size compared to other *S. agalactiae* genomes due to the absence of large DNA fragments. Most of the missing DNA fragments correspond to regions that have been well defined in *S. agalactiae* NEM316 and referred to as putative pathogenicity islands (PAI) because they exhibit a mixture of genetic signatures of different mobile elements and are usually found next to tRNA genes (Glaser et al., 2002). These islands are known regions of genomic plasticity that shaped the evolution of *S. agalactiae* through the acquisition of traits for virulence and adaptation (Tettelin et al., 2005; Glaser et al., 2002; Brochet et al., 2008a). Accordingly, the fish-associated genes from Loci 2, 3 and 5 were found within islands IV or VI (**Figure 4.9**). However the complete or partial absence of most putative PAI (as well as other regions) is uncommon and has never been reported in previous studies of *S. agalactiae* of human origin.

In recent years, genome sequences of long-established and recent host-restricted bacteria have provided insights into the evolutionary forces driven by this lifestyle. The transition to host-pathogen or symbiotic restriction has been associated with a sharp increase of mobile elements (mostly insertion sequences) followed by genomic rearrangements and deletions of varying size (Moran and Plague, 2004). The accumulation of mobile elements is believed to be the consequence of a reduced effective population size that lowers the efficiency by which purifying selection maintains genes, thereby providing opportunities for insertions (Parkhill et al., 2003). The proliferation of mobile elements has consequences, both beneficial and deleterious: (1) transposition often leads to the inactivation of genes; (2) mobile elements may carry or activate genes involved in virulence or adaptation; (3)

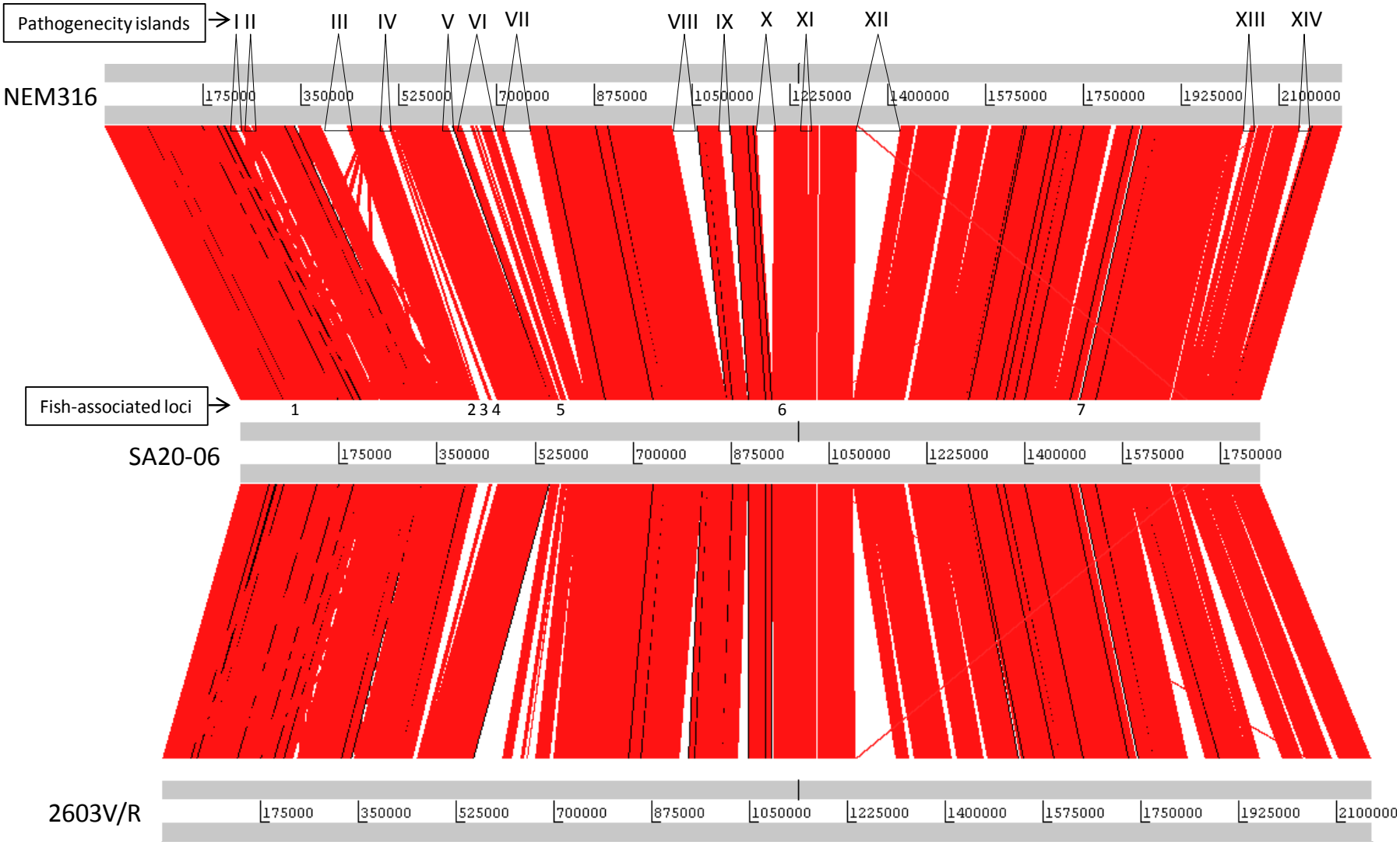
repeated transposition of elements results in homologous regions scattered in the genome, which yields the basis for rearrangements and deletion of large fragments (Al Safadi et al., 2010; Moran and Plague, 2004). Another hallmark during or following host-restriction is the abundance of pseudogenes, which are putatively inactivated coding sequences due to mutations (Ochman and Davalos, 2006). The abundance of pseudogenes is likely due to a combination of factors including genetic drift associated with population bottleneck, but also the result of a constant and rich environment provided by the host that renders some genes useless and eliminated (Moran and Plague, 2004). Therefore, if proliferation of mobile elements is expected in the short term, inactivation and loss of large genetic elements are expected in the longer run. Genomic characteristics of long-term host dependence include genome size reduction, deletion or mutation of mobile elements beyond recognition, and reduced metabolic capacities (Moran and Plague, 2004).

The recently described bovine *S. agalactiae* genome was found to be of much larger size than *S. agalactiae* genomes from human source due to large acquisitions of genetic material by horizontal gene transfer including material from other mastitis-causing streptococci (Richards et al., 2011). More importantly, a distinctive feature of the bovine *S. agalactiae* genome was the high frequency of recently acquired insertion sequences (IS) (97 IS in the bovine strain as compared to 30 or less in the human strains); the authors hypothesised a possible evolutionary bottleneck, which is in agreement with other studies showing a more limited genomic diversity in the bovine-associated population compared to the human (Bisharat et al., 2004). The bovine *S. agalactiae* genome therefore presents features of niche adaptation and possible transit towards a more restricted niche.

The fish-adapted *S. agalactiae* subgroup exhibits a limited genomic diversity (Chapter 2), but it contrasts with the bovine-adapted *S. agalactiae* genome at numerous levels. Molecular epidemiological studies failed to identify any of the frequently encountered IS (Chapter 2) and, accordingly, no well conserved IS was identified in the STIR-CD-17 genome.

Moreover, the genome is of reduced size due to absence or deletion of large genomic fragments and it contains a high number of pseudogenes (102 pseudogenes in STIR-CD-17 in contrast to 1 in 2603V/R, 32 in A909 and 40 in NEM316). Finally, among all the strains investigated during this project, isolates belonging to the fish-adapted subgroup were the most laborious, and it is indeed not without reason that these strains have been misclassified in the past as a new species, *S. difficile*, characterised by: (1) an absence of growth at 37°C, (2) slow growth on BHI agar plates at 30°C (1mm colonies after 48h incubation), (3) absence of β-haemolysis on sheep blood agar plates, and (3) a restricted metabolic capability, as assessed by biochemical tests (Eldar et al., 1994). Metabolic pathways were not investigated in this genomic study; however and as mentioned in section 4.3.1., genes considered to be conserved among the *S. agalactiae* species were reported absent from STIR-CD-17 by the PGAAP, including numerous genes encoding for proteins with metabolic functions like the fructose-6-phosphate aldolase, a 1,4 dihydroxy-2-naphthoate octaprenyltransferase, or a glutamate dehydrogenase. BLASTN searches failed in identifying these enzyme-coding genes within SA20-06, whereas it was present in other *S. agalactiae* genomes. This observation highlights the putative loss of metabolic pathways, but further investigation is required. In conclusion, genomic features of the fish-adapted subgroup of *S. agalactiae* suggest that, unlike bovine *S. agalactiae*, it has undergone full niche restriction.

Figure 4.9 (next page) Pairwise comparison of the genomic sequences of SA20-06 (middle) against NEM316 (top) and 2603V/R (bottom) using ACT (Carver et al., 2005). The sequences have been aligned from the predicted replication origins (*oriC*; right). The red bars separating each genome represent similarity matches identified by reciprocal BLASTN analysis, with a score cutoff of 300. The position of putative pathogenicity islands (Herbert et al., 2005) is highlighted in the NEM316 genome, whereas the position of putative fish-associated loci is highlighted in SA20-06.


Pathogenecity islands
I II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV
NEM316
175000
350000
525000
700000
875000
1050000
1225000
1400000
1575000
1750000
1925000
2100000
Fish-associated loci
1
2 3 4
5
6
7
SA20-06
175000
350000
525000
700000
875000
1050000
1225000
1400000
1575000
1750000
2603V/R
175000
350000
525000
700000
875000
1050000
1225000
1400000
1575000
1750000
1925000
2100000

4.5 CONCLUSIONS

The non-haemolytic *S. agalactiae* of CC552 have a distinctive genome characterised by a limited number of IS, an abundance of pseudogenes, a small genomic size, and an apparent loss of metabolic functions considered conserved within the species, indicating that this fish-adapted subgroup of isolates has undergone niche restriction. Accordingly, these isolates are the most divergent of all *S. agalactiae* based on phylogenetic analysis of the core genome, which corroborate, at the genomic level, previous work showing that non-haemolytic *S. agalactiae* recovered from fish exhibit limited genomic diversity and are distinct from human or bovine isolates (Chapter 2). Conversely, β-haemolytic isolates recovered from fish and belonging to the human-associated CC7 were found to cluster with human isolates of CC7, further supporting the possibility that some strains may represent a zoonotic or anthroponotic hazard (Evans et al., 2008 and 2009; Chapter 2). Comparative analysis of the accessory genome enabled the identification of a cluster of genes uniquely shared between CC552 and CC7 among the *S. agalactiae* species (Locus 3) and encoding for proteins involved in carbohydrate transport and metabolism, possibly necessary in providing precursors for EPS formation, CPS formation, or metabolic advantages required for their fitness as fish pathogens. Other loci identified were specific to strain STIR-CD-17 or to CC552 based on limited genomic comparisons; however the extension of this analysis through the PCR screening of a larger population of *S. agalactiae* suggested that some of these genes may be present in isolates of CC7. In the future, a larger and more diverse bacterial population should be screened and more putative fish-associated genes incorporated in the study to confirm these preliminary results. Some of the loci identified corresponded either to MME (Locus 1 and 7) or showed a mixture of genetic "signatures" of MGE (Locus 5, 6 and 8), suggesting the acquisition of these through horizontal gene transfer. It is not possible at present to identify whether these genes were acquired through intraspecies transfers or interspecies transfers from the aquatic environment. These putative MGEs are however unlikely to have been acquired recently as major components of these

have lost their function (*e.g.* resolvases appeared as pseudogenes in Locus 5). Among the genes identified within the putative MGEs, some may encode for proteins involved in adaptation to the aquatic environment (*e.g.* SHMT, Locus 8) or virulence (*e.g.* surface antigen, Locus 5); however numerous fish-associated genes identified were found to encode hypothetical proteins predicted to be localised in the cytoplasm. Finding predicted cytoplasmic proteins as genuine components of the extracellular proteome ('moonlighting' proteins) is a recognized phenomenon that has been the focus of recent discussion (Henderson and Martin, 2011; Copley, 2012), suggesting that these genes may still play a role in virulence. The STIR-CD-17 genome was mostly characterised by genetic material loss, with limited additional genetic material. Therefore, it may be hypothesised that the virulence of this strain in fish does not fully result from the acquisition of new genes but from the expression of virulence genes well defined in human isolates.

4.6 REFERENCES

Al Safadi, R., Amor, S., Hery-Arnaud, G., Spellerberg, B., Lanotte, P., Mereghetti, L., Gannier, F., Quentin, R., Rosenau, A., 2010. Enhanced expression of *lmb* gene encoding laminin-binding protein in *Streptococcus agalactiae* strains harboring IS1548 in scpB-lmb intergenic region. PloS One 5, e10794.

Bateman, A. and Rawlings, N.D., 2003. The CHAP domain: A large family of amidases including GSP amidase and peptidoglycan hydrolases. Trends Biochem. Sci. 28, 234-237.

Bisharat, N., Crook, D.W., Leigh, J., Harding, R.M., Ward, P.N., Coffey, T.J., Maiden, M.C., Peto, T., Jones, N., 2004. Hyperinvasive neonatal group B streptococcus has arisen from a bovine ancestor. J. Clin. Microbiol. 42, 2161-2167.

Bowater, R.O., Forbes-Faulkner, J., Anderson, I.G., Condon, K., Robinson, B., Kong, F., Gilbert, G.L., Reynolds, A., Hyland, S., McPherson, G., Brien, J.O., Blyde, D., 2012. Natural outbreak of *Streptococcus agalactiae* (GBS) infection in wild giant queensland grouper, *Epinephelus lanceolatus* (Bloch), and other wild fish in northern Queensland, Australia. J. Fish Dis. 35, 173-186.

Briggs, R.E., Tatum, F.M., Casey, T.A., Frank, G.H., 1994. Characterization of a restriction endonuclease, PhaI, from pasteurella haemolytica serotype A1 and protection of heterologous DNA by a cloned PhaI methyltransferase gene. Appl. Environ. Microbiol. 60, 2006-2010.

Brochet, M., Rusniok, C., Couvé, E., Dramsi, S., Poyart, C., Trieu-Cuot, P., Kunst, F., Glaser, P., 2008a. Shaping a bacterial genome by large chromosomal replacements, the evolutionary history of *Streptococcus agalactiae*. PNAS 105, 15961-15966.

Brochet, M., Couvé, E., Zouine, M., Vallaeys, T., Rusniok, C., Lamy, M.C., Buchrieser, C., Trieu-Cuot, P., Kunst, F., Poyart, C., 2006. Genomic diversity and evolution within the species *Streptococcus agalactiae*. Microb. Infect. 8, 1227-1243.

Brochet, M., Couvé, E., Glaser, P., Guédon, G., Payot, S., October 15, 2008b. Integrative conjugative elements and related elements are major contributors to the genome diversity of *Streptococcus agalactiae*. J. Bacteriol. 190, 6913-6917.

Carver, T., Berriman, M., Tivey, A., Patel, C., Böhme, U., Barrell, B.G., Parkhill, J., Rajandream, M., 2008. Artemis and ACT: Viewing, annotating and comparing sequences stored in a relational database. Bioinformatics 24, 2672-2676.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution 17, 540-552.

Chopin, M., Chopin, A., Bidnenko, E., 2005. Phage abortive infection in lactococci: Variations on a theme. Curr. Opin. Microbiol. 8, 473-479.

Cieslewicz, M.J., Chaffin, D., Glusman, G., Kasper, D., Madan, A., Rodrigues, S., Fahey, J., Wessels, M.R., Rubens, C.E., 2005. Structural and genetic diversity of group B streptococcus capsular polysaccharides. Infection and Immunity 73, 3096-3103.

Clamp, M., Cuff, J., Searle, S.M., Barton, G.J., 2004. The Jalview java alignment editor. Bioinformatics 20, 426-427.

Copley, S.D., 2012. Moonlighting is mainstream: Paradigm adjustment required. Bioessays 34, 578-588.

Davies, M.R., Shera, J., Van Domselaar, G.H., Sriprakash, K.S., McMillan, D.J., April 1, 2009. A novel integrative conjugative element mediates genetic transfer from group G streptococcus to other β-hemolytic streptococci. J. Bacteriol. 191, 2257-2265.

Dawid, S., Roche, A.M., Weiser, J.N., January 2007. The blp bacteriocins of streptococcus pneumoniae mediate intraspecies competition both in vitro and in vivo. Infect. Immun. 75, 443-451.

Delannoy, C.M.J., Zadoks, R.N., Lainson, F.A., Ferguson, H.W., Crumlish, M., Turnbull, J.F., Fontaine, M.C., 2012. Draft genome sequence of a nonhemolytic fish-pathogenic *Streptococcus agalactiae* strain. J. Bacteriol. 194, 6341-6342.

Eldar, A., Bejerano, Y., Bercovier, H., 1994. *Streptococcus shiloi* and *Streptococcus difficile*: Two new streptococcal species causing a meningoencephalitis in fish. Curr. Microbiol. 28, 139-143.

Evans, J.J., Klesius, P.H., Pasnik, D.J., Bohnsack, J.F., 2009. Human *Streptococcus agalactiae* isolate in Nile tilapia (*Oreochromis niloticus*). Emerg. Infect. Diseases 15, 774.

Evans, J.J., Bohnsack, J.F., Klesius, P.H., Whiting, A.A., Garcia, J.C., Shoemaker, C.A., Takahashi, S., 2008. Phylogenetic relationships among *Streptococcus agalactiae* isolated from piscine, dolphin, bovine and human sources: A dolphin and piscine lineage associated with a fish epidemic in Kuwait is also associated with human neonatal infections in Japan. J. Med. Microbiol. 57, 1369-1376.

Evans, J.J., Klesius, P.H., Gilbert, P.M., Shoemaker, C.A., Al Sarawi, M.A., Landsberg, J., Duremdez, R., Al Marzouk, A., Al Zenki, S., 2002. Characterization of β-haemolytic group B *Streptococcus agalactiae* in cultured seabream, *Sparus auratus* L., and wild mullet, *Liza klunzingeri* (Day), in Kuwait. J. Fish Dis. 25, 505-513.

Finch, L.A. and Martin, D.R., 1984. Human and bovine group B streptococci: Two distinct populations. J. Appl. Microbiol. 57, 273-278.

Forde, A. and Fitzgerald, G.F., 2003. Molecular organization of exopolysaccharide (EPS) encoding genes on the lactococcal bacteriophage adsorption blocking plasmid, pCI658. Plasmid 49, 130-142.

Furrer, B., Candrian, U., Hoefelein, C., Luethy, J., 1991. Detection and identification of listeria monocytogenes in cooked sausage products and in milk by in vitro amplification of haemolysin gene fragments. J. Appl. Microbiol. 70, 372-379.

Glaser, P., Rusniok, C., Buchrieser, C., Chevalier, F., Frangeul, L., Msadek, T., Zouine, M., Couvé, E., Lalioui, L., Poyart, C., Trieu-Cuot, P., Kunst, F., 2002. Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. Mol. Microbiol. 45, 1499-1513.

Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221-224.

Grossiord, B., Vaughan, E., E., Luesink, E., de Vos, W., M., 1998. Genetics of galactose utilisation via the Leloir pathway in lactic acid bacteria. Lait 78, 77-84.

Haguenoer, E., Baty, G., Pourcel, C., Lartigue, M., Domelier, A., Rosenau, A., Quentin, R., Mereghetti, L., Lanotte, P., 2011. A multi locus variable number of tandem repeat analysis (MLVA) scheme for *Streptococcus agalactiae* genotyping. BMC Microbiol. 11, 171.

Halio, S.B., Blumentals, I.I., Short, S.A., Merrill, B.M., Kelly, R.M., 1996. Sequence, expression in escherichia coli, and analysis of the gene encoding a novel intracellular protease (PfpI) from the hyperthermophilic archaeon *Pyrococcus furiosus*. J. Bacteriol. 178, 2605-2612.

Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symposium Series, 1999, 41, 95-98.

Henderson, B. and Martin, A., 2011. Bacterial virulence in the moonlight: Multitasking bacterial moonlighting proteins are virulence determinants in infectious disease. Infect. Immun. 79, 3476-3491.

Herbert, M., Beveridge, C., McCormick, D., Aten, E., Jones, N., Snyder, L., Saunders, N., 2005. Genetic islands of *Streptococcus agalactiae* strains NEM316 and 2603VR and their presence in other group B streptococcal strains. BMC Microbiol. 5, 31.

Horvath, P. and Barrangou, R., 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science 327, 167-170.

Huson, D., Richter, D., Rausch, C., Dezulian, T., Franz, M., Rupp, R., 2007. Dendroscope: An interactive viewer for large phylogenetic trees. BMC Bioinformatics 8, 460.

Jones, N., Oliver, K.A., Barry, J., Harding, R.M., Bisharat, N., Spratt, B.G., Peto, T., Crook, D.W., 2006. Enhanced invasiveness of bovine-derived neonatal sequence type 17 group B streptococcus is independent of capsular serotype. Clin. Infect. Dis. 42, 915-924.

Jones, N., Bohnsack, J.F., Takahashi, S., Oliver, K.A., Chan, M.S., Kunst, F., Glaser, P., Rusniok, C., Crook, D.W., Harding, R.M., Bisharat, N., Spratt, B.G., 2003. Multilocus sequence typing system for group B streptococcus. J. Clin. Microbiol. 41, 2530-2536.

Kong, F., Gowan, S., Martin, D., James, G., Gilbert, G.L., 2002. Molecular profiles of group B streptococcal surface protein antigen genes: Relationship to molecular serotypes. J. Clin. Microbiol. 40, 620-626.

Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R., Wilson, G., 1994. The DNA (cytosine-5) methyltransferases. Nucleic Acids Res. 22, 1-10.

Laing, C., Buchanan, C., Taboada, E., Zhang, Y., Kropinski, A., Villegas, A., Thomas, J., Gannon, V., 2010. Pan-genome sequence analysis using panseq: An online tool for the rapid analysis of core and accessory genomic regions. BMC Bioinformatics 11, 461.

Lancefield, R.C., McCarty, M., Everly, W.N., 1975. Multiple mouse-protective antibodies directed against group B streptococci. Special reference to antibodies effective against protein antigens. The Journal of Experimental Medicine 142, 165-179.

Levander, F. and Rådström, P., 2001. Requirement for phosphoglucomutase in exopolysaccharide biosynthesis in glucose- and lactose-utilizing *Streptococcus thermophilus.* Appl. Environ. Microbiol. 67, 2734-2738.

Liu, G., Zhang, W., Lu, C., 2012. Complete genome sequence of *Streptococcus agalactiae GD201008*-001, isolated in china from tilapia with meningoencephalitis. Journal of Bacteriology 194, 6653-6653.

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

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., 2011. Evolution and classification of the CRISPR–Cas systems. Nature Reviews Microbiology 9, 467-477.

Mian, G.F., Godoy, D.T., Leal, C.A.G., Yuhara, T.Y., Costa, G.M., Figueiredo, H.C.P., 2009. Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. Vet. Microbiol. 136, 180-183.

Milne, I., Wright, F., Rowe, G., Marshall, D.F., Husmeier, D., McGuire, G., 2004. TOPALi: Software for automatic identification of recombinant sequences within DNA multiple alignments. Bioinformatics 20, 1806-1807.

Moran, N.A. and Plague, G.R., 2004. Genomic changes following host restriction in bacteria. Curr. Opin. Genet. Dev. 14, 627-633.

Ochman, H. and Davalos, L.M., 2006. The nature and dynamics of bacterial genomes. Science 311, 1730-1733.

Ohnishi, Y., Yamazaki, H., Kato, J., Tomono, A., Horinouchi, S., 2005. AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. Biosci. Biotechnol. Biochem. 69, 431-439.

Pao, S.S., Paulsen, I.T., Saier, M.H., 1998. Major facilitator superfamily. Microb. Mol. Biol. Rev. 62, 1-34.

Parkhill, J., Sebaihia, M., Preston, A., Murphy, L.D., Thomson, N., Harris, D.E., Holden, M.T.G., Churcher, C.M., Bentley, S.D., Mungall, K.L., 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. Nat. Genet. 35, 32-40.

Pereira, U.P., Mian, G.F., Oliveira, I.C.M., Benchetrit, L.C., Costa, G.M., Figueiredo, H.C.P., 2010. Genotyping of *Streptococcus agalactiae* strains isolated from fish, human and cattle and their virulence potential in Nile tilapia. Vet. Microbiol. 140, 186-192.

Podbielski, A., Blankenstein, O., Lütticken, R., 1994. Molecular characterization of the *cfb* gene encoding group B streptococcal CAMP-factor. Med. Microbiol. Immunol. (Berl.) 183, 239-256.

Pritzlaff, C.A., Chang, J.C.W., Kuo, S.P., Tamura, G.S., Rubens, C.E., Nizet, V., 2001. Genetic basis for the β-haemolytic/cytolytic activity of group B streptococcus. Mol. Microbiol. 39, 236-248.

Rasko, D., Myers, G., Ravel, J., 2005. Visualization of comparative genomic analyses by BLAST score ratio. BMC Bioinformatics 6, 2.

Richards, V.P., Lang, P., Pavinski Bitar, P.D., Lefébure, T., Schukken, Y.H., Zadoks, R.N., Stanhope, M.J., 2011. Comparative genomics and the role of lateral gene transfer in the evolution of bovine adapted *Streptococcus agalactiae*. Infect. Genet. Evol. 11, 1263-1275.

Rodríguez, F., Oliver, J.L., Marín, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142, 485-501.

Ronquist, F. and Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Sadowski, P., 1986. Site-specific recombinases: Changing partners and doing the twist. J. Bacteriol. 165, 341-347.

Saunders, N.J. and Snyder, L.A.S., 2002. The minimal mobile element. Microbiology 148, 3756-3760.

Sørensen, U.B.S., Poulsen, K., Ghezzo, C., Margarit, I., Kilian, M., 2010. Emergence and global dissemination of host-specific *Streptococcus agalactiae* clones. Mbio 1.3.

Stingele, F., Neeser, J.R., Mollet, B., 1996. Identification and characterization of the eps (exopolysaccharide) gene cluster from *Streptococcus thermophilus* Sfi6. J. Bacteriol. 178, 1680-1690.

Suanyuk, N., Kong, F., Ko, D., Gilbert, G.L., Supamattaya, K., 2008. Occurrence of rare genotypes of *Streptococcus agalactiae* in cultured red tilapia *Oreochromis* sp. and nile tilapia *O. niloticus* in Thailand-Relationship to human isolates? Aquaculture 284, 35-40.

Sutcliffe, I.C., Black, G.W., Harrington, D.J., 2008. Bioinformatic insights into the biosynthesis of the group B carbohydrate in *Streptococcus agalactiae*. Microbiology 154, 1354-1363.

Tatusov, R., Fedorova, N., Jackson, J., Jacobs, A., Kiryutin, B., Koonin, E., Krylov, D., Mazumder, R., Mekhedov, S., Nikolskaya, A., Rao, B.S., Smirnov, S., Sverdlov, A., Vasudevan, S., Wolf, Y., Yin, J., Natale, D., 2003. The COG database: An updated version includes eukaryotes. BMC Bioinformatics 4, 41.

Tettelin, H., Riley, D., Cattuto, C., Medini, D., 2008. Comparative genomics: The bacterial pan-genome. Curr. Opin. Microbiol. 11, 472-477.

Tettelin, H., Masignani, V., Cieslewicz, M.J., Eisen, J.A., Peterson, S., Wessels, M.R., Paulsen, I.T., Nelson, K.E., Margarit, I., Read, T.D., Madoff, L.C., Wolf, A.M., Beanan, M.J., Brinkac, L.M., Daugherty, S.C., DeBoy, R.T., Durkin, A.S., Kolonay, J.F., Madupu, R., Lewis, M.R., Radune, D., Fedorova, N.B., Scanlan, D., Khouri, H., Mulligan, S., Carty, H.A., Cline, R.T., Van Aken, S.E., Gill, J., Scarselli, M., Mora, M., Iacobini, E.T., Brettoni, C., Galli, G., Mariani, M., Vegni, F., Maione, D., Rinaudo, D., Rappuoli, R., Telford, J.L., Kasper, D.L., Grandi, G., Fraser, C.M., 2002. Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. Proceedings of the National Academy of Sciences 99, 12391-12396.

Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., DeBoy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sullivan, S.A., Daugherty, S.C., Haft, D.H., Selengut, J., Gwinn, M.L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K.J.B., Smith, S., Utterback, T.R., White, O., Rubens, C.E., Grandi, G., Madoff, L.C., Kasper, D.L., Telford, J.L., Wessels, M.R., Rappuoli, R., Fraser, C.M., 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". Proc. NatL. AcaD. Sci. U S A 102, 13950-13955.

Tocher, D.R., 2003. Metabolism and functions of lipids and fatty acids in teleost fish. Rev. Fish. Sci. 11, 107-184.

Urwin, R. and Maiden, M.C.J., 2003. Multi-locus sequence typing: A tool for global epidemiology. Trends Microbiol. 11, 479-487.

Waditee-Sirisattha, R., Sittipol, D., Tanaka, Y., Takabe, T., 2012. Overexpression of serine hydroxymethyltransferase from halotolerant cyanobacterium in *Escherichia coli* results in increased accumulation of choline precursors and enhanced salinity tolerance. FEMS Microbiol. Lett. 333, 46-53.

Wang, B., Jian, J., Lu, Y., Cai, S., Huang, Y., Tang, J., Wu, Z., 2012. Complete genome sequence of *Streptococcus agalactiae* ZQ0910, a pathogen causing meningoencephalitis in the GIFT strain of Nile tilapia (*Oreochromis niloticus*). J. Bacteriol. 194, 5132-5133.

Ye, X., Li, J., Lu, M., Deng, G., Jiang, X., Tian, Y., Quan, Y., Jian, Q., 2011. Identification and molecular typing of *Streptococcus agalactiae* isolated from pond-cultured tilapia in China. Fish. Sci. 77, 623-632.

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., 2010. PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26, 1608-1615.

Zadoks, R.N., Middleton, J.R., McDougall, S., Katholm, J., Schukken, Y.H., 2011. Molecular epidemiology of mastitis pathogens of dairy cattle and comparative relevance to humans. J. Mammary Gland Biol. Neoplasia 1-16.

CHAPTER 5

Comparative analysis of *Streptococcus agalactiae* virulence determinants

Christian M.J. Delannoy, Ruth N. Zadoks, Margaret Crumlish, Jimmy Turnbull, Michael C. Fontaine

Status: This manuscript is in preparation for submission in a relevant peer-reviewed journal.

Contribution: The candidate designed the study, performed the analytical work and wrote the present manuscript with corrections and editing from all co-authors. The phenotypic CAMP test was performed by an MSc student (David Rodgers) under the supervision of the candidate.

5.1 ABSTRACT

Streptococcus agalactiae is a pathogen of multiple host species. Numerous genes encoding virulence determinants have been well-characterised in *S. agalactiae* isolated from humans, but the presence of these genes and the function of their encoded proteins have not been thoroughly investigated in *S. agalactiae* of fish origin. In the present study, genes encoding recognised virulence factors in human *S. agalactiae* were selected and their presence evaluated within the genome of STIR-CD-17, an isolate belonging to a fish-adapted subpopulation of *S. agalactiae*. A literature review of those genes found in STIR-CD-17 is presented, accompanied by the results of bioinformatic analysis to assess structural homologies between human and piscine *S. agalactiae*. Numerous genes were completely absent in STIR-CD-17, while the *cyl* operon responsible for the β-haemolysin production was found to be only partially present, indicating that their encoded proteins are not important contributors to pathogenicity of *S. agalactiae* in fish. The gene encoding an immunogenic bacterial adhesin in certain *S. agalactiae* from humans (gbs2018) was identified as a distinct variant, unique to fish isolates, which possibly reflects differences or adaptations in the function of the protein. Finally, a set of genes were found to be well-conserved in STIR-CD-17 and included *fbsA*, *pavA*, *srr1*, *cfb*, *hylB*, *ponA* and *sodA*. The *cspA* gene was also found to be well conserved, but a deletion responsible for a frameshift suggested that, if the gene is expressed, the protein would be secreted and not cell-wall anchored. Moreover, the *cps* genes were also found to be well-conserved, with the exception of *cpsK*, but whether the variations in *cpsK* affect the biosynthesis of the capsule is unknown. This work provides a basis for further post-genomic studies *in vitro* and *in vivo*.

Keywords: Virulence, genome, *Streptococcus agalactiae*, ST260

5.2 INTRODUCTION

In the late 19th century, Robert Koch formulated guidelines that laid the foundation for establishing the link between pathogens and disease. For over 100 years, his postulate was gold standards to define microbial virulence, despite the fact that it was not possible to establish the link between pathogen and disease for a number of microorganisms that do not grow *in vitro*, only cause the disease in partnership with other organisms, or require specific environmental factors or host predispositions. In order to overcome some of these limitations, later studies provided revisions of the Koch postulate which encompassed immunological and/or epidemiological proofs of disease causation (reviewed by Fredricks and Relmans, 1996).

With the advent of molecular microbiology in the late twentieth century, studies then focused on specific genes and the importance of their corresponding protein to virulence. In order to define a virulence factor, Stanley Falkow devised a molecular version of Koch's postulate that consist of 3 criteria: (1) the phenotype or trait under investigation should be associated with pathogenic strains and absent from the non-pathogenic ones, (2) the specific inactivation of the gene(s) encoding for the suspected virulence trait should lead to a measurable loss of pathogenicity in an appropriate animal model, (3) the subsequent reintroduction of the mutated gene(s) should restore the pathogenicity in the same animal model (Falkow, 1988). However, this postulate also has limitations since it suggests a clear distinction between pathogens and non-pathogens, which is not always the case. For example, *S. agalactiae* is carried as part of the commensal vaginal flora in woman and causes neonatal disease through vertical transmission, suggesting adaptive changes to a new environment with up-regulation or activation of specific genes involved in virulence (Yang et al., 2010).

In the last decade, the search for virulence genes has been revolutionized by the exponential increase in the number of available genome sequences and the development of comparative

genomic methodologies (Raskin et al., 2006). For example, genome analysis of *S. agalactiae* enabled the discovery of genes encoding an immunoprotective protein later found to be involved in pilus formation (Lauer et al., 2005). Purely *in silico* approaches based on identifying virulence genes by comparing sequence similarity with known virulence genes have been developed (Lin et al., 2011). These approaches have led to an era where the term "putativism" is legion. Indeed, genes may have niche-adapted functions and identification of genetic conservation can therefore only lead to an assumption of virulence that needs to be tested in an adapted experimental model to demonstrate functional conservation (Wassenaar and Gaastra, 2001). In human *S. agalactiae*, a computational approach has led to the identification of a myriad of putative virulence genes which have not undergone *in vitro* or *in vivo* studies thus-far (Lin et al., 2011).

However, numerous virulence genes have been well characterised in human *S. agalactiae* (Maisey et al., 2008). Those genes do not always fulfil Falkow's postulates, *e.g.* the *cps* operon considered to encode a virulence factor, the polysaccharide capsule, is present in both pathogenic and non pathogenic bacteria possibly as part of the core genome (Cieslewicz et al., 2005). Other virulence genes, however, have been shown to be part of mobile genetic elements (*e.g.* the *scpB* and *lmb* genes, encoding fibronectin-binding and laminin-binding surface proteins, are part of a composite transposon; Herbert et al., 2005) and their acquisition may enhance the virulence of specific strains in specific niches. Accordingly, evidence from molecular epidemiological studies using multi-locus sequence typing (MLST) suggest that specific strains are associated with specific niche, *e.g.* sequence type (ST) 17 with neonates, ST67 with cows, and ST260 with fish (Bisharat et al., 2004; Brochet et al., 2006; Chapter 2). Moreover, further evidence indicate that this association is not only due to differences in exposure, but to true differences in virulence as shown in experimental infection studies of tilapia comparing the virulence of *S. agalactiae* strains of distinct ST and/or host of origin (Garcia et al., 2008; Pereira et al., 2010; Chapter 3).

Comparative genomic analysis (Chapter 4) enabled the identification of genetic elements that are uniquely or predominantly associated with fish strains, but virulence elements that may be shared across host-species were not considered. In the present study, genes encoding recognised virulence factors in human *S. agalactiae* were selected and their presence assessed in the piscine *S. agalactiae* STIR-CD-17. A literature review of those genes found in STIR-CD-17 is presented, accompanied by bioinformatic analysis to assess structural homologies between human and piscine *S. agalactiae*. This approach enabled the identification of a set of genes which are potential candidates for further *in vitro* and *in vivo* studies in order to evaluate their putative role in virulence in the specific context of fish infections.

5.3 MATERIALS & METHODS

5.3.1 Virulence gene selection and identification

A review of the literature was performed in order to identify factors for which a role in virulence is supported by *in vitro* and/or *in vivo* studies of *S. agalactiae* from humans. Based upon their apparent primary function, these factors were classified into three major categories (Doran and Nizet, 2004): (1) adhesins, or factors responsible for the adherence and colonization to host epithelia, endothelia, or extracellular matrix, (2) invasins, or factors responsible for the invasion of host cells or the dissemination of bacteria across the extracellular matrix, and (3) immune evasins, or factors involved in the evasion of the host immune system. The systematic names (Locus tag) of the selected virulence factors in the three complete reference genomes available at the time of analysis (NEM316, A909 and 2603V/R; Feb 2012) were further retrieved from the literature (Lin et al., 2011) or through key words and/or BLAST searches into the NCBI website (http://www.ncbi.nlm.nih.gov/). Finally, the presence/absence of these virulence coding genes was evaluated in the piscine STIR-CD-17 genome. For this, BLASTN searches of the nucleotide sequences from the three complete genomes were performed against the STIR-CD-17 genome using the BioEdit

software (Hall, 1999; http://www.mbio.ncsu.edu/bioedit/bioedit.html) and as previously described (Chapter 4). When a partial or complete match was obtained using BioEdit, the Contig containing the matching sequence in STIR-CD-17 was used in pairwise comparisons against the three complete genomes using ACT (Chapter 4), which allow assessment of homology and determination of conservation of position within the genomes. Because the STIR-CD-17 genome is incomplete (Chapter 4), the apparent absence of a gene may represent a true negative result (the gene is missing from the Contig) or a false negative result (the relevant Contig is missing). In order to provide further confidance of missing genes, the Contigs in STIR-CD-17 that would contain genes neighbouring a putatively missing gene were retrieved and ACT comparisons were performed. The selected virulence factors, their categorization and their systematic names in genomes are listed in **Table 5.1**.

For the genes found to be well conserved among genomes, the homology of their encoded proteins was further evaluated using normalised BLAST score ratio analyses as previously described (Chapter 4). For genes found to be only partially conserved and for which multiple BLAST hits were obtained due to variable repeat segments or non homologous regions within the sequence, homology was evaluated through nucleotide and amino acid alignments using the MAFFT multiple sequence alignment web service implemented in JalView 2.4 (Clamp et al., 2004).

Table 5.1 List of known virulence genes with systematic gene names in the 3 complete reference genomes of human *S. agalactiae* and in the draft genome sequence of the piscine strain STIR-CD-17.

Category	Gene/island	Function/annotation	Systematic name/locus in genomes			
			NEM316	A909	2603	STIR-CD-17°
Adhesins						
	fbsA	Fibrinogen-binding protein FbsA	GBS1087	SAK1142	SAG1052	M3M_07935
	fbsB	Fibrinogen-binding protein FbsB	GBS0850	SAK0955	SAG0832	-
	pavA	Fibronectin-binding protein	GBS1263	SAK1277	SAG1190	M3M_03075
	srr 1	Serine-rich repeat protein 1	GBS1529	SAK1493	SAG1462	M3M_05192
	srr 2	Serine-rich repeat protein 2	-	-	-	-
	bibA	Immunogenic bacterial adhesin	GBS2018	SAK1999/2002	SAG2063	M3M_09338
	PI-1	Pilus island 1	GBS1474-1478	-	SAG0645-0649	-
	PI-2a	Pilus island 2a	-	-	SAG1404-1408	-
	PI-2b	Pilus island 2b	-	SAK1437-1442	-	M3M_06269-06299
	lmb	Laminin-binding protein	GBS1307	SAK1319	SAG1234	-
Invasins						
	cyl	β-hemolysin/cytolysin	GBS0644-55	SAK0790-0801	SAG0662-73	M3M_00345-00355
	cfb	CAMP factor	GBS2000	SAK1983	SAG2043	M3M_09048
	hylB	Hyaluronate lyase	GBS1270	SAK1284	SAG1197	M3M_03035
	rib	Surface protein rib	GBS0470	-	SAG0433	-
	bca	C-α protein	-	SAK0517	-	-
Immune evasins						
	cps	Cps gene cluster	GBS1237-47	SAK1251-1262	SAG1162-75	M3M_08948-09003
	neu	*neu* gene cluster	GBS1233-36	SAK1247-1250	SAG1158-61	M3M_09008-09023
	pbp1A/ponA	Penicillin-binding protein 1A	GBS0288	SAK0370	SAG0298	M3M_06939
	soda	Superoxide dismutase	GBS0808	SAK_0913	SAG0788	M3M_06204
	cspA	Serine protease cspA	GBS2008	SAK1991	SAG2053	Contig31
	scpB	C5a peptidase	GBS1308	SAK1320	SAG1236	-
	bac	C-β protein	-	SAK0186	-	-

°The Contig number is provided when genes are not annotated but found to be partially present.

5.3.2 Structural and phylogenetic analysis of *gbs2018*

A total of 38 complete *gbs2018* nucleotide sequences for which MLST results are known were retrieved from the NCBI website. These comprise the *gbs2018* from all published *S. agalactiae* genomes available at the time of analysis, plus sequences deposited in GenBank by Lamy et al. (2006) and Brochet et al. (2006).

The *gbs2018* gene is known to display sequence diversity: 3 major allelic variants have been described in isolates from human sources (Lamy et al., 2006). Even though all *gbs2018* genes display homologous regions, structural diversity is revealed by the acquisition or loss of nucleotide segments and a total of 11 distinct segments have been reported thus far (Lamy et al., 2006). Work by Lamy et al. (2006) thereby served as a reference for deciphering the structure of new variants encountered. In the current bioinformatic analysis, *gbs2018* sequences were aligned and compared using the MAFFT multiple sequence alignment web service implemented in JALVIEW 2.4. Distinct *gbs2018* variants identified were then compared pairwise using dot-plots (DOTTUP program with default setting with the exception of the word size of 12; http://mobyle.pasteur.fr), which allowed to determine the segments shared between different variants and the boundaries of these segments.

Phylogenetic analysis was then performed based on the concatenation of segments previously aligned independently. For this aim, specific segments identified within the *gbs2018* variants were manually extracted and aligned independently using MAFFT. Files were created and comprised a specific segment type with aligned sequences from all strains comprising that segment type. These files were then loaded in SEAVIEW (Gouy et al., 2010) and concatenated. Finally the output file was loaded once more into TOPALi v2.5 for phylogenetic tree construction. Tree construction was optimised using the model selection feature. The selected model (HKY model of Hasegawa, Kishino, and Yano; Hasegawa et al., 1985) was then used to estimate a Bayesian phylogenetic tree using the MrBayes program (Ronquist et al., 2003) launched from TOPALi v2.5. The MrBayes settings were 2

runs of 625,000 generations and a burn-in period of 125,000 generations, with trees sampled every 10 generations. The consensus tree was imported into DENDROSCOPE v3.2.1 (Huson et al., 2007), where visualization and editing were performed as required.

5.3.3 Phenotypic test: CAMP reaction

The Christie/Atkins/Munch-Petersen (CAMP) test (Wilkinson, 1977) was carried out to evaluate the CAMP activity of all the aquatic *S. agalactiae* collection of fish and sea mammals' origin (Chapter 2, section 2.3.1). Isolates were streaked on 5% (v/v) sheep blood agar plates (E&C Laboratories, Scotland, UK) at right angles to a β-haemolysin-producing strain of *Staphylococcus aureus* (ID No: NCTC7428) in duplicate, and incubated aerobically at 28°C for up to 48h. Also included were 2 *Streptococcus canis* isolates (Isolate FSL Z3-022; FSL-Z3-023) as CAMP- positive controls, and a *Streptococcus dysagalactiae* subsp. *dysagalactiae* (Isolate QMP Z3-580) as a CAMP-negative control.

5.4 RESULTS AND DISCUSSION

5.4.1 Adhesins

Among the putative adhesins investigated, 5 were detected within the piscine STIR-CD-17 genome (**Table 5.1**). These comprised genes encoding the fibronectin-binding protein PavA and sortase-processed surface proteins such as the fibrinogen-binding protein FbsA, the pilus type 2 (variant PI-2b), the immunogenic bacterial adhesin BibA and the serine-rich repeat protein Srr1.

5.4.1.1 Fibrinogen-binding protein

The ability to bind fibrinogen is an important characteristic of strains belonging to the hyper-virulent ST17 clone responsible for neonatal infectious disease in humans (Rosenau et al., 2007). Two proteins with an ability to bind both soluble and immobilized human fibrinogen have been identified and referred to as FbsA and FbsB/Fgag (Schubert et al., 2002; Jacobsson et al., 2003). Both proteins are not found in all *S. agalactiae* strains

(Rosenau et al., 2007) but they are usually found conjointly in the hyper-virulent ST17 clone (Brochet et al., 2006). In STIR-CD-17, only *fbsA* was detected.

In vitro and *in vivo* studies of *S. agalactiae* of human origin have confirmed the role of FbsA as an important virulence factor. *In vitro*, an *fbsA* knock-out mutant resulted in loss of fibrinogen-binding abilities and impaired growth in human blood, suggesting that the protein may also contribute to phagocytosis resistance (Schubert et al., 2002). Moreover, assessment of an *fbsA* knock-out mutant in a mouse model of infection resulted in significant reduction of mortalities, less pronounced weight loss, and less severe arthritis (Jonsson et al., 2005).

FbsA has typical features of a cell-surface located protein with an N-terminal signal peptide sequence and a C-terminal wall anchoring region preceded by an LPXTG motif. Its inner structure contains variable 16-amino acid long repetitive units that have been described to interact with human fibrinogen and to promote adhesion to human epithelial and endothelial cells (Schubert et al., 2002 and 2004; Tenenbaum et al., 2005). The number of repeats of this unit, which is composed of consensus sequence G-N/S/T-V-L-A/E/M/Q-R-R-X-K/R/W-A/D/E/N/Q-A/F/I/L/V/Y-X-X-K/R-X-X, varies between 2 and 30 amongst different strains (Rosenau et al., 2007; Shubert et al., 2002), but a single repeat binds fibrinogen, as demonstrated using synthetic peptides (Shubert et al., 2002).

The predicted fibrinogen binding region in STIR-CD-17 is composed of 2 repeat sequences of G-N-V-L-E-R-R-Q-R/H-D-V-D/E-N-K-S-Q (**Figure 5.1**). It is not known if the repeat segment binds piscine fibrinogen, as all previous experiments have focused on human fibrinogen. However, the high degree of conservation of the gene sequence and of the repeat segments suggests that FbsA might play a role in STIR-CD-17 pathogenesis in fish, although further studies using piscine *S. agalactiae* and/or fish models of infection are required.

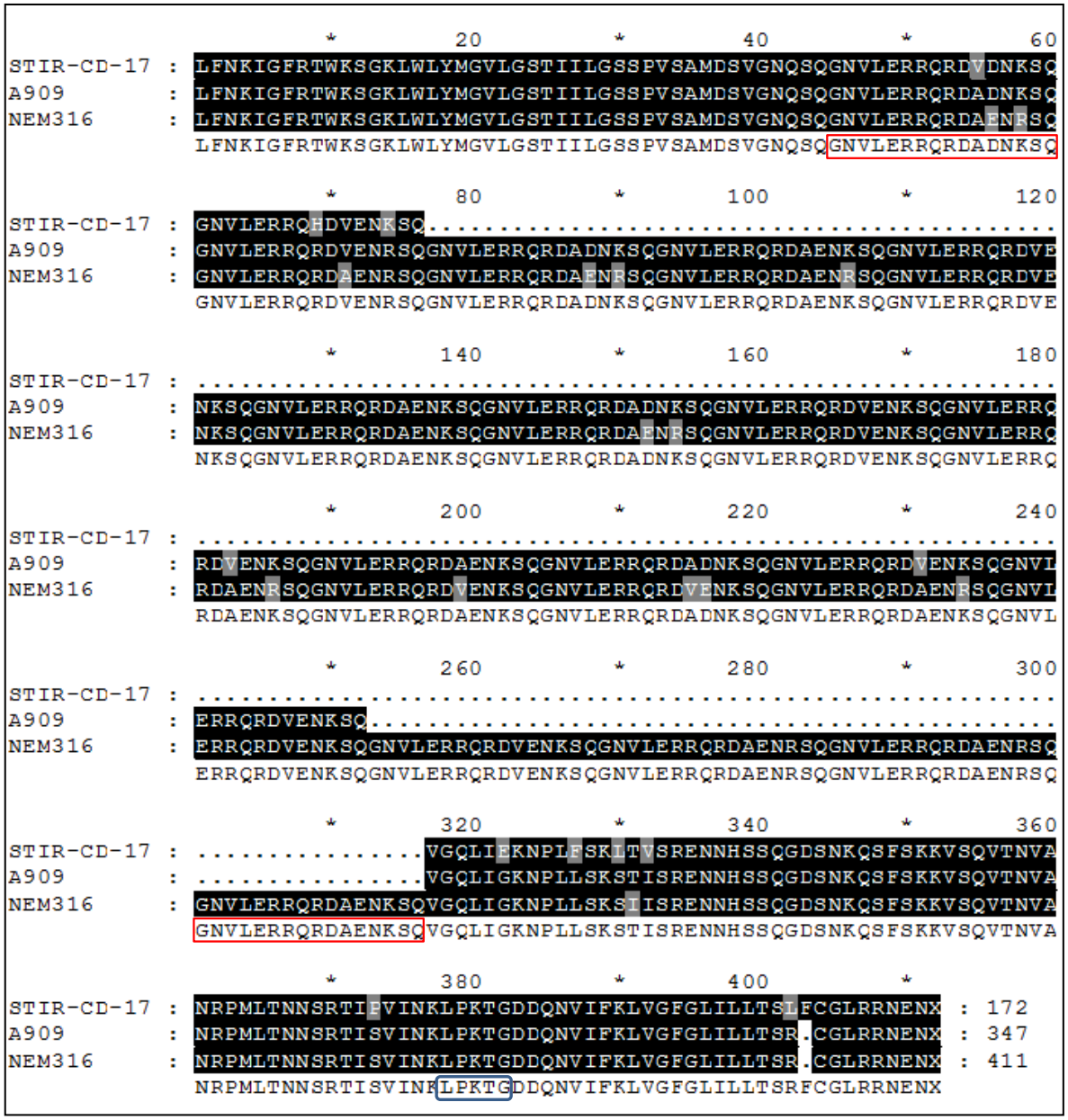


Figure 5.1 Sequence alignment of predicted FbsA amino acid sequences from STIR CD-17, reference strain A909 and reference strain NEM316. The fibrinogen-binding site, composed of repetitive units of 16 bases, extends from consensus positions 40 to 316 (the first and last repeat units from the consensus sequence are framed in red). For the reference strains, a total of 13 (A909) and 17 (NEM316) repeat units are found, whereas STIR-CD-17 has only 2 of these repeat segments. For all sequences, the N-terminal region comprises the secretion signal peptide sequence and the C-terminal region comprises the cell-wall anchoring region preceded by an LPXTG motif (framed in blue). Shading levels were set using Genedoc software (http://www.psc.edu/biome-d/genedoc/), with the grey colour highlighting single nucleotide differences. Dots correspond to missing nucleotides. The consensus sequence is represented in the bottomline of each block.

5.4.1.2 Fibronectin-binding protein

PavA was first identified in *S. pneumoniae* but homologues are found in a variety of Gram-positive bacteria, including *S. agalactiae* (Glaser et al., 2002; Henderson et al., 2011). Despite its lack of an N-terminal signal peptidase cleavage site and a C-terminal wall anchoring signature, it was shown by immune-electron microscopy that PavA is located on the outer cell surface of *S. pneumoniae* (Holmes et al., 2001). Pneumococcal PavA binds fibronectin, which is a major glycoprotein of the host extracellular matrix (Holmes et al., 2001). Experiments using knock-out mutants showed that PavA affects pneumococcal adherence and invasion of epithelial and endothelial cells *in vitro*, whereas *in vivo* experiments in mouse models revealed a severe decrease in the virulence of the corresponding mutants (Holmes et al., 2001; Pracht et al., 2005).

PavA is well conserved among *S. agalactiae,* including STIR-CD-17, with a calculated BSR > 0.99 between the protein in STIR-CD-17 and the proteins in NEM316, A909 or 2603V/R. However there are no reports on the role of the *pavA* gene in *S. agalactiae* and the expression, cellular location and role of the corresponding protein in human and piscine *S. agalactiae* requires further investigation.

5.4.1.3 Serine-rich repeat protein

Two variants of the serine-rich repeat protein Srr have been described in the literature, with the second variant (Srr2) mostly associated with strains from the hypervirulent ST17 clone (Brochet et al., 2006; Seifert et al., 2006). However, in STIR-CD-17 it is the variant Srr1 that is found.

In vitro and *in vivo* experiments using knock-out mutants of *S. agalactiae* of human origin revealed that Srr1 has an important role in the binding of *S. agalactiae* to human keratin 4 and fibrinogen (Samen et al., 2007; Seo et al., 2012), that it contributes to vaginal colonisation through host-cell attachment (Sheen et al., 2011), and that it facilitates penetration of the blood-brain-barrier (van Sorge et al., 2009; Seo et al., 2012). In a neonatal

sepsis model of infection in mice, strains possessing Srr1 were, however, found to be less virulent than those with Srr2 (Seifert et al., 2006).

Defining features of the protein Srr include an N-terminal signal peptide, one or two unique non repeat regions (NR domain), two serine-repeat regions (SRR1 and SRR2) of which one comprises the majority of the protein (SRR2), and a C-terminal LPXTG cell-wall anchoring motif (Seifert et al., 2006). The NR region mediates adhesion whereas the SRR2 domain has no adhesive properties and appears to be highly variable in size (Pyburn et al., 2011). It has been hypothesised that the SRR2 domain has evolved in order to extend the NR domain beyond the cell-wall and other surface components to facilitate attachment of the NR domain (Pyburn et al., 2011). The amino acids that alternate with the serine in the SRR regions of Srr1 are alanine, threonine and methionine (Seifert et al., 2006). In *S. agalactiae* A909, Srr1 is 1258 amino acids long with 280 repeats of the amino acid sequence S-A/T/M. In contrast, Srr1 in STIR-CD-17 is much smaller in size (730 amino acids; **Figure 5.2**) with an SRR2 domain containing only 16 S-A/T/M repeats. This observation suggests that cell-wall components may be thinner in STIR-CD-17 than in other *S. agalactiae* strains or that Srr1 is of limited functional relevance in STIR-CD-17.

Homologues of the gene exist in several streptococcal species (Lizcano et al., 2012) and the genes encoding for Srr1 or Srr2 are part of the so-called *sec*Y2/A2 locus. This locus includes 2 glycosyltransferases (*Gtf*A and *Gtf*B) involved in Srr protein glycosylation, which is required for protein stability, and 5 genes (*sec*Y2, *asp*1-3, *sec*A2) encoding components of the *secY2/A2* complex necessary for Srr transport (Mistou et al., 2009; Seifert et al., 2006). Species and strain-specific structural variations occur and can include additional glycosyltransferases (for example a total of 8 glycosyltransferases are observed in *S. agalactiae* A909; **Figure 2**). In STIR-CD-17, *secY2* is absent, suggesting that Srr1 might not be exported extracellularly. Disruption of *secY2* in strains of *S. gordonii* resulted in accumulation of the protein within the cytoplasm (Bensing and Sullam, 2002), but in

Mycobacterium smegmatis, *Mycobacterium tuberculosis* and *Listeria monocytogenes*, *SecA2*-only systems have been identified and reported to be functional through an as-yet undefined mechanism (Rigel and Braunstein, 2008). However, *secA2* appears as a pseudogene in STIR-CD-17 and it is therefore unlikely that the Srr1 protein is present on the surface of STIR-CD-17.

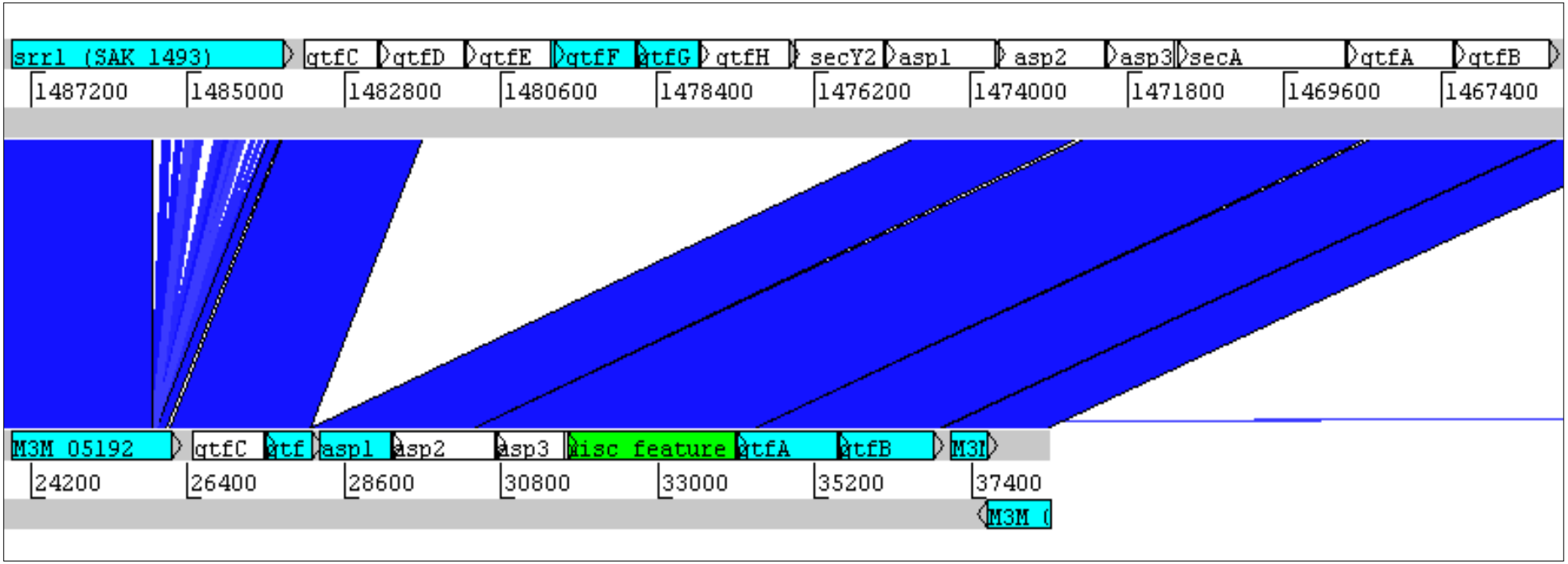


Figure 5.2 Characterisation of the *secY2/A2* locus in STIR-CD-17 by linear comparison between *S. agalactiae* A909 (top sequence) and Contig 622 from STIR CD-17 (bottom sequence), created using the Artemis Comparison Tool. A homologous block of genomic sequence (BLASTN matches) is indicated by blue bars between the sequences. In *S. agalactiae* A909, this locus includes 1 gene (*srr*1) encoding a serine-rich cell-wall anchored protein, 8 glycosyltransferases (*Gtf*A to *Gtf*H) involved in the Srr protein glycosylation, which is required for protein stability, and 5 genes (*sec*Y2, *asp*1-3, *sec*A2) encoding components of the *secY2/A2* complex and necessary for Srr1 transport. In comparison, in *S. agalactiae* STIR-CD-17, the locus lacks a 6,805 bp region comprising the genes *gtf*E-*gtf*H plus *secY2*; s*ecA2* appears as a pseudogene (green) due to potential frameshift, and the *srr1* gene (M3M_05192) is of reduced size at its C-terminal region.

5.4.1.4 Peptidoglycan anchored protein

Brochet et al. (2006) first described the presence of a putative peptidoglycan-anchored protein referred to as gbs2018 and sag2063 based on their locus tag in the published genomes of strains NEM316 and 2603V/R respectively. Through the exploration of the gene structure in a global collection of *S. agalactiae* of human and animal origin, a total of 7 allelic variants (*gbs2018-1* to *-6*, plus *sag2063*) were reported. *Gbs2018-4* and *-6* correspond to variants only identified in *S. agalactiae* from cattle and fish respectively, whereas other variants have been identified in *S. agalactiae* from humans and other host species (Brochet et al., 2006). A strong correlation exists between chimeric variants and the STs of the strains, with for example *gbs2018-3* found exclusively among the hypervirulent ST17 clone (Brochet et al., 2006; Santi et al, 2007; Lamy et al., 2006).

The *gbs2018* variants are present at the same genetic locus in all *S. agalactiae* (Santi et al, 2007). The resulting protein may either be surface anchored or truly extracellular, but not all strains demonstrate expression of the protein *in vitro* possibly due to regulatory differences (Santi et al., 2007 and 2009). Using deletion mutants of the reference strain 2603V/R (containing the *sag2063* variant), an *in vitro* study demonstrated the protein to have adhesive and antiphagocytic properties (Santi et al., 2007). Moreover, vaccination of mice with a recombinant derivative of *sag2063* conferred protection against bacteria expressing either the same variant or the *gbs2018-1* variant. However, evidence suggests that there is no cross protection between all variants found among human isolates (Santi et al., 2007). Work on *gbs2018-3* highlighted this variant to be essential for the hyper-virulence of ST17 clone as it is required for adhesion and translocation across the intestinal barrier and the blood brain barrier (Tazi et al., 2010). No work has been published on other variants.

Structurally, these proteins display similar N-terminal (secretion signal peptide) and C-terminal (LPXTG cell-wall anchoring motif) regions but diverge at their central domains. The structure of allelic variants found among human isolates has been described by different

research teams; each of those teams, however, attributed a different name to similar chimeric variants, therefore rendering the nomenclature somewhat confusing (**Table 5.2**). Based on multiple alignments, Lamy et al. (2006) defined the presence of 11 segments (S1 to S11). In that study, the boundaries of most segments were defined by blocks displaying identical nucleotide sequences, whereas the basis of identification for segments S5 and S7 relied on the fact that the segment encodes variable KPXX repeats. The structure of chimeric variants as reported by Lamy et al. (2006) exclusively focused on *sag2063* and *gbs2018-1* to *-3*; in the current study, the structural analysis is extended to include the variants identified by Brochet et al (2006) and the one identified in STIR-CD-17, consisting of *gbs2018-4* (strain 549.12 isolated from a cow and of ST247), *gbs2018-5* (CCH30 isolated from a human and of ST1) and *gbs2018-6* (strain 2.22 isolated from a fish and of ST246; STIR-CD-17 isolated from a fish and of ST260) (**Figure 5.3**). This was achieved using multiple alignments and dot plot analyses as described in **Figure 5.4**. The *gbs2018-5* variant, as deposited in GenBank by Brochet et al. (2006), was found to be only partially complete as analysis of this sequence revealed that it only comprises the 3' segments corresponding to S7 and S8, from which we cannot deduce whether it actually corresponds to a new variant. In the current study, a total of four new segments are reported among the *gbs2018* variants. As shown in **Figure 5.3**, each *gbs2018* variant contains at least one distinctive segment (S3 for *gbs2018-1*, S9 for *gbs2018-2*, S10 for *gbs2018-3*, S15 for *gbs2018-4* and S14 for *gbs2018-6*). Both sequences from fish (2.22 and STIR-CD-17) are identical (*gbs2018-6*) with the exception of S7 (3 repeats of KPXX for STIR-CD-17 and 2 repeats for 2.22). For this variant, most of the inner part of the protein is unique (S14), but interestingly one of its segments (S11) is also shared with the variant found among hypervirulent strains of ST17 (**Figure 5.3**).

The putative *gbs2018* encoding proteins each have a highly conserved region at their N and C termini: whilst this allows sequences from diverse strains to be aligned, the resulting phylogenetic tree as generated by Lamy et al. (2006) would not accurately reflect the

heterogeneity found among the different segments as these would not be aligned correctly. In our study, segments were extracted and aligned independently prior to concatenation and phylogenetic tree construction. Major clustering was therefore based on segments shared between all variants (S1 and S8), whereas other segments enabled the tree to be refined within the major clusters. Furthermore S7/S5 were removed from the analysis as the number of repeats appeared to correspond to an adaptive evolution specific for each isolate, not reflecting the overall picture. It is possible that S7/S5 have evolved in order to extend the N-terminal sequence beyond the cell-wall, as hypothesised for the repeat domain of the Srr1 protein (section 5.4.1.3). In the tree (**Figure 5.5**), each major cluster corresponds to a specific variant, whereas sub-clusters denote variations within a specific variant. Most importantly the analysis reveals that the *gbs2018-6* from fish and *gbs2018-4* from bovine are clustering with the *gbs2018-3* variant found among ST17 hypervirulent human strains.

The variant found among fish pathogenic strains possibly has an important role in virulence and could be an interesting target for further studies.

Table 5.2 Names attributed by different research teams to chimeric variants of the gene encoding a peptidoglycan-anchored protein originally called *gbs2018*/*sag2063*.

Reference strains		Nomenclature in different publications				Hosts	STs
Name	Accession N°	Brochet et al., 2006	Lamy et al., 2006	Santi et al., 2007	Tazi et al., 2010		
2603V/R	AE009948	SAG2063	Gbs2018 cluster A	BibA variant I	BibA	H, F	19, 110, 28
NEM316	NC_004368.1	Gbs2018-1	Gbs2018 cluster A	BibA variant II	BibA	H, B, C, F	23, 249, 253
NEM1010	AM183361	Gbs2018-2	Gbs2018 cluster B	BibA variant III	NA	H, B, C, F, R	1, 2, 3, 6, 7, 8, 9, 10, 12, 41, 50, 103, 196, 226, 248, 251
BM110	AM051291	Gbs2018-3	Gbs2018 cluster C	BibA variant IV	HvgA	H	17, 252
549.12	AM51292	Gbs2018-4	NA	NA	NA	B	61, 247
CCH330	AM051293	Gbs2018-5	NA	NA	NA	H, B	1, 8
2.22	AM051294	Gbs2018-6	NA	NA	NA	P	246

The host range and sequence type (ST) of strains found to carry each of the specific variant is provided. A reference strain and accession number is also provided for each variant. NA, Non Available, H. Human, B., Bovine, C., Canine, F., Feline, R. Rodent, P., Piscine; HvgA, Hypervirulent GBS adhesin; BibA, Group B *Streptococcus* immunogenic bacterial adhesin.

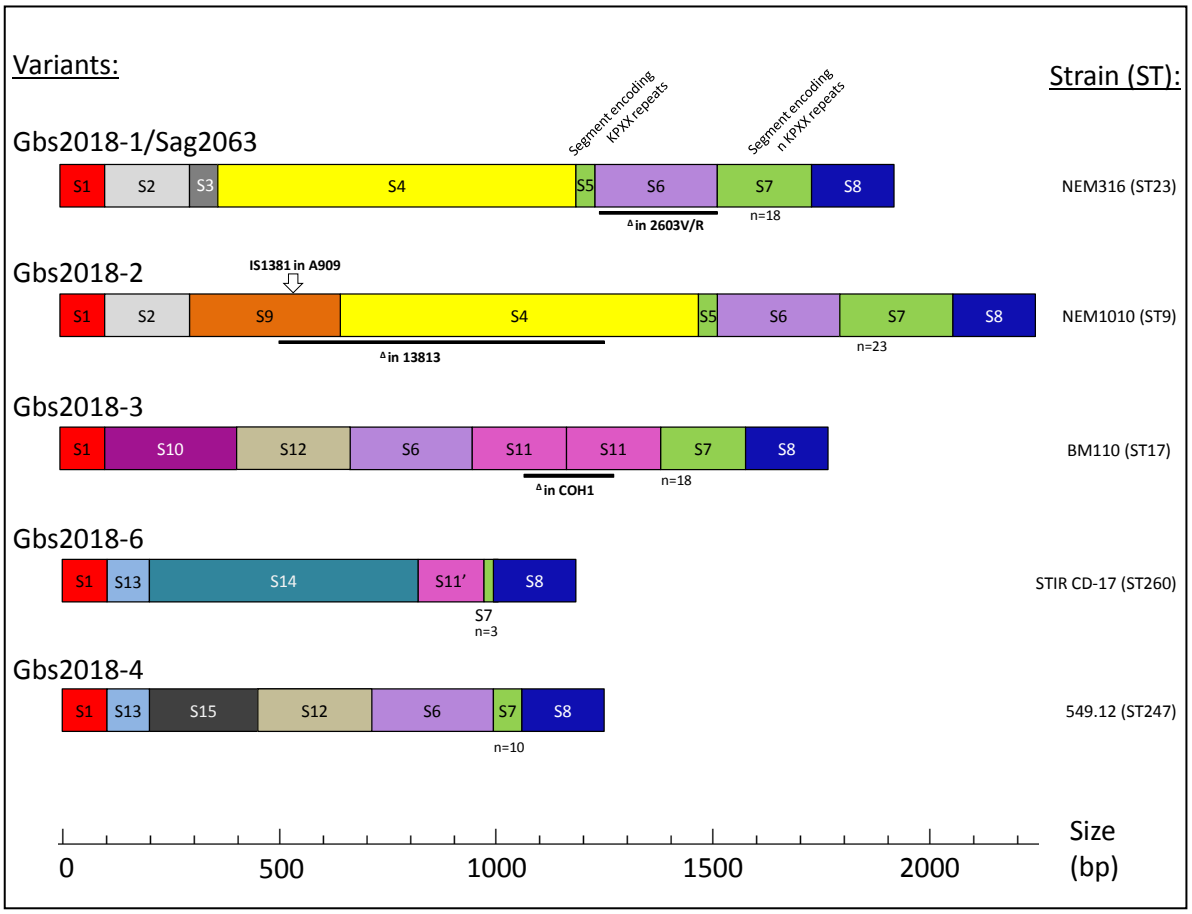


Figure 5.3 Mosaic structure of the gene encoding the 5 chimeric *gbs2018* variants encountered in *S. agalactiae* from humans (variants *gbs2018-1* to *-3*; after Lamy et al., 2006), bovines (variant *gbs2018-4*; mentioned by Brochet et al., 2006) and fish (variant *gbs2018-6* from genome STIR-CD-17 and mentioned by Brochet et al., 2006). The 5'-terminal segment S1 encodes for the secretion signal peptide whereas the 3'-terminal segment S8 contains the cell-wall anchor signature. The segments S5 and S7 contain variable numbers of KPXX repeats. Other segments are found in 1 to 4 of the gbs2018 variants. Distinction between *gbs2018-1* and *sag2063* is based on the presence or absence of S6. For variants *gbs2018-1* to *-3*, the segments are named as in Lamy et al. (2006) with adaptation: S10 as referred in Lamy et al. (2006) has been subdivided and referred to as S10 and S12 in the present figure. Structure of variants *gbs2018-4* and *-6* is based on analysis from the present study. The segment S11' identified in this study is identical to S11 with the exception of an internal deletion. Segments S12, S13, S14 and S15 were newly identified in the current study.

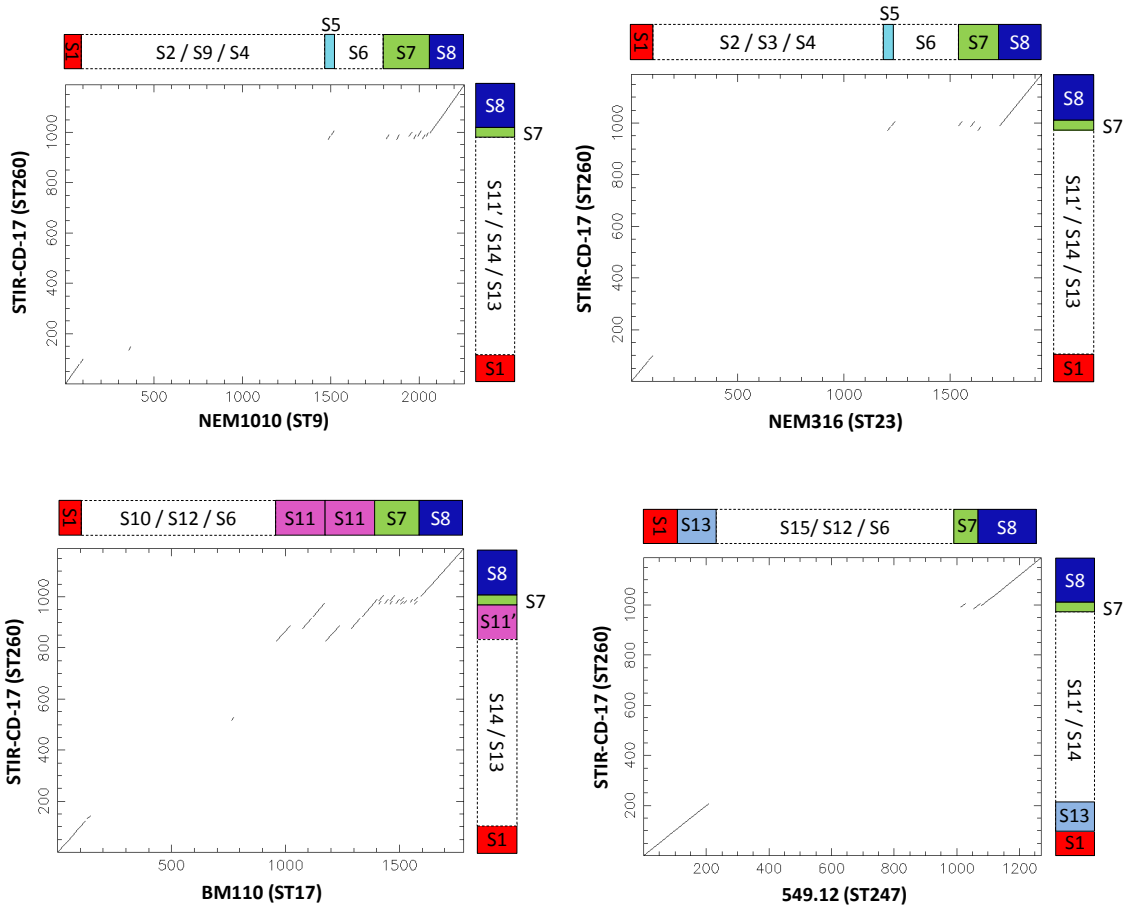


Figure 5.4 Pairwise comparisons of *gbs2018* from STIR-CD-17 against representatives of previously described *gbs2018* variants using dot-plots. Abscissas and ordinates indicate the size (bp) of the genes compared and lines indicate matching regions. The structure of the variants compared is presented along with the graphs, with matching segments represented by blocks framed with continuous lines, whereas non-matching segments are represented by block with dashed lines. Comparison of STIR-CD-17 with NEM1010 and NEM316 showed that all variants share the conserved N-terminal signal peptide region S1 (shown in red) and the conserved C-terminal regions S7 and S8 (shown in green and navy, respectively). Comparison with BM110 identifies homology of the S11 region, which is present with 2 copies in BM110 and a single copy with internal deletion in STIR-CD-17 (shown in pink). Finally, comparison with 549.12 shows homology of the S13 element (shown in light blue). The central segment of STIR-CD-17 did not show homology with any known sequence and is named S14.

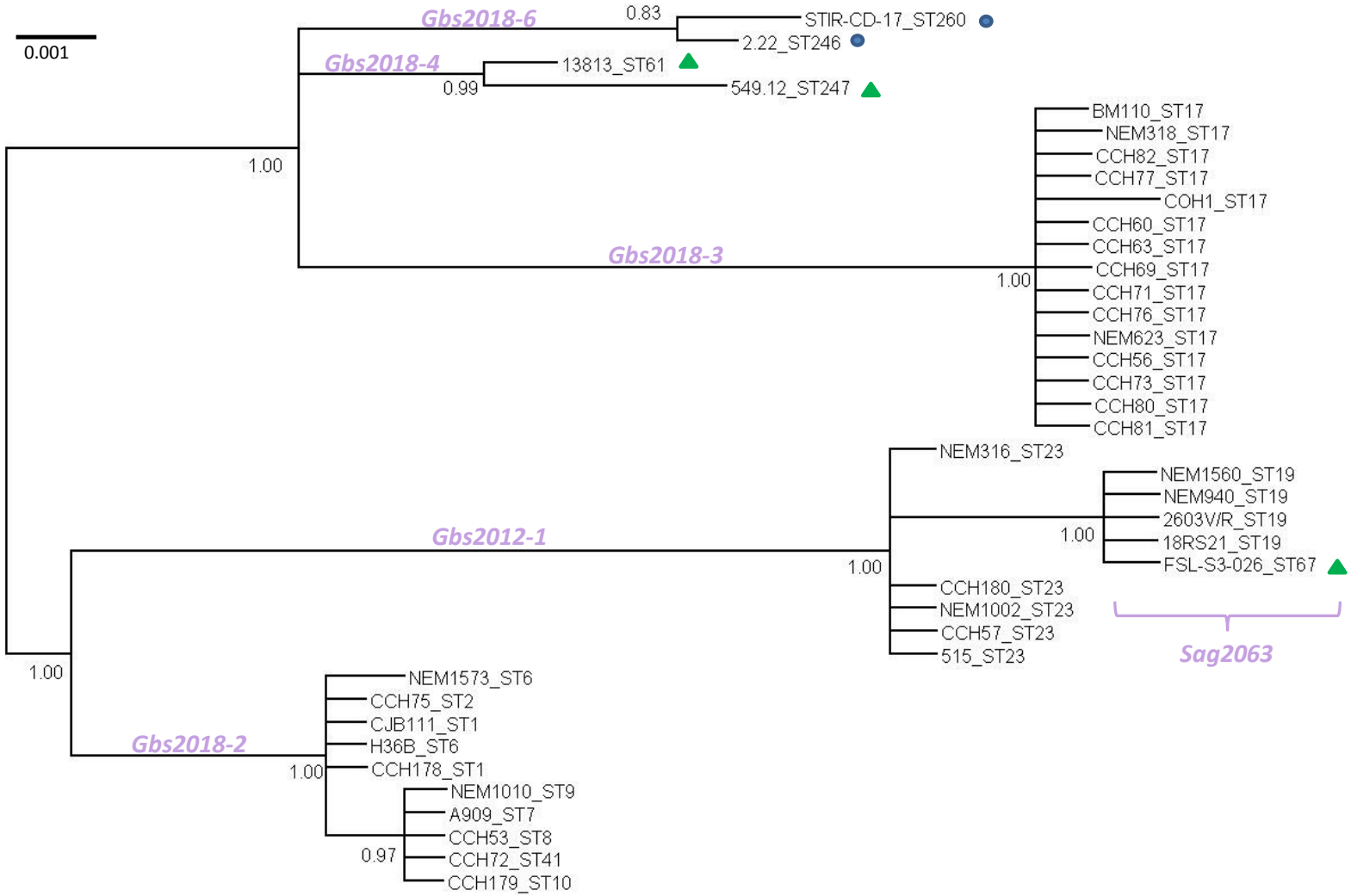


Figure 5.5 Bayesian phylogenetic tree of the *gbs2018* gene. Names of strains and corresponding sequence types (STs) are provided within the figure. Symbols adjacent to the strain information indicate the origin of the strain being fish (blue circles) or bovine (green triangle), whereas the absence of symbol indicates a human or unknown origin (NEM316). Major clusters are in agreement with the different variants reported in **Figure 5.3**. Posterior probabilities are shown at each node and the scale bar represents the substitutions per site.

5.4.1.5 Pili

Pili were unrecognized in decades of studies of *S. agalactiae* until comparative genomics enabled the identification of genes responsible for expression of these cell surface appendages (Lauer et al., 2005). Two distinct loci called PI-1 and PI-2 are responsible for the biosynthesis of pili in *S. agalactiae*, with PI-2 existing as 2 chimeric variants known as PI-2a and PI-2b (Rosini et al., 2006; review by Telford et al., 2006).

Both loci have a similar genetic organization; they are composed of 2 sortase C (SrtC) encoding-genes implicated in the assembly and polymerization of the pilus components, plus 3 structural protein encoding-genes with cell-wall anchoring motifs, namely the backbone protein (BP) and 2 ancillary proteins (AP1 and AP2). Cell-wall anchoring of pili is mediated by another sortase not encoded within the PI locus, the sortase A (SrtA) which covalently attaches pili to the cell-wall utilizing one of the AP as the anchor protein (Nobbs et al., 2008). Analysis of the distribution of pili within a collection of 289 human clinical isolates revealed that all strains carried at least one of these 2 loci and up to 94% of the strains expressed pili on their surface (Margarit et al., 2009).

Most studies to-date have focused on deciphering the role of pilus type 2a in pathogenesis. The AP1-2a (ancillary protein AP1 of pilus type 2a), has adhesion properties and the pilus type 2a variant has been implicated in mediating attachment to human respiratory epithelial cells (Dramsi et al., 2006, Konto-Ghioghi et al., 2009), human vaginal and cervical epithelial cells (Sheen et al., 2011) and to brain microvascular endothelial cells (Maisey et al., 2007). Pili are involved in biofilm formation (Konto-Ghioghi et al., 2009) and it was recently demonstrated that the biofilm-forming biotype is strictly conferred by pilus type 2a and not by pili of types 1 or 2b (Rinaudo et al., 2010). However conflicting results have been published concerning resistance to phagocytosis and resistance to intracellular killing by macrophages for pilus type 2a (Maisey et al., 2008; Papasergi et al., 2011), which suggests that not only pilus type but also level of expression and host or environmental

factors may influence pilus functions. Finally, *in vivo* challenges of 6 weeks old and newborn mice revealed that PI-2a is an important virulence factor but in the neonatal context only (Papasergi et al., 2011).

Studies on pilus type 1 and 2a suggested that these variants promote survival of *S. agalactiae* within macrophages (Jiang et al., 2012) and invasion of epithelial cells (Adderson et al., 2003; Krishnan et al., 2007). Conflicting results in terms of adhesion have, however, been obtained for the pilus type 1: using recombinant AP2-1 proteins, one study revealed that it acts as a ligand to specific tissue receptors, thereby promoting adherence to pulmonary epithelial cells (Krishnan et al., 2007), whereas another study using deletion mutants could not demonstrate that PI-1 pili mediate adhesion to respiratory epithelial, cervical or vaginal cells (Jiang et al., 2012).

Only one study evaluated the role of the pilus type 2b, which is the pilus type present in STIR-CD-17, and this *in vitro* study used knock out mutants to demonstrate that this variant promotes phagocytosis and survival of the bacteria within macrophages (Chattopadhyay et al., 2011).

STIR-CD-17 lacks the PI-1 locus. Strains with PI-2b but lacking PI-1 have been rarely reported in human *S. agalactiae*; of a collection of 289 isolates, only 1% was PI-2b positive and PI-1 negative (Margarit et al., 2009). In that study, the structure of the pilus islands was investigated and it was shown that PI-2b has a good degree of conservation for BP and AP2, whereas AP1 clusters into 2 allelic variants (1 variant represented by the reference strain COH1 and the other by A909). With STIR-CD-17, a third API-1 allele, closer to that of COH1 (ST17), was identified in this study (**Figure 5.6**). Moreover BP-2b is truncated and a sortase appears as a pseudogene due to potential frameshift causing the introduction of stop codons. Delta-BP knock-out mutants completely eliminate pilus polymerisation (Nobbs et al., 2008; Rosini et al., 2006), suggesting that BP is essential. Therefore no pili might be present on the surface of STIR-CD-17.

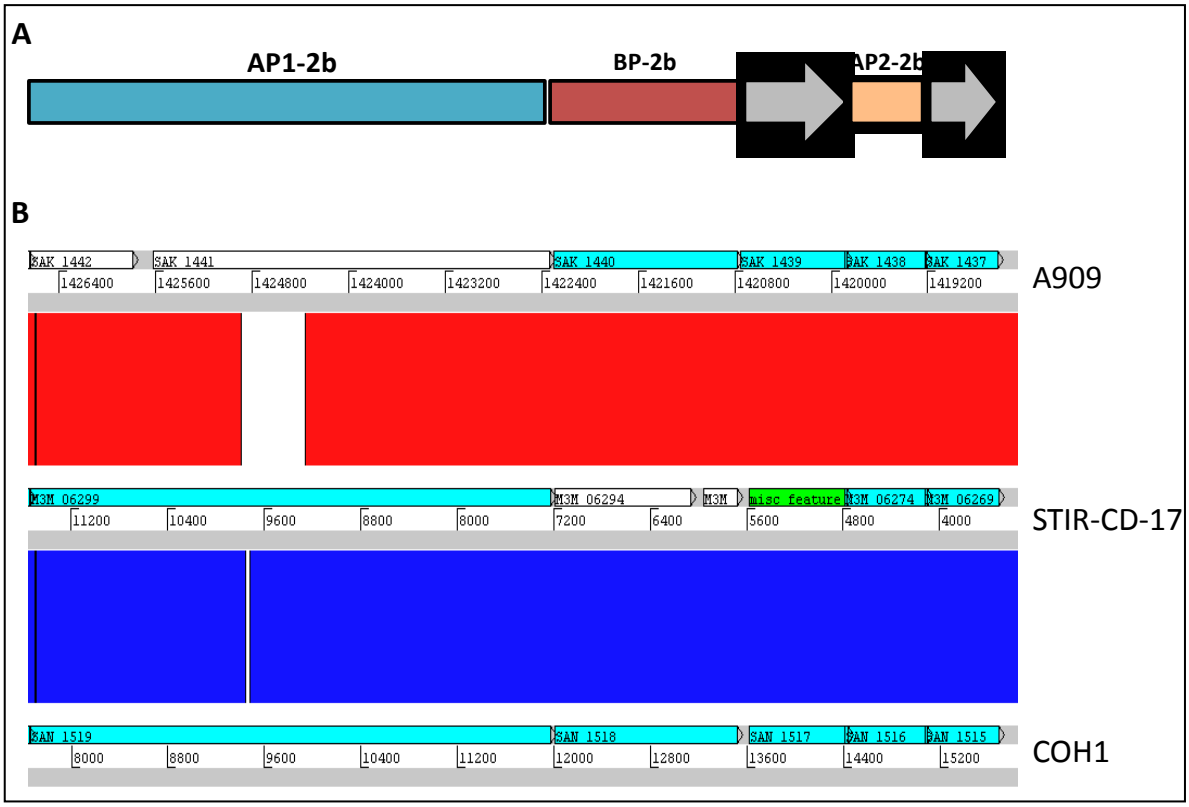


Figure 5.6 ***Streptococcus agalactiae*** **pilus type 2b.** ***Panel A:*** schematic representation of the genetic structure encoding for the pilus type 2b, which is composed of genes encoding a backbone protein (BP), 2 ancillary proteins (AP1 and AP2) and 2 sortases (grey arrows). ***Panel B:*** pairwise comparison of *S. agalactiae* STIR-CD-17 (middle sequence) with A909 (top sequence) or COH1 (bottom sequence) using the Artemis Comparison Tool. A homologous block of genomic sequence is indicated by blue or red bars between the sequences. The AP1 from STIR-CD-17 shows a region with non-homology with the AP1 from A909 or COH1, suggesting it is a new variant. In STIR-CD-17, the BP is truncated and one of the sortases is shown as a pseudoprotein (green), further suggesting that no pilus in expressed by STIR-CD-17**.**

5.4.2 Invasins

Among the 5 invasion genes investigated, 3 were found to be fully or partially present in the genome of STIR-CD-17. These include genes encoding β-haemolysin, CAMP factor and hyaluronate lyase HylB.

5.4.2.1 Beta-haemolysin/cytolysin

A hallmark phenotype for most human *S. agalactiae* is the appearance of a clearing zone surrounding colonies grown on blood agar, also called β-haemolysis. This phenotypic characteristic is due to a surface-associated toxin referred to as beta-hemolysin/cytolysin (βH/C) and capable of forming pores in red blood cell membranes (Marchlewicz and Duncan, 1981). The toxin is responsible for cytolytic and/or apoptotic injuries in a broad range of eukaryotic cells, including pneumocytes (Nizet et al., 1996), cardiomyocytes (Hensler et al., 2008), hepatocytes (Ring et al., 2002), neuronal cells (Reiß et al., 2011), brain and lung micro-vascular endothelial cells (Nizet et al., 1997; Gibson et al., 1999) and macrophages (Liu et al., 2004). *In vivo*, the βH/C was found to contribute to virulence in rabbit or murine models of arthritis (Puliti et al., 2000), pneumonia (Hensler et al., 2005), sepsis (Ring et al., 2002) and meningitis (Doran et al., 2003).

The toxin has not been purified to-date and there is therefore little information about its chemical structure. However, the genetic basis of the βH/C production has been identified to lie within the so called *cyl* operon (Spellerberg et al., 1999; Pritzlaff et al., 2001). This discovery enabled the generation of isogenic deficient mutants for investigation of the toxin function using *in vitro* and *in vivo* models. Within the operon, *cyl*E represents the structural gene for the toxin as the gene alone is able to confer the haemolytic capacity to *E. coli* (Pritzlaff et al., 2001); however, evidence suggests that other genes within the operon (*cyl*J and *cyl*K) are necessary for the full expression of the phenotype (Forquin et al., 2007).

The diagnosis of *S. agalactiae* infection in humans often includes observation of β-haemolysis as a phenotypic tool (Jones et al., 2003), and *cyl*B has recently been suggested as a real time PCR target for detection of *S. agalactiae* in clinical samples (de Zoysa et al., 2012). Non-haemolytic isolates, however, exist and their prevalence among carriage isolates is estimated between 5 to 8 %, although this value may be underestimated as most studies use β-haemolysis as diagnostic criterion for identification of *S. agalactiae* (Nickmans et al., 2012). Non-haemolytic isolates have been reported not only in humans, but also in bovines, mice, amphibians, shrimp and fish (Elliot et al., 1990; Duarte et al., 2004; Hasson et al., 2009). In fish, non-haemolytic *S. agalactiae* isolates are a frequent finding during outbreaks (Bowater et al., 2012; Mian et al., 2009; Chapter 2) and these isolates have previously been erroneously classified as a new bacterial species called *Streptococcus difficile* (Vandamme et al., 1997). The piscine STIR-CD-17 is also non-haemolytic, as are all of the isolates belonging to the fish-adapted CC552 (Chapter 2). Genomic analysis of the only non-piscine, non-haemolytic strain available (the bovine strain ATCC13813) revealed that the complete *cyl* operon is missing. In STIR-CD-17, the *cylA* and incomplete *cylB* plus *cylE* are present, whereas other genes from the operon are absent, explaining the non-haemolytic phenotype (**Figure 5.7**). The partial or complete loss of these genes in STIR-CD-17 and the absence of haemolysis in numerous fish pathogenic *S. agalactiae* demonstrate that the toxin is not an important contributor to pathgenicity in this host.

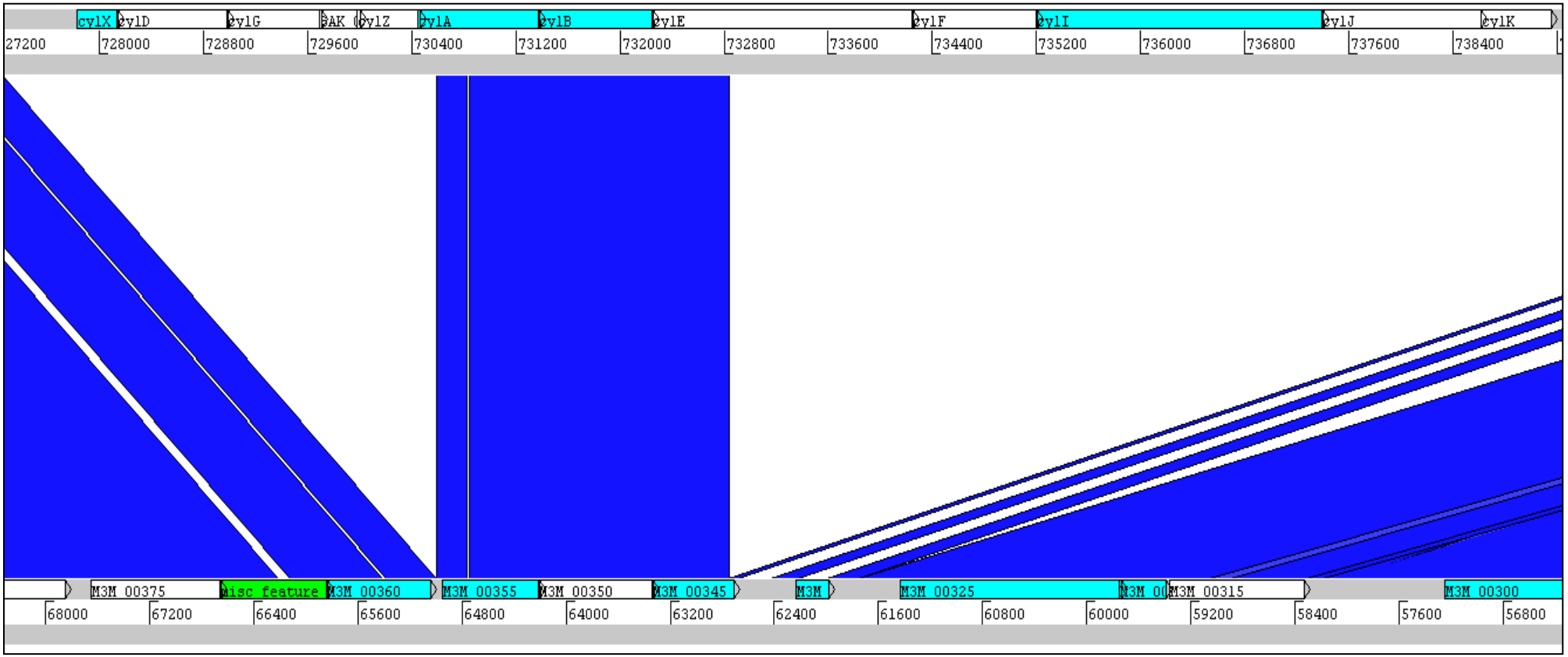


Figure 5.7 Linear comparison of the *cyl* operon (and adjacent genes) between the β-heamolytic *S. agalactiae* A909 (top sequence) and the non-heamolytic *S. agalactiae* STIR CD-17 (bottom sequence), created using the Artemis Comparison Tool. A homologous block of genomic sequence (BLASTN matches) is indicated by blue bars between the sequences. All genes from the *cyl* operon are present within *S. agalactiae* A909 (*cylX* to *cylK*), whereas most corresponding genes are absent from the *S. agalactiae* STIR CD-17, with the exception of *cylB* (M3M_00350; complete sequence) and partial sequences for *cylA* (M3M_00355) and *cylE* (M3M _00345).

5.4.2.2 CAMP-factor

Christie et al. (1944) first described a phenotypic reaction consisting of the lysis of red blood cells due to the synergystic interaction between an undetermined agent produced by *S. agalactiae* and the *Staphylococcus aureus* sphingomyelinase C (beta-toxin). The reaction and the *S. agalactiae* secreted agent are now commonly known as CAMP reaction and CAMP-factor respectively, after the discoverers initials (Christie, Atkins & Munch-Petersen). Evidence that the CAMP factor is an important virulent factor is based on two *in vivo* studies, one indicating that purified CAMP factor is lethal to mice and rabbits (Skalka and Smola, 1981), and the other revealing that sub-lethal doses of *S. agalactiae* injected along with the purified toxin induces fatal septicaemia in mice (Jürgens et al., 1987). It was also found that the CAMP factor binds to glycosylphosphotidylinosito-anchored proteins, suggesting that binding of the toxin to a cell receptor is necessary for its cell pore-forming action (Lang et al., 2007). More recently, the virulence properties of a CAMP-deficient mutant was evaluated and, based on an *in vivo* experiment, it was concluded that the mutant retains full virulence (Hensler et al., 2008). This new study is in contradiction with previous observations and suggests that the CAMP factor might therefore be nonessential for systemic virulence of *S. agalactiae*.

The complete 226 amino acid sequence of the CAMP factor was identified manually from peptide fragments in 1988 (Rühlmann et al., 1988), which enabled subsequent identification in 1994 of the corresponding gene named *cfb* (Podbielski et al., 1994). Homologues of this gene have been identified in other streptococcal species such as *S. pyogenes* (Gase et al., 1999), *S. uberis* (Jiang et al., 1996) and *S. canis* (Gürtük and Lämmler, 1990). The CAMP reaction and *cfb*-targeted PCRs have found their use as a diagnostic tool for *S. agalactiae* identification (Bae and Bottone, 1980; Ke et al., 2000), but strains that are phenotypically and/or PCR CAMP-negative have been reported on rare occasions in human and bovine isolates, suggesting that *cfb* is not always expressed and could be present in a mutant form or be absent from some strains (Hassan et al., 2000 and 2002; Kong et al., 2002; Podbielski

et al., 1994). Moreover non-haemolytic fish pathogenic isolates are commonly reported as phenotypically CAMP-negative (Bowater et al., 2012, Evans et al., 2008).

Along with many other reports of *S. agalactiae* from fish (Evans et al., 2007; Bowater et al., 2012), all piscine *S. agalactiae* belonging to the fish-adapted subgroup from our study were found to be CAMP-negative, including STIR-CD-17 (data not shown). However, the gene was found to be present within the STIR-CD-17 genome and it is well-conserved in comparison with the reference genomes (BSR > 0.97). The presence of the gene suggests that it is not expressed either due to *in vitro* conditions, possibly due to a repressor activation, or due to a gene defect along the pathway of expression. In order to investigate the latter hypothesis, the *CovSR* operon (Contig581, M3M03940 to M3M_03965) responsible for expression regulation of the CAMP gene was further investigated and compared to other sequences using ACT. The operon appeared to be well conserved. Within that operon, the *covR* gene is of particular interest, as *covR* deletion mutants are deficient for CAMP factor expression (Lamy et al., 2004). In STIR-CD-17, *covR* (M3M_03940) only shows 3 non-synonymous bp differences with the corresponding gene from NEM316 (gbs1672). This suggests that expression of the CAMP factor might not be due to a gene defect but to an undetermined repression under laboratory conditions used when growing the bacterium on blood agar. However, preliminary RT-PCR results suggest that expression in STIR-CD-17 takes place *in vitro* after a overnight growth in brain and heart infusion (data not shown). The overall good conservation of the gene within *S. agalaciae* species suggests that, from an evolutionary point of view, the toxin is beneficial to bacterial fitness in specific niches or for pathogenesis in hosts including fish. However, the actual function performed remains to be determined.

5.4.2.3 Hyaluronate lyase

Hyaluronate lyase is a secreted protease encoded by *hylB* that cleaves one of the major components of the extracellular matrix referred to as hyaluronan (Baker and Pritchard,

2000). Because of its enzymatic property, *hylB* is thought to be an important virulence factor contributing to host tissue invasion. Consistent with this hypothesis, higher levels of hyaluronate lyase production have been associated with strains causing invasive infections in humans (Milligan et al., 1978; Musser et al., 1989). However, the role of *hylB* in the course of infection has not been defined. Moreover, conflicting results were obtained when another study (Granlund et al., 1998) identified a group of invasive *S. agalactiae* strains with no hyaluronidase activity due to integration of IS1548 within the *hylB* gene. The role of this gene in *S. agalactiae* pathogenesis is therefore still under debate.

Structurally, *hylB* contains a promoter region, a ribosome binding site and a secretion signal peptide-encoding sequence (Gase et al., 1998). In STIR-CD-17, the gene is globally well conserved in comparison to the reference genomes (BSR > 0.9) and it does not contain any insertion sequence, suggesting that it could be expressed.

5.4.3 Immune evasins

Among the 7 putative evasins investigated, 4 were found to be fully or partially present in the genome of STIR-CD-17 (**Table 5.1**). These include genes encoding the enzymes responsible for production of the polysaccharide capsule, the serine protease CspA, the penicillin-binding protein 1A and superoxide dismutase.

5.4.3.1 Polysaccharide capsule

Virtually all *S. agalactiae* are considered to be encapsulated, even though a recent study revealed that some strains lack the capsular locus (Creti et al., 2012). The capsule is a major virulence factor formed by different arrangements of polysaccharides (including glucose, galactose and N-acetylneuraminic acid) organised in repeat unit structures (Cieslewicz et al., 2005). The coating of encapsulated bacteria by N-acetylneuraminic acid is considered to be an adaptative evolutionary feature of molecular mimicry, since the bacterial N-acetylneuraminic acid is identical to predominant sialic acids found on glycans of vertebrate

cells, thereby preventing the host from recognising the bacteria as non-self (Angata and Varki, 2002; Carlin et al., 2007; Severi et al., 2007). Consequently, mutants lacking sialic acid demonstrated a loss of virulence in a neonatal model of mice infection (Wessels et al., 1989).

Genes responsible for polysaccharide production and transportation are clustered in the so-called *cps* operon (*cpsA* to *cpsL*). Structural diversity of capsule architecture is determined by polymorphism of the *cps* operon leading to variation in serotypes (Cieslewicz et al., 2005). To date, ten capsular serotypes have been identified: Ia, Ib and II to IX (Slotved et al., 2007). In humans, ST17 is predominantly associated with serotype III, but this serotype has been identified in other STs and other host species, including ST283 in fish (Chapter 2) and ST23 in cattle (Sørensen et al., 2010). In cattle, many strains are non-typeable (Dogan et al., 2005). In fish, 3 distinct serotypes have been described: Ia, Ib and III (Evans et al., 2008, Suanyuk et al., 2008; Chapter 2).

In this study, STIR-CD-17 was found to be of serotype Ib by PCR (Chapter 2). In order to confirm this result with the genomic information, the *cps* locus was investigated and compared with the only 2 *cps*Ib sequences available from the NCBI website (accession numbers NZ_AAJS010 and AB050723). Specific *cps* gene organisations and homologies as described (Cieslewicz et al., 2005) confirmed STIR-CD-17 to be of serotype Ib, even though previously unreported gene polymorphisms were observed. Indeed, sequence comparison of *cps* loci revealed the presence of numerous single nucleotide differences specific for STIR-CD-17 as well as a non-homologous region of 218bp within the *cpsK* gene (**Figure 5.8**). The predicted CpsK from STIR-CD-17 and other isolates of serotype Ib however shows homology (BSR = 0.81), whereas the comparison with the predicted CpsK from the reference genomes shows divergence (BSC < 0.54), as expected, due to different serotypes. Whether the non-homologous region affects the biosynthesis of the capsule in serotype Ib remains to be investigated.

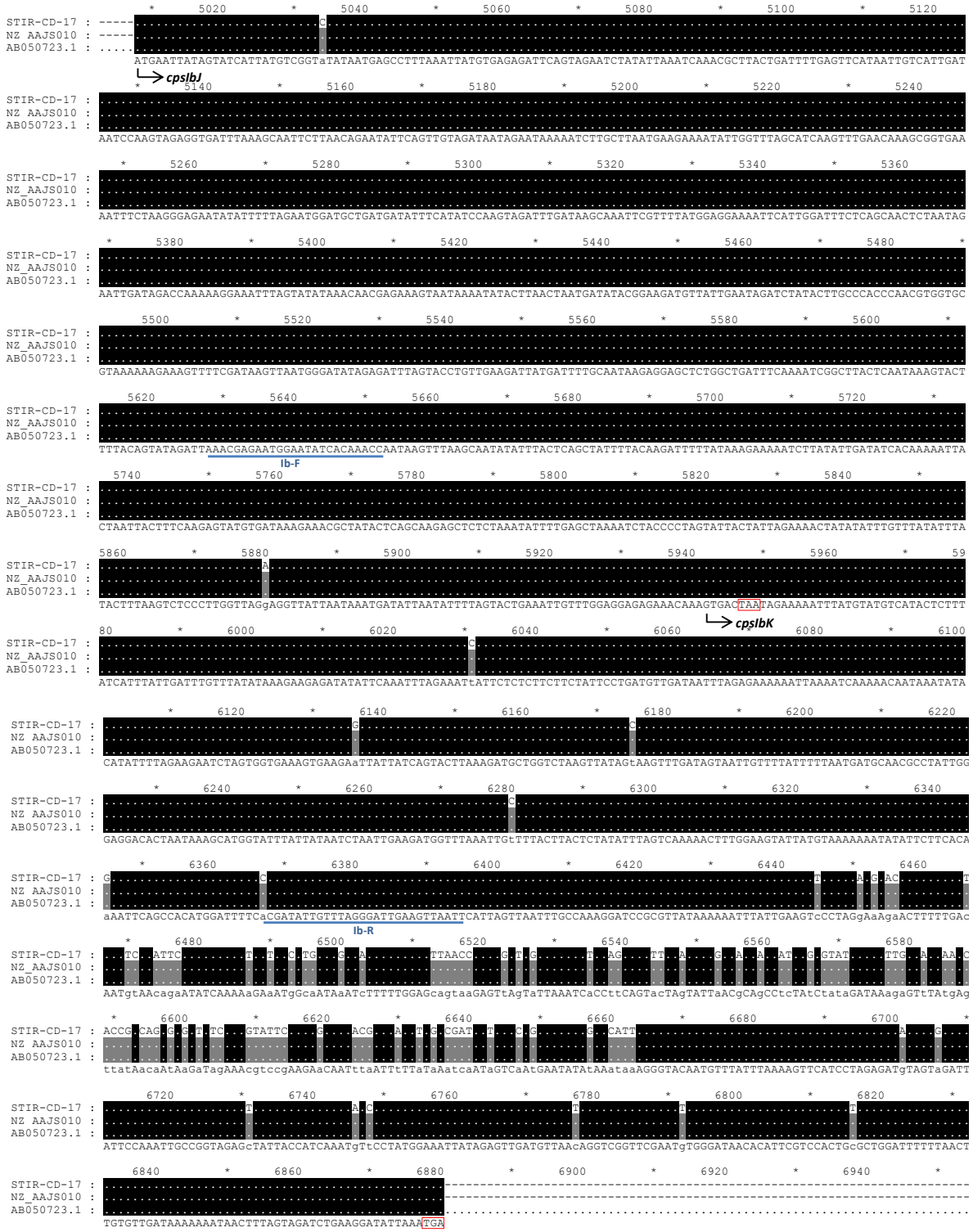


Figure 5.8 Sequence alignment of *cpsIbJ* and *cpsIbK* genes using sequence from isolate STIR CD-17 and sequences with accession number NZ_AAJS010 (H36B isolate) and AB050723 (undefined isolate). Specific primers annealing sites for Ib serotyping from Poyart et al. (2007) are shown in the figure (Ib-F, Ib-R). Stop codons are framed in red and start codons are indicated with arrows. Nucleotide position refers to the sequence with Genbank accession number AB050723. Dots indicate nucleotide sequences identical to that of the consensus sequence. Two shading levels are set using Genedoc software: black for 100% homology between the 3 sequences and grey for 2 sequences being homologous out of the 3.

Immediately downstream of the *cps* operon is the *neu* gene cluster (*neu* A to D) responsible for terminal sialysation of the polysaccharide capsule. This region is well conserved within STIR-CD-17 as compared with the reference genomes.

5.4.3.2 Penicillin-binding protein 1A

Since their discovery as targets of beta-lactam antimicrobials, the penicillin binding proteins (PBPs) have been the subject of intense research and have been identified as enzyme families involved in peptidoglycan biosynthesis (Georgopapadakou et al., 1980; Zapun et al., 2008). Bacterial species can express numerous PBPs with different and potentially redundant functions (Denome et al., 1999). One of those proteins, referred to as PBP1a and encoded by *pon*A, was shown to be required for virulence in a neonatal sepsis model of infection in rats using signature-tagged mutagenesis (Jones et al., 2000). Subsequently, it was found that *pon*A mutants were more susceptible to phagocytic killing and were cleared more rapidly after lung exposure in an *in vivo* aerosol model of rat pups infection (Jones et al., 2003 and 2007). One of the mechanisms by which PonA may act as an immune evasin was revealed when the protein was linked with increased resistance to the human antimicrobial peptides (AMPs) that are released by phagocytes for the killing of engulfed microorganisms (Hamilton et al., 2006). Other proteins have been described to confer AMP resistance through changes in surface charges (Peschel, 2002), but *pon*A mutants do not demonstrate any alteration of charge at their surface (Hamilton et al., 2006). It is therefore not known how the PBP1a protein confers resistance to AMPs.

In fish, AMPs constitute the first line of defense against pathogens and have been identified in tissues such as mucosal surfaces and skin (Cole et al. 1997), and in immune cells such as phagocytic granulocytes (Mulero et al., 2008) and eosinophilic granule cells (Silphaduang and Noga, 2001). Furthermore, the expression of an AMP with demonstrated antimicrobial properties has been shown to be up-regulated in the gill and spleen of *S. agalactiae*-infected tilapia (Peng et al., 2012). Antimicrobial peptides are thereby considered an important

component of the innate immune system and phagocytic killing in fish (Gómez & Balcázar, 2008). The *ponA* gene was identified in STIR-CD-17 and its corresponding protein is well conserved in comparison to reference genomes (BSR > 0.99). Thus, the ability of the bacterium to resist the action of piscine AMPs through *ponA* expression deserves further investigation.

5.4.3.3 Superoxide dismutase

The manganese-dependant superoxide dismutase (Mn-SodA or SodA) is a 202 aa long periplasmic enzyme that converts superoxide anions (O_2^-) into molecular oxygen (O_2) and hydrogen peroxide (H_2O_2), which in turn is further detoxified by catalases or peroxidases (Poyart et al., 2001). Superoxide anions are toxic compounds released by activated phagocytes intra- and extra-cellularly during the respiratory burst. Intraphagosomal release of O_2^- is regarded as a component in the killing of bacteria, whereas extracellular release of O_2^- is seen as a defence mechanism that kills foreign cells without phagocytosis and mediates the inflammatory process (Bannister et al., 1987). Consequently, SodA may constitute an important immune evasin for pathogenic bacteria. Using a mouse infection model, survival of a *sod*A knock-out mutant was severely impaired in the bloodstream and the brain but unchanged in the liver or the spleen (Poyart et al., 2001), suggesting that SodA confers survival benefits in specific host niches.

Fish macrophages and neutrophils are known to produce bactericidal superoxide anions during the respiratory burst (Gómez & Balcázar, 2008). In STIR-CD-17, the predicted SodA is well conserved in comparison with the reference genomes (BSR > 0.99), shedding light on a putative immune evasin of importance in fish infections.

5.4.3.4 Serine proteinase

The cell-surface associated protein A (CspA), encoded by the *csp*A gene, was initially described by Harris et al. (2003) as a putative cell-wall serine protease capable of cleaving human fibrinogen into fibrin-like compounds. Coating of *Streptococcus agalactiae* by

fibrin-like substances might then prevent recognition of the bacterium by the host immune system. In accordance with this assumption, inactivation of *cspA* decreased *S. agalactiae* virulence and resistance to opsonophagocytic killing by neutrophils in a neonatal rat model of sepsis (Harris et al., 2003). CspA also has the ability to cleave and inactivate numerous chemokines that act on neutrophil activation (Bryan and Shelver, 2009).

Structurally, CspA is composed of an N-terminal secretion signal peptide, a protease domain, a domain A, and a C-terminal LPXTG cell-wall anchoring motif. In STIR-CD-17, no gene corresponding to *cspA* has been annotated through the NCBI pipeline but it is mentioned as a pseudogene due to a potential frameshift. Sequence comparison using ACT revealed a deletion sequence of 352 bases within CspA in STIR-CD-17, which is responsible for the frameshift. However, two open reading frames (ORFs) can still be detected: ORF-A contains the signal peptide encoding domain, the protease domain and the domain A of CspA, whereas ORF-B contains the cell wall anchor domain (**Figure 5.9**). In the event that ORF-A is expressed, which has yet to be determined, these observations suggest that CspA cannot be a cell-wall anchored protein but instead would be a secreted protein in STIR-CD-17. This possibility is further supported by examples from the literature. In the case of another streptococcal species which has undergone genome reduction, namely *Streptococcus equi*, frameshift mutations that truncated cell-wall proteins have also been described and the secreted products found to remain active (Holden et al., 2009).

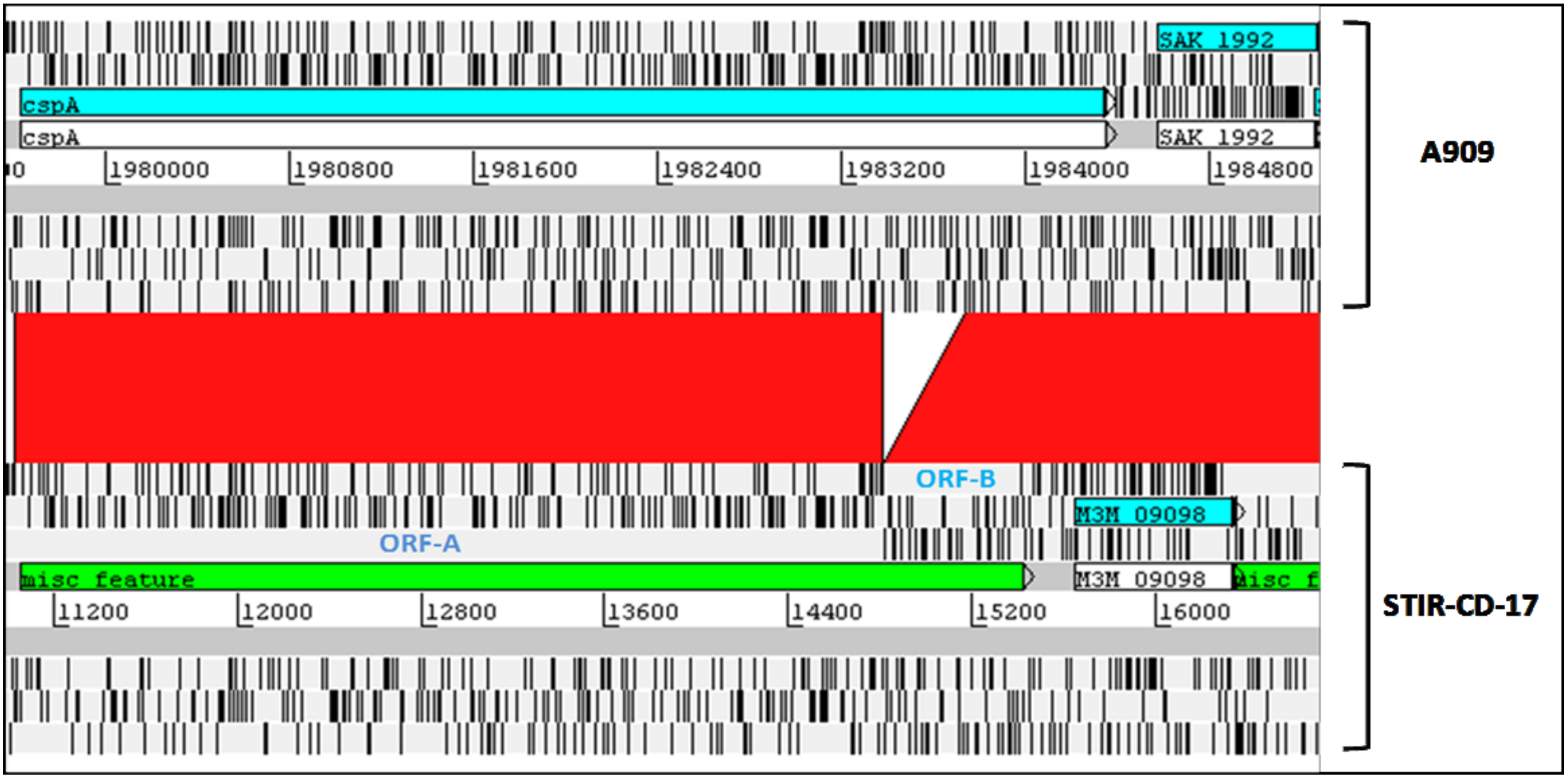


Figure 5.9 Investigation of the *CspA* locus by linear comparison between *S. agalactiae* A909 (top) and STIR-CD-17 (bottom) using the Artemis Comparison Tool. The 3 reading frames and gene positions for both reverse and forward sequences are presented. Bars within reading frames denote stop codons. A homologous block of genomic sequence (BLASTN matches) is indicated by red bars between the sequences. This shows that the gene has been annotated as a pseudogene due to a deletion responsible for a frameshift. However 2 open reading frames (ORFA and ORFB) can be distinguished. The ORFA comprises most of the gene, including the signal peptide encoding motif, suggesting that it could be expressed as a secreted protein.

5.4.4 Virulence genes absent from STIR-CD-17

Numerous virulence genes present in other *S. agalactiae* were absent from the STIR-CD-17 genome and are therefore non essential for pathogenesis in fish.

In the case of the fibrinogen binding protein FbsB, encoded by the *fbsB* gene, most of the corresponding sequence in STIR-CD-17 has been replaced by a non-homologous sequence with no conclusive blastN hit within the GenBank repository (**Figure 5.10**). In other cases, the genes are part of putative pathogenicity islands or mobile genetic elements found to be absent from the STIR-CD-17 genome: this includes the pilus type 1, the alpha-like proteins encoding genes *rib* and *bca*, the laminin binding protein encoding gene *lmb*, the C5a-peptidase encoding gene *scp*B, and the *bac* gene encoding for the β-C protein (**Figure 5.11**).

In some of these cases, the corresponding region is occupied by putative fish-associated genes: Locus 5 (Chapter 4, section 4.4.4.5) is found in the place of the PI-1 coding genes, whereas Locus 2 (Chapter 4, section 4.4.4.2) is found in the place of *rib* and *bca*.

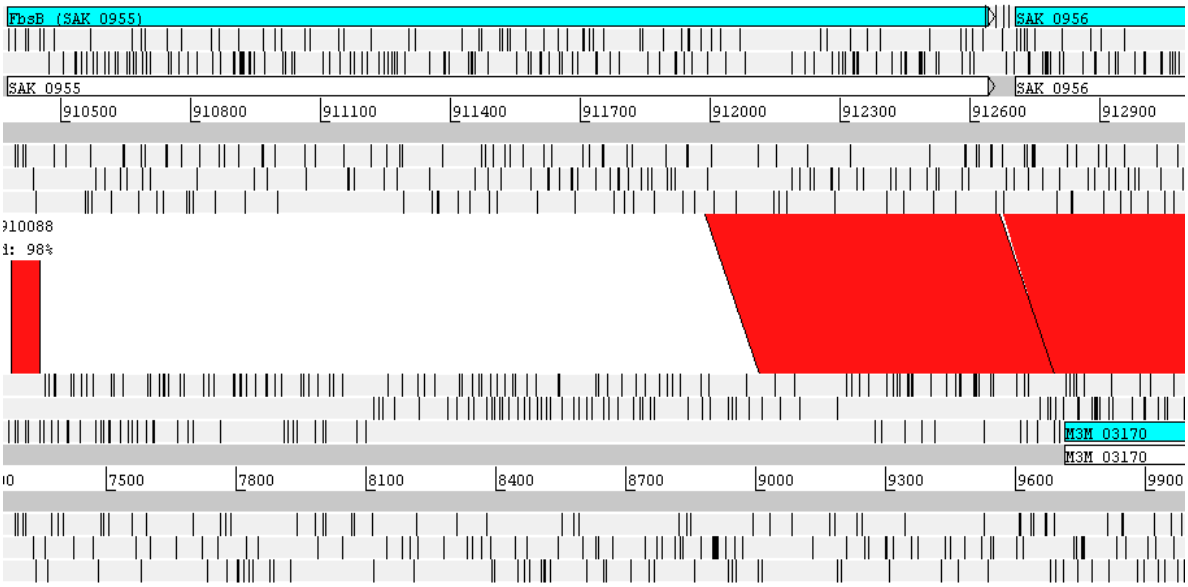


Figure 5.10 Investigation of the *fbsB* gene by linear comparison between *S. agalactiae* A909 (top sequence) and Contig 531 from STIR CD-17 (bottom sequence) created using the Artemis Comparison Tool. The 3 reading frames and gene positions for both reverse and forward sequences are represented. Bars within reading frames denote stop codons. A homologous block of genomic sequence (BLASTN matches) is indicated by red bars between the sequences. The *fbsB* gene within *S. agalactiae* A909 is annotated as SAK_0955. The *fbsB* gene within STIR-CD-17 is absent with the corresponding region mostly replaced by a non-homologous sequence that shows no significant blastN hit within the NCBI website.

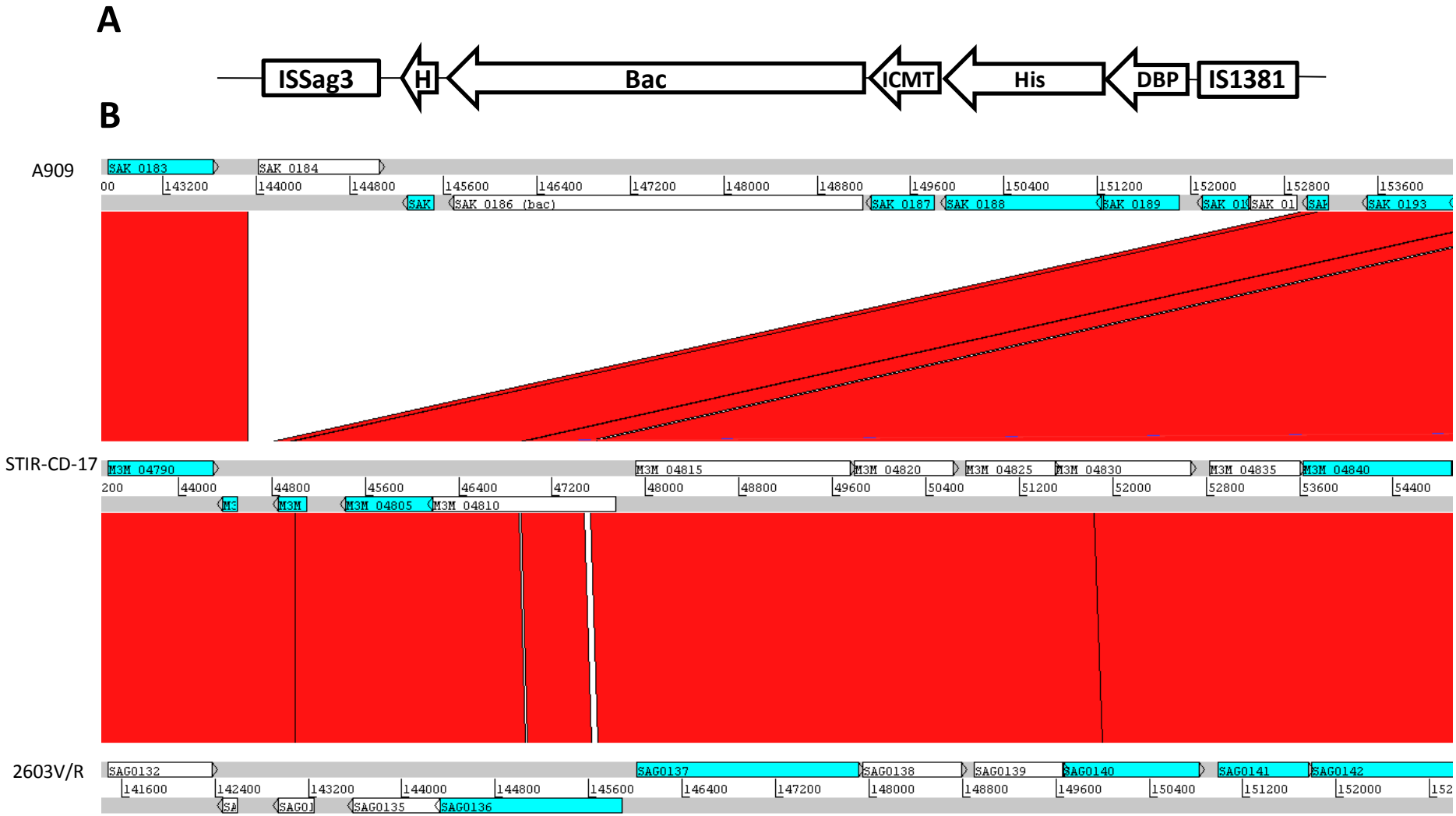


Figure 5.11 Analysis of the genomic region comprising and flanking *bac*. *Panel A:* in A909, *bac* is part of a putative composite transposon of 9kB flanked by the insertion sequences ISSag3 and IS1381. ***Panel B:*** based on ACT comparison, the corresponding region appears to be absent from STIR-CD-17 and 2603V/R. H. Hypothetical protein, ICMT. isoprenylcysteine carboxyl methyltransferase, His. histidine kinase, DBP. DNA binding response regulator.

5.5 CONCLUSIONS

Numerous genes encoding recognised virulence factors in *S. agalactiae* from humans were absent in STIR-CD-17, including genes encoding putative adhesins (pilus type 1, fibrinogen binding protein FbsB, laminin-binding protein), invasins (surface protein Rib, C-α protein), or evasins (C5a peptidase, C-β protein). Moreover, a cluster of genes encoding a virulence determinant important for pathogenicity in humans *S. agalactiae* was found to be only partially present, *i.e.* the *cyl* operon responsible for β-haemolysin production. Genes encoding the β-haemolysin therefore appear to have originally been present within the STIR-CD-17 genome, but their function has been lost, possibly during the niche-reduction process (Chapter 4). The complete or partial absence of all these genes indicates that their encoded virulence determinants are non-essential to the fish-adapted subpopulation of *S. agalactiae*.

Virulence factors may be present as different variants among distinct *S. agalactiae* strains, as described for the pilus type 2 (variants 2a and 2b), the gbs2018 (variants 1 to 6), and the serine-rich repeat protein Srr (variant 1 and 2). In STIR-CD-17, genes encoding the serine-rich repeat protein 1, the pilus type 2b and the gbs2018 variant 6 were identified. Genomic evidence, however, suggested that no pilus might be present on the surface of STIR-CD-17 due to frameshifts. Srr1 is also found in *S. agalactiae* from humans, but the gbs2018-6 variant identified in STIR-CD-17 has uniquely been identified among isolates belonging to the fish-adapted subpopulation. The existence of distinct variants among different bacterial clones suggests differences or adaptations in the function of the protein. Based on the importance of the gbs2018 variants studied in *S. agalactiae* from humans (Tazi et al., 2010), it can be hypothesised that gbs2018-6 enhances the fitness of isolates in fish, but further work is required to investigate expression and function of this variant. Furthermore, a set of genes encoding recognised virulence factors in *S. agalactiae* from humans were found to be well-conserved in STIR-CD-17, including *fbsA*, *pavA*, *srr1*, *cfb*, *hylB*, *ponA* and *sodA*. The *cspA* gene, encoding a serine protease, was also found to be well-conserved, but a deletion

responsible for a frameshift suggested that, if the gene is expressed, the protein would be secreted and not cell-wall anchored. The *cps* genes were also found to be well-conserved, with the exception of *cpsK*, but whether the variations in *cpsK* affect the biosynthesis of the capsule is unknown at the present time.

Two complete genomes of *S. agalactiae* of piscine origin have been recently deposited in GenBank (after May 2012) and include the SA20-06 isolate belonging to the fish-adapted subgroup of *S. agalactiae* (CC552), and the GD201008-001 isolate belonging to the human-associated subpopulation (CC7). Because of time constraints, the analysis presented here was not updated with these genomes. However, preliminary analyses revealed that all observations of absence, presence or variations of genes as made for STIR-CD-17 are consistently found in SA20-06, indicating that these observations are conserved among the fish-adapted subgroup of *S. agalactiae*. This work provides a basis for further post-genomic studies that may include, for example, expression studies and immunogold labelling for assessment of capsular localisation. Generation of knockout mutants and *in vivo* studies are, however, required to further confirm the implication of the corresponding proteins in virulence according to the tenet of the molecular version of Koch's postulate (Falkow, 1988).

5.6 REFERENCES

Adderson, E.E., Takahashi, S., Wang, Y., Armstrong, J., Miller, D.V., Bohnsack, J.F., 2003. Subtractive hybridization identifies a novel predicted protein mediating epithelial cell invasion by virulent serotype III group B *Streptococcus agalactiae*. Infect. Immun. 71, 6857-6863.

Angata, T. and Varki, A., 2002. Chemical diversity in the sialic acids and related a-keto acids: An evolutionary perspective. Chemical Reviews-Columbus 102, 439-470.

Bae, B.H.C. and Bottone, E.J., 1980. Modified Christie-Atkins-Munch-petersen (CAMP) test for direct identification of hemolytic and nonhemolytic group B streptococci on primary plating. Can. J. Microbiol. 26, 539-542.

Baker, J.R. and Pritchard, D.G., 2000. Action pattern and substrate specificity of the hyaluronan lyase from group B streptococci. Biochem. J. 348, 465

Bannister, J.V., Bannister, W.H., Rotilio, G., 1987. Aspects of the structure, function, and applications of superoxide dismutase. Crit. Rev. Biochem. Mol. Biol. 22, 111-180.

Bensing, B.A. and Sullam, P.M., 2002. An accessory sec locus of *Streptococcus gordonii* is required for export of the surface protein GspB and for normal levels of binding to human platelets. Mol. Microbiol. 44, 1081-1094.

Bisharat, N., Crook, D.W., Leigh, J., Harding, R.M., Ward, P.N., Coffey, T.J., Maiden, M.C., Peto, T., Jones, N., 2004. Hyperinvasive neonatal group B streptococcus has arisen from a bovine ancestor. J. Clin. Microbiol. 42, 2161-2167.

Bowater, R.O., Forbes-Faulkner, J., Anderson, I.G., Condon, K., Robinson, B., Kong, F., Gilbert, G.L., Reynolds, A., Hyland, S., McPherson, G., Brien, J.O., Blyde, D., 2012. Natural outbreak of *Streptococcus agalactiae* (GBS) infection in wild giant Queensland grouper, *Epinephelus lanceolatus* (Bloch), and other wild fish in northern Queensland, Australia. J. Fish Dis. 35, 173-186.

Brochet, M., Couvé, E., Zouine, M., Vallaeys, T., Rusniok, C., Lamy, M., Buchrieser, C., Trieu-Cuot, P., Kunst, F., Poyart, C., Glaser, P., 2006. Genomic diversity and evolution within the species *Streptococcus agalactiae*. Microb. Infect. 8, 1227-1243.

Bryan, J.D. and Shelver, D.W., March 15, 2009. *Streptococcus agalactiae* CspA is a serine protease that inactivates chemokines. J. Bacteriol. 191, 1847-1854.

Carlin, A.F., Lewis, A.L., Varki, A., Nizet, V., February 15, 2007. Group B streptococcal capsular sialic acids interact with siglecs (immunoglobulin-like lectins) on human leukocytes. J. Bacteriol. 189, 1231-1237.

Christie, K., Atkins, N.E., Munch-Petersen, E., 1944. A note on a lytic phenomenon shown by group b streptococci. Aust J Exp Biol Med 22, 197-200.

Cieslewicz, M.J., Chaffin, D., Glusman, G., Kasper, D., Madan, A., Rodrigues, S., Fahey, J., Wessels, M.R., Rubens, C.E., 2005. Structural and genetic diversity of group B streptococcus capsular polysaccharides. Infect. Immun. 73, 3096-3103.

Clamp, M., Cuff, J., Searle, S.M., Barton, G.J., 2004. The jalview java alignment editor. Bioinformatics 20, 426-427.

Cole, A.M., Weis, P., Diamond, G., 1997. Isolation and characterization of pleurocidin, an antimicrobial peptide in the skin secretions of winter flounder. J. Biol. Chem. 272, 12008-12013.

Creti, R., Imperi, M., Pataracchia, M., Alfarone, G., Recchia, S., Baldassarri, L., 2012. Identification and molecular characterization of a *S. agalactiae* strain lacking the capsular locus. Eur. J. Clin. Microbiol. Infect. Dis. 31, 233-235.

de Zoysa, A., Edwards, K., Gharbia, S., Underwood, A., Charlett, A., Efstratiou, A., 2012. Non-culture detection of *Streptococcus agalactiae* (lancefield group B streptococcus) in clinical samples by real-time PCR. J. Med. Microbiol. 61, 1086-1090.

Denome, S.A., Elf, P.K., Henderson, T.A., Nelson, D.E., Young, K.D., 1999. *Escherichia coli* mutants lacking all possible combinations of eight penicillin binding proteins: Viability, characteristics, and implications for peptidoglycan synthesis. J. Bacteriol. 181, 3981-3993.

Dogan, B., Schukken, Y.H., Santisteban, C., Boor, K.J., December 2005. Distribution of serotypes and antimicrobial resistance genes among *Streptococcus agalactiae* isolates from bovine and human hosts. J. Clin. Microbiol. 43, 5899-5906.

Doran, K.S. and Nizet, V., 2004. Molecular pathogenesis of neonatal group B streptococcal infection: No longer in its infancy. Mol. Microbiol. 54, 23-31.

Doran, K.S., Liu, G.Y., Nizet, V., 2003. Group B streptococcal β-hemolysin/cytolysin activates neutrophil signaling pathways in brain endothelium and contributes to development of meningitis. J. Clin. Invest. 112, 736-744.

Dramsi, S., Caliot, E., Bonne, I., Guadagnini, S., Prévost, M., Kojadinovic, M., Lalioui, L., Poyart, C., Trieu-Cuot, P., 2006. Assembly and role of pili in group B streptococci. Mol. Microbiol. 60, 1401-1413.

Duarte, R.S., Miranda, O.P., Bellei, B.C., Brito, M.A.V.P., Teixeira, L.M., 2004. Phenotypic and molecular characteristics of *Streptococcus agalactiae* isolates recovered from milk of dairy cows in Brazil. J. Clin. Microbiol. 42, 4214-4222.

Elliott, J.A., Facklam, R.R., Richter, C.B., 1990. Whole-cell protein patterns of nonhemolytic group B, type Ib, streptococci isolated from humans, mice, cattle, frogs, and fish. J. Clin. Microbiol. 28, 628-630.

Evans, J.J., Bohnsack, J.F., Klesius, P.H., Whiting, A.A., Garcia, J.C., Shoemaker, C.A., Takahashi, S., 2008. Phylogenetic relationships among *Streptococcus agalactiae* isolated from piscine, dolphin, bovine and human sources: A dolphin and piscine lineage associated with a fish epidemic in Kuwait is also associated with human neonatal infections in Japan. J. Med. Microbiol. 57, 1369-1376.

Falkow, S., 1988. Molecular Koch's postulates applied to microbial pathogenicity. Rev. Infect. Dis. 10 Suppl 2, S274-6.

Forquin, M., Tazi, A., Rosa-Fraile, M., Poyart, C., Trieu-Cuot, P., Dramsi, S., April 2007. The putative glycosyltransferase-encoding gene *cylJ* and the group B streptococcus (GBS)-specific gene *cylK* modulate hemolysin production and virulence of GBS. Infect. Immun. 75, 2063-2066.

Fredericks, D.N. and Relman, D.A., 1996. Sequence-based identification of microbial pathogens: A reconsideration of Koch's postulates. Clin. Microbiol. Rev. 9, 18-33.

Garcia, J.C., Klesius, P.H., Evans, J.J., Shoemaker, C.A., 2008. Non-infectivity of cattle *Streptococcus agalactiae* in Nile tilapia, *Oreochromis niloticus* and channel catfish, *Ictalurus punctatus*. Aquaculture 281, 151-154.

Gase, K., Ozegowski, J., Malke, H., 1998. The *Streptococcus agalactiae hylB* gene encoding hyaluronate lyase: Completion of the sequence and expression analysis. Biochimica Et Biophysica Acta (BBA) - Gene Structure and Expression 1398, 86-98.

Gase, K., Ferretti, J.J., Primeaux, C., McShan, W.M., 1999. Identification, cloning, and expression of the CAMP factor gene (*cfa*) of group A streptococci. Infect. Immun. 67, 4725-4731

Georgopapadakou, N.H. and Liu, F.Y., 1980. Penicillin-binding proteins in bacteria. Antimicrob. Agents Chemother. 18, 148-157.

Gibson, R.L., Nizet, V., Rubens, C.E., 1999. Group B streptococcal beta-hemolysin promotes injury of lung microvascular endothelial cells. Pediatr. Res. 45, 626-634.

Glaser, P., Rusniok, C., Buchrieser, C., Chevalier, F., Frangeul, L., Msadek, T., Zouine, M., Couve, E., Lalioui, L., Poyart, C., Trieu-Cuot, P., Kunst, F., 2002. Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. Mol. Microbiol. 45, 1499-1513.

Gómez, G.D. and Balcázar, J.L., 2008. A review on the interactions between gut microbiota and innate immunity of fish. FEMS Immunol. Med. Microbiol. 52, 145-154.

Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221-224.

Granlund, M., Öberg, L., Sellin, M., Norgren, M., 1998. Identification of a novel insertion element, IS1548, in group B streptococci, predominantly in strains causing endocarditis. J. Infect. Dis. 177, 967-976.

Gürtük, K. and Lämmler, C., 1990. Purification and partial characterization of a cohaemolysin (CAMP-factor) produced by *Streptococcus canis*. FEMS Microbiol. Lett. 64, 97-101.

Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symp. Ser., 1999, 41, 95-98.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160-174.

Hamilton, A., Popham, D.L., Carl, D.J., Lauth, X., Nizet, V., Jones, A.L., November 2006. Penicillin-binding protein 1a promotes resistance of group B streptococcus to antimicrobial peptides. Infect. Immun. 74, 6179-6187.

Harris, T.O., Shelver, D.W., Bohnsack, J.F., Rubens, C.E., 2003. A novel streptococcal surface protease promotes virulence, resistance to opsonophagocytosis, and cleavage of human fibrinogen. J. Clin. Invest. 111, 61-70.

Hassan, A.A., Akineden, Ö., Lämmler, C., Huber-Schlenstedt, R., 2002. Molecular characterization of phenotypically CAMP-negative *Streptococcus agalactiae* isolated from bovine mastitis. J. Vet. Med., Series B 49, 257-259.

Hassan, A.A., Abdulmawjood, A., Yildirim, A.A., Fink, K., Lämmler, C., Schlenstedt, R., 2000. Identification of streptococci isolated from various sources by determination of *cfb* gene and other CAMP-factor genes. Can. J. Microbiol. 46, 946-951.

Hasson, K.W., Wyld, E.M., Fan, Y., Lingsweiller, S.W., Weaver, S.J., Cheng, J., Varner, P.W., 2009. Streptococcosis in farmed *Litopenaeus vannamei*: A new emerging bacterial disease of penaeid shrimp. Dis. Aquat. Org. 86, 93.

Henderson, B., Nair, S., Pallas, J., Williams, M.A., 2011. Fibronectin: A multidomain host adhesin targeted by bacterial fibronectin-binding proteins. FEMS Microbiol. Rev. 35, 147-200.

Hensler, M.E., Miyamoto, S., Nizet, V., 2008. Group B streptococcal β-Hemolysin/Cytolysin directly impairs cardiomyocyte viability and function. Plos One 3, 2446.

Hensler, M.E., Quach, D., Hsieh, C., Doran, K.S., Nizet, V., 2008. CAMP factor is not essential for systemic virulence of group B streptococcus. Microb. Pathog. 44, 84-88.

Hensler, M.E., Liu, G.Y., Sobczak, S., Benirschke, K., Nizet, V., Heldt, G.P., 2005. Virulence role of group B streptococcus β-Hemolysin/Cytolysin in a neonatal rabbit model of early-onset pulmonary infection. J. Infect. Dis. 191, 1287-1291.

Herbert, M., Beveridge, C., McCormick, D., Aten, E., Jones, N., Snyder, L., Saunders, N., 2005. Genetic islands of *Streptococcus agalactiae* strains NEM316 and 2603VR and their presence in other group B streptococcal strains. BMC Microbiol. 5, 31.

Holmes, A.R., McNab, R., Millsap, K.W., Rohde, M., Hammerschmidt, S., Mawdsley, J.L., Jenkinson, H.F., 2001. The *pavA* gene of *Streptococcus pneumoniae* encodes a fibronectin-binding protein that is essential for virulence. Mol. Microbiol. 41, 1395-1408.

Holden M.T.G., Heather Z., Paillot R., Steward K.F., Webb K., et al. (2009) Genomic Evidence for the Evolution of *Streptococcus equi*: Host Restriction, Increased Virulence, and Genetic Exchange with Human Pathogens. PLoS Pathog 5(3): e1000346.

Huson, D., Richter, D., Rausch, C., Dezulian, T., Franz, M., Rupp, R., 2007. Dendroscope: An interactive viewer for large phylogenetic trees. BMC Bioinformatics 8, 460.

Jacobsson, K., 2003. A novel family of fibrinogen-binding proteins in *Streptococcus agalactiae*. Vet. Microbiol. 96, 103-113.

Jiang, M., Babiuk, L.A., Potter, A.A., 1996. Cloning, sequencing and expression of the CAMP factor gene of *Streptococcus uberis*. Microb. Pathog. 20, 297-307.

Jiang, S., Park, S.E., Yadav, P., Paoletti, L.C., Wessels, M.R., 2012. Regulation and function of pilus island 1 in group B streptococcus. J. Bacteriol. 194, 2479-2490.

Jones, A.L., Knoll, K.M., Rubens, C.E., 2000. Identification of *Streptococcus agalactiae* virulence genes in the neonatal rat sepsis model using signature-tagged mutagenesis. Mol. Microbiol. 37, 1444-1455.

Jones, A.L., Mertz, R.H., Carl, D.J., Rubens, C.E., 2007. A streptococcal penicillin-binding protein is critical for resisting innate airway defenses in the neonatal lung. J. Immunol. 179, 3196-3202.

Jones, A.L., Needham, R.H.V., Clancy, A., Knoll, K.M., Rubens, C.E., 2003. Penicillin-binding proteins in *Streptococcus agalactiae*: A novel mechanism for evasion of immune clearance. Mol. Microbiol. 47, 247-256.

Jones, N., Bohnsack, J.F., Takahashi, S., Oliver, K.A., Chan, M., Kunst, F., Glaser, P., Rusniok, C., Crook, D.W.M., Harding, R.M., Bisharat, N., Spratt, B.G., 2003. Multilocus sequence typing system for group B streptococcus. J. Clin. Microbiol. 41, 2530-2536.

Jonsson, I., Pietrocola, G., Speziale, P., Verdrengh, M., Tarkowski, A., 2005. Role of fibrinogen-binding adhesin expression in septic arthritis and septicemia caused by *Streptococcus agalactiae*. J. Infect. Dis. 192, 1456-1464.

Jürgens, D., Sterzik, B., Fehrenbach, F.J., 1987. Unspecific binding of group B streptococcal cocytolysin (CAMP factor) to immunoglobulins and its possible role in pathogenicity. J. Exp. Med. 165, 720-732.

Ke, D., Ménard, C., Picard, F.J., Boissinot, M., Ouellette, M., Roy, P.H., Bergeron, M.G., 2000. Development of conventional and real-time PCR assays for the rapid detection of group B streptococci. Clin. Chem. 46, 324-331.

Kong, F., Gowan, S., Martin, D., James, G., Gilbert, G.L., 2002. Serotype identification of group B streptococci by PCR and sequencing. J. Clin. Microbiol. 40, 216-226.

Konto-Ghiorghi, Y., Mairey, E., Mallet, A., Dumanil, G., Caliot, E., Trieu-Cuot, P., Dramsi, S., 2009. Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. PLoS Pathog 5, e1000422.

Krishnan, V., Gaspar, A.H., Ye, N., Mandlik, A., Ton-That, H., Narayana, S.V.L., 2007. An IgG-like domain in the minor pilin GBS52 of *Streptococcus agalactiae* mediates lung epithelial cell adhesion. Structure 15, 893-903.

Lamy, M., Dramsi, S., Billoët, A., Réglier-Poupet, H., Tazi, A., Raymond, J., Guérin, F., Couvé, E., Kunst, F., Glaser, P., Trieu-Cuot, P., Poyart, C., 2006. Rapid detection of the "highly virulent" group B streptococcus ST-17 clone. Microb. Infect. 8, 1714-1722.

Lamy, M., Zouine, M., Fert, J., Vergassola, M., Couve, E., Pellegrini, E., Glaser, P., Kunst, F., Msadek, T., Trieu-Cuot, P., Poyart, C., 2004. CovS/CovR of group B streptococcus: A two-component global regulatory system involved in virulence. Mol. Microbiol. 54, 1250-1268.

Lang, S., Xue, J., Guo, Z., Palmer, M., 2007. *Streptococcus agalactiae* CAMP factor binds to GPI-anchored proteins. Med. Microbiol. Immunol. (Berl.) 196, 1-10.

Larsson, C., Stålhammar-Carlemalm, M., Lindahl, G., 1996. Experimental vaccination against group B streptococcus, an encapsulated bacterium, with highly purified preparations of cell surface proteins rib and alpha. Infect. Immun. 64, 3518-3523.

Lauer, P., Rinaudo, C.D., Soriani, M., Margarit, I., Maione, D., Rosini, R., Taddei, A.R., Mora, M., Rappuoli, R., Grandi, G., Telford, J.L., 2005. Genome analysis reveals pili in Group B Streptococcus. Science 309, 105-105.

Lin, F.P., Lan, R., Sintchenko, V., Gilbert, G.L., Kong, F., Coiera, E., 2011. Computational bacterial genome-wide analysis of phylogenetic profiles reveals potential virulence genes of *Streptococcus agalactiae*. Plos One 6, e17964.

Liu, G.Y., Doran, K.S., Lawrence, T., Turkson, N., Puliti, M., Tissi, L., Nizet, V., 2004. Sword and shield: Linked group B streptococcal β-hemolysin/cytolysin and carotenoid pigment function to subvert host phagocyte defense. Proc. Natl. Acad. Sci. U S A 101, 14491-14496.

Lizcano, A., Sanchez, C.J., Orihuela, C.J., 2012. A role for glycosylated serine-rich repeat proteins in Gram-positive bacterial pathogenesis. Molecular Oral Microbiology 27, 257-269.

Maisey, H.C., Hensler, M., Nizet, V., Doran, K.S., February 15, 2007. Group B streptococcal pilus proteins contribute to adherence to and invasion of brain microvascular endothelial cells. J. Bacteriol. 189, 1464-1467.

Maisey, H.C., Doran, K.S., Nizet, V., 2008. Recent advances in understanding the molecular basis of group B Streptococcus virulence. Expert Reviews in Molecular Medicine 10, e27.

Maisey, H.C., Quach, D., Hensler, M.E., Liu, G.Y., Gallo, R.L., Nizet, V., Doran, K.S., 2008. A group B streptococcal pilus protein promotes phagocyte resistance and systemic virulence. The FASEB Journal 22, 1715-1724.

Marchlewicz, B.A. and Duncan, J.L., 1981. Lysis of erythrocytes by a hemolysin produced by a group B streptococcus sp. Infect. Immun. 34, 787-794.

Margarit, I., Rinaudo, C.D., Galeotti, C.L., Maione, D., Ghezzo, C., Buttazzoni, E., Rosini, R., Runci, Y., Mora, M., Buccato, S., Pagani, M., Tresoldi, E., Berardi, A., Creti, R., Baker, C.J., Telford, J.L., Grandi, G., 2009. Preventing bacterial infections with pilus-based vaccines: The group B streptococcus paradigm. J. Infect. Dis. 199, 108-115.

Mian, G.F., Godoy, D.T., Leal, C.A.G., Yuhara, T.Y., Costa, G.M., Figueiredo, H.C.P., 2009. Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. Vet. Microbiol. 136, 180-183.

Milligan, T.W., Baker, C.J., Straus, D.C., Mattingly, S.J., 1978. Association of elevated levels of extracellular neuraminidase with clinical isolates of type III group B streptococci. Infect. Immun. 21, 738-746.

Mistou, M., Dramsi, S., Brega, S., Poyart, C., Trieu-Cuot, P., July 1, 2009. Molecular dissection of the secA2 locus of group B streptococcus reveals that glycosylation of the Srr1 LPXTG protein is required for full virulence. J. Bacteriol 191, 4195-4206.

Mulero, I., Noga, E.J., Meseguer, J., García-Ayala, A., Mulero, V., 2008. The antimicrobial peptides piscidins are stored in the granules of professional phagocytic granulocytes of fish and are delivered to the bacteria-containing phagosome upon phagocytosis. Dev. Comp. Immun. 32, 1531-1538

Musser, J.M., Mattingly, S.J., Quentin, R., Goudeau, A., Selander, R.K., 1989. Identification of a high-virulence clone of type III *Streptococcus agalactiae* (group B streptococcus) causing invasive neonatal disease. Proc Natl Acad Sci 86, 4731-4735.

Nickmans, S., Verhoye, E., Boel, A., Van Vaerenbergh, K., De Beenhouwer, H., 2012. Possible solution to the problem of nonhemolytic group B streptococcus on granada medium. J. Clin. Microbiol. 50, 1132-1133.

Nizet, V., Gibson, R.L., Chi, E.Y., Framson, P.E., Hulse, M., Rubens, C.E., 1996. Group B streptococcal beta-hemolysin expression is associated with injury of lung epithelial cells. Infect. Immun. 64, 3818-3826.

Nizet, V., Kim, K.S., Stins, M., Jonas, M., Chi, E.Y., Nguyen, D., Rubens, C.E., 1997. Invasion of brain microvascular endothelial cells by group B streptococci. Infect. Immun. 65, 5074-5081.

Nobbs, A.H., Rosini, R., Rinaudo, C.D., Maione, D., Grandi, G., Telford, J.L., August 2008. Sortase A utilizes an ancillary protein anchor for efficient cell wall anchoring of pili in *Streptococcus agalactiae*. Infect. Immun. 76, 3550-3560.

Papasergi, S., Brega, S., Mistou, M., Firon, A., Oxaran, V., Dover, R., Teti, G., Shai, Y., Trieu-Cuot, P., Dramsi, S., 2011. The GBS PI-2a pilus is required for virulence in mice neonates. Plos One 6, e18747.

Pereira, U., Mian, G., Oliveira, I., Benchetrit, L., Costa, G., Figueiredo, H., 2010. Genotyping of *Streptococcus agalactiae* strains isolated from fish, human and cattle and their virulence potential in Nile tilapia. Vet. Microbiol. 140, 186-192.

Peng, K., Lee, S., Hour, A., Pan, C., Lee, L., Chen, J., 2012. Five different piscidins from Nile tilapia, *Oreochromis niloticus*: Analysis of their expressions and biological functions. PloS One 7, e50263.

Peschel, A., 2002. How do bacteria resist human antimicrobial peptides? Trends Microbiol. 10, 179-186.

Podbielski, A., Blankenstein, O., Lütticken, R., 1994. Molecular characterization of the *cfb* gene encoding group B streptococcal CAMP-factor. Med. Microbiol. Immunol. (Berl.) 183, 239-256.

Poyart, C., Pellegrini, E., Gaillot, O., Boumaila, C., Baptista, M., Trieu-Cuot, P., 2001. Contribution of Mn-cofactored superoxide dismutase (SodA) to the virulence of *Streptococcus agalactiae*. Infect. Immun. 69, 5098-5106.

Poyart, C., Tazi, A., Reglier-Poupet, H., Billoet, A., Tavares, N., Raymond, J., Trieu-Cuot, P., 2007. A multiplex PCR assay for rapid and accurate capsular typing of group B streptococcus. J. Clin. Microbiol. 45, 1985-1988.

Poyart, C., Tazi, A., Réglier-Poupet, H., Billoët, A., Tavares, N., Raymond, J., & Trieu-Cuot, P. (2007). Multiplex PCR assay for rapid and accurate capsular typing of group B streptococci. J. Clinl. microbiol., *45*(6), 1985-1988.

Pracht, D., Elm, C., Gerber, J., Bergmann, S., Rohde, M., Seiler, M., Kim, K.S., Jenkinson, H.F., Nau, R., Hammerschmidt, S., 2005. PavA of *Streptococcus pneumoniae* modulates adherence, invasion, and meningeal inflammation. Infect. Immun. 73, 2680-2689.

Pritzlaff, C.A., Chang, J.C.W., Kuo, S.P., Tamura, G.S., Rubens, C.E., Nizet, V., 2001. Genetic basis for the ?-haemolytic/cytolytic activity of group B streptococcus. Mol. Microbiol. 39, 236-248.

Puliti, M., Nizet, V., von Hunolstein, C., Bistoni, F., Mosci, P., Oreflci, G., Tissi, L., 2000. Severity of group B streptococcal arthritis is correlated with β-hemolysin expression. J. Infect. Dis. 182, 824-832.

Pyburn, T.M., Bensing, B.A., Xiong, Y.Q., Melancon, B.J., Tomasiak, T.M., Ward, N.J., Yankovskaya, V., Oliver, K.M., Cecchini, G., Sulikowski, G.A., Tyska, M.J., Sullam,

P.M., Iverson, T.M., 2011. A structural model for binding of the serine-rich repeat adhesin GspB to host carbohydrate receptors. PLoS Pathog 7, e1002112.

Raskin, D., Seshadri, R., Pukatzki, S., Mekalanos, J., 2006. Bacterial genomics and pathogen evolution. Cell 124, 703-714.

Reiß, A., Braun, J.S., Jäger, K., Freyer, D., Laube, G., Bührer, C., Felderhoff-Müser, U., Stadelmann, C., Nizet, V., Weber, J.R., 2011. Bacterial pore-forming cytolysins induce neuronal damage in a rat model of neonatal meningitis. J. Infect. Dis. 203, 393-400.

Rigel, N.W. and Braunstein, M., 2008. A new twist on an old pathway accessory secretion systems. Mol. Microbiol. 69, 291-302.

Rinaudo, C.D., Rosini, R., Galeotti, C.L., Berti, F., Necchi, F., Reguzzi, V., Ghezzo, C., Telford, J.L., Grandi, G., Maione, D., 2010. Specific involvement of pilus type 2a in biofilm formation in group B *Streptococcus*. Plos One 5, e9216.

Ring, A., Braun, J.S., Pohl, J., Nizet, V., Stremmel, W., Shenep, J.L., 2002. Group B streptococcal β-hemolysin induces mortality and liver injury in experimental sepsis. J. Infect. Dis. 185, 1745-1753.

Rodríguez, F., Oliver, J.L., Marín, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142, 485-501.

Ronquist, F. and Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574

Rosenau, A., Martins, K., Amor, S., Gannier, F., Lanotte, P., van der Mee-Marquet, N., Mereghetti, L., Quentin, R., March 2007. Evaluation of the ability of *Streptococcus agalactiae strains* isolated from genital and neonatal specimens to bind to human fibrinogen and correlation with characteristics of the *fbsA* and *fbsB* genes. Infect. Immun. 75, 1310-1317.

Rosini, R., Rinaudo, C.D., Soriani, M., Lauer, P., Mora, M., Maione, D., Taddei, A., Santi, I., Ghezzo, C., Brettoni, C., Buccato, S., Margarit, I., Grandi, G., Telford, J.L., 2006. Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. Mol. Microbiol. 61, 126-141.

Rühlmann, J., Wittmann-Liebold, B., Jürgens, D., Fehrenbach, F.J., 1988. Complete amino acid sequence of protein B. FEBS Lett. 235, 262-266.

Samen, U., Eikmanns, B.J., Reinscheid, D.J., Borges, F., November 2007. The surface protein srr-1 of streptococcus agalactiae binds human keratin 4 and promotes adherence to epithelial HEp-2 cells. Infect. Immun. 75, 5405-5414.

Santi, I., Maione, D., Galeotti, C.L., Grandi, G., Telford, J.L., Soriani, M., 2009. BibA induces opsonizing antibodies conferring in vivo protection against group B streptococcus. J. Infect. Dis. 200, 564-570.

Santi, I., Scarselli, M., Mariani, M., Pezzicoli, A., Masignani, V., Taddei, A., Grandi, G., Telford, J.L., Soriani, M., 2007. BibA: A novel immunogenic bacterial adhesin contributing to group B streptococcus survival in human blood. Mol. Microbiol. 63, 754-767.

Schubert, A., Zakikhany, K., Pietrocola, G., Meinke, A., Speziale, P., Eikmanns, B.J., Reinscheid, D.J., 2004. The fibrinogen receptor FbsA promotes adherence of *Streptococcus agalactiae* to human epithelial cells. Infect. Immun. 72, 6197-6205.

Schubert, A., Zakikhany, K., Schreiner, M., Frank, R., Spellerberg, B., Eikmanns, B.J., Reinscheid, D.J., 2002. A fibrinogen receptor from group B streptococcus interacts with fibrinogen by repetitive units with novel ligand binding sites. Mol. Microbiol. 46, 557-569.

Seifert, K.N., Adderson, E.E., Whiting, A.A., Bohnsack, J.F., Crowley, P.J., Brady, L.J., 2006. A unique serine-rich repeat protein (srr-2) and novel surface antigen (ε) associated with a virulent lineage of serotype III *Streptococcus agalactiae*. Microbiology 152, 1029-1040.

Seo, H.S., Mu, R., Kim, B.J., Doran, K.S., Sullam, P.M., 2012. Binding of glycoprotein Srr1 of *Streptococcus agalactiae* to fibrinogen promotes attachment to brain endothelium and the development of meningitis. PLoS Pathog 8, 1002947.

Severi, E., Hood, D.W., Thomas, G.H., 2007. Sialic acid utilization by bacterial pathogens. Microbiology 153, 2817-2822.

Sheen, T.R., Jimenez, A., Wang, N., Banerjee, A., van Sorge, N.M., Doran, K.S., 2011. Serine-rich repeat proteins and pili promote *Streptococcus agalactiae* colonization of the vaginal tract. J. Bacteriol. 193, 6834-6842.

Silphaduang, U. and Noga, E.J., 2001. Antimicrobials: Peptide antibiotics in mast cells of fish. Nature 414, 268-269.

Skalka, B. and Smola, J., 1981. Lethal effect of CAMP-factor and UBERIS-factor - a new finding about diffusible exosubstances of *Streptococcus agalactiae* and *Streptococcus uberis*. Zentralblatt Für Bakteriologie, Mikrobiologie Und Hygiene.1.Abt.Originale.A, Medizinische Mikrobiologie, Infektionskrankheiten Und Parasitologie 249, 190-194.

Slotved, H., Kong, F., Lambertsen, L., Sauer, S., Gilbert, G.L., September 2007. Serotype IX, a proposed new *Streptococcus agalactiae* serotype. J. Clin. Microbiol. 45, 2929-2936.

Sørensen, U.B.S., Poulsen, K., Ghezzo, C., Margarit, I., Kilian, M., 2010. Emergence and global dissemination of host-specific *Streptococcus agalactiae* clones. Mbio 1.3.

Spellerberg, B., Martin, S., Brandt, C., Lütticken, R., 2000. The cyl genes of *Streptococcus agalactiae* are involved in the production of pigment. FEMS Microbiol. Lett. 188, 125-128.

Spellerberg, B., Pohl, B., Haase, G., Martin, S., Weber-Heynemann, J., Lütticken, R., 1999. Identification of genetic determinants for the hemolytic activity of *Streptococcus agalactiae* by ISS1Transposition. J. Bacteriol. 181, 3212-3219.

Suanyuk, N., Kong, F., Ko, D., Gilbert, G.L., Supamattaya, K., 2008. Occurrence of rare genotypes of streptococcus agalactiae in cultured red tilapia *Oreochromis* sp. and Nile tilapia *O. niloticus* in Thailand-Relationship to human isolates? Aquaculture 284, 35-40.

Tazi, A., Disson, O., Bellais, S., Bouaboud, A., Dmytruk, N., Dramsi, S., Mistou, M., Khun, H., Mechler, C., Tardieux, I., Trieu-Cuot, P., Lecuit, M., Poyart, C., 2010. The surface protein HvgA mediates group B streptococcus hypervirulence and meningeal tropism in neonates. J. Exp. Med. 207, 2313-2322.

Telford, J.L., Barocchi, M.A., Margarit, I., Rappuoli, R., Grandi, G., 2006. Pili in Gram-positive pathogens. Nat Rev Micro 4, 509-519.

Tenenbaum, T., Bloier, C., Adam, R., Reinscheid, D.J., Schroten, H., 2005. Adherence to and invasion of human brain microvascular endothelial cells are promoted by fibrinogen-binding protein FbsA of *Streptococcus agalactiae*. Infect. Immun. 73, 4404-4409.

van Sorge, N.M., Quach, D., Gurney, M.A., Sullam, P.M., Nizet, V., Doran, K.S., 2009. The group B streptococcal serine-rich repeat 1 glycoprotein mediates penetration of the blood-brain barrier. J. Infect. Dis. 199, 1479-1487.

Vandamme, P., Devriese, L.A., Pot, B., Kersters, K., Melin, P., 1997. *Streptococcus difficile* is a nonhemolytic group B, type Ib streptococcus. Int. J. Syst. Bacteriol. 47, 81-85.

Wassenaar, T.M. and Gaastra, W., 2001. Bacterial virulence: Can we draw the line? FEMS Microbiol. Lett. 201, 1-7.

Wessels, M.R., Rubens, C.E., Benedí, V.J., Kasper, D.L., 1989. Definition of a bacterial virulence factor: Sialylation of the group B streptococcal capsule. Proc. Natl. Acad. Sci. 86, 8983-8987.

Wilkinson, H.W., 1977. CAMP-disk test for presumptive identification of group B streptococci. J. Clin. Microbiol. 6, 42-45.

Yang, Q., Zhang, M., Harrington, D.J., Black, G.W., Sutcliffe, I.C., 2010. A proteomic investigation of *Streptococcus agalactiae* grown under conditions associated with neonatal exposure reveals the upregulation of the putative virulence factor C protein β antigen. International J. Med. Microbiol. 300, 331-337.

Zapun, A., Contreras-Martel, C., Vernet, T., 2008. Penicillin-binding proteins and β-lactam resistance. FEMS Microbiol. Rev. 32, 361-385.

Chapter 6

General discussion

6.1 Context of this study

The objective of this study was to enhance our understanding of the molecular epidemiology, host-adaptation and pathogenicity of *S. agalactiae* in aquatic species, with particular emphasis on tilapia. *Streptococcus agalactiae* is a multi-host pathogen and numerous typing methods have been used in an attempt to decipher the population structure and relatedness of isolates of bovine and human origin (Brochet et al., 2006; Sørensen et al., 2010). Concerning *S. agalactiae* from aquatic origins, scientific information on this important threat for the aquaculture industry had mainly focused on case reporting (Hernández et al., 2009) and/or experimental challenges (Mian et al., 2009), with limited information in terms of pathogenesis, virulence determinants and genotypes of the strains involved. However, two studies published shortly before the work described in this thesis commenced provided the first insights into the population structure of *S. agalactiae* from aquatic animals. Evans et al. (2008) and Suanyuk et al. (2008) used standardized typing systems, covering either the core genome or the accessory genome, to compare genotypes of *S. agalactiae* between aquatic species, humans and bovines. Based on comparison with equivalent *S. agalactiae* typing information already existing in public databases, their results revealed that genetically distinct subpopulations of *S. agalactiae* do occur in fish, with some being apparently unique to isolates occurring in fish and others more similar to isolates occurring in humans. Furthermore, experimental infections of tilapia using bovine mastitis-causing isolates revealed that these bacteria could not infect tilapia, suggesting that *S. agalactiae* from cattle and from fish belong to distinct subpopulations. When this thesis project was initiated, no information regarding the genome content of *S. agalactiae* recovered from clinically affected fish was available. Therefore, in order to expand the

available knowledge, the present study was designed to provide a sequential and multidisciplinary approach. Starting with a global collection of aquatic *S. agalactiae* isolates, the population structure and relatedness of these isolates with those from human or bovine origins was evaluated using a combination of molecular typing methods. Specific strains were then studied further, *in vivo*, using an experimental tilapia challenge model, which allowed an assessment of the virulence of strains representing specific *S. agalactiae* subpopulations and the associated disease pathology. Based on these results, a single strain identified as fish-adapted and extremely virulent in tilapia was selected for whole genome sequencing. This study provided the first comparative genomic study to include *S. agalactiae* from fish, with the aim of identifying genes with a putative role in virulence or host adaptation.

6.2 *Streptococcus agalactiae* and host-adapted subpopulations

Streptococcus agalactiae has been the subject of numerous typing studies in the past decades, with multi-locus sequence typing (MLST) generally accepted as the standard typing method for global population studies (Urwin and Maiden, 2003). Whole genome sequencing may replace MLST in the future, but MLST remains an important tool for molecular epidemiology because the MLST database includes more isolates than have been characterized based on whole genome sequencing (Spratt, 2011), thus providing a less detailed but broader frame of reference than offered by whole genome sequences. MLST is an unambiguous method based on sequencing of the internal portion of selected housekeeping genes (Jones et al., 2003). It is used to define sequence types (STs) and clonal complexes (CCs). Studies have identified that some STs or CCs correspond to subpopulations associated with specific host-species (**Figure 6.1**).

Among the different CCs identified, CC17 corresponds to a very homogenous group that is adapted to humans and which is over-represented among isolates responsible for invasive infections in neonates (Brochet et al., 2006), while CC67 corresponds to a specific cluster of

strains pathogenic for cattle. Phylogenetic analysis based solely on MLST suggests a close relationship between CC17 and CC67, and authors have hypothesised that the hyper-invasive neonatal clone arose from a bovine ancestor (Bisharat et al., 2004). However, subsequent studies considering a larger part of the genome or the whole core genome did not support this close relationship between CC17 and CC67 (Brochet et al., 2006; Sørensen et al., et al., 2010; Chapter 4). In bovines, more recent work showed that other strains are commonly involved in mastitis (Sørensen et al., 2010; Zadoks et al., 2011), including those belonging to ST103. Phylogenetic analysis based on the core genome of *S. agalactiae*, including the genome of a ST103 strain recently made publicly available in GenBank, indicated that these non-haemolytic isolates are very distinct from others found in cattle or humans (**Figure 6.2)** and that they are most closely related to the non-haemolytic CC552 that is associated with fish and other poikilotherms (Chapter 2). The adaptation of CC552 to poikilotherms is based on a reductive evolutionary process and seems to have occurred anciently (Chapter 4).

In contrast, some CCs are shared between host species (**Figure 6.1**), suggesting that cross-species transmissions may occur. For example, CC23 is frequently identified among human carriage strains (Jones et al., 2003) and it has also been reported in numerous other host-species but never in fish (Chapter 2). Another example is CC7; isolates belonging to that CC are considered to be human-associated and are genetically very distant from CC552 (Chapter 2 and 4), yet it has been frequently reported in fish. This suggests that host adaptation is not only reflected at the level of the core genome (for CC552) but may also be driven by genes present in the accessory genome (for CC7).

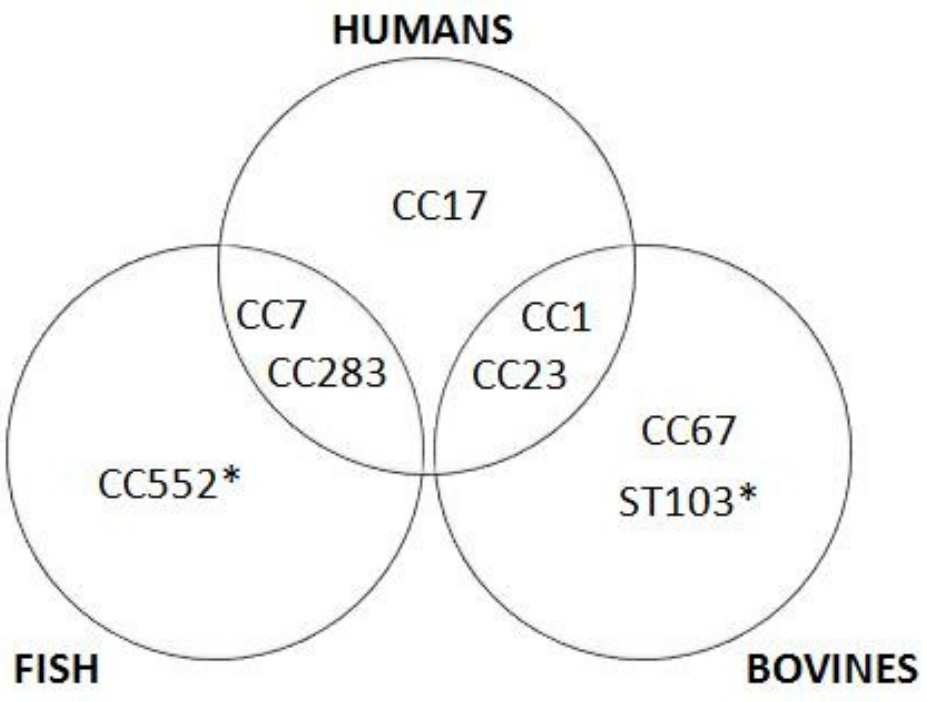


Figure 6.1 Venn diagram illustrating *S. agalactiae* genotypes predominantly identified in a single host or shared between humans, bovines or fish. Clonal complexes (CCs) or sequence types (STs) as obtained by multi-locus sequence typing are displayed.*clusters of non-haemolytic isolates**.**

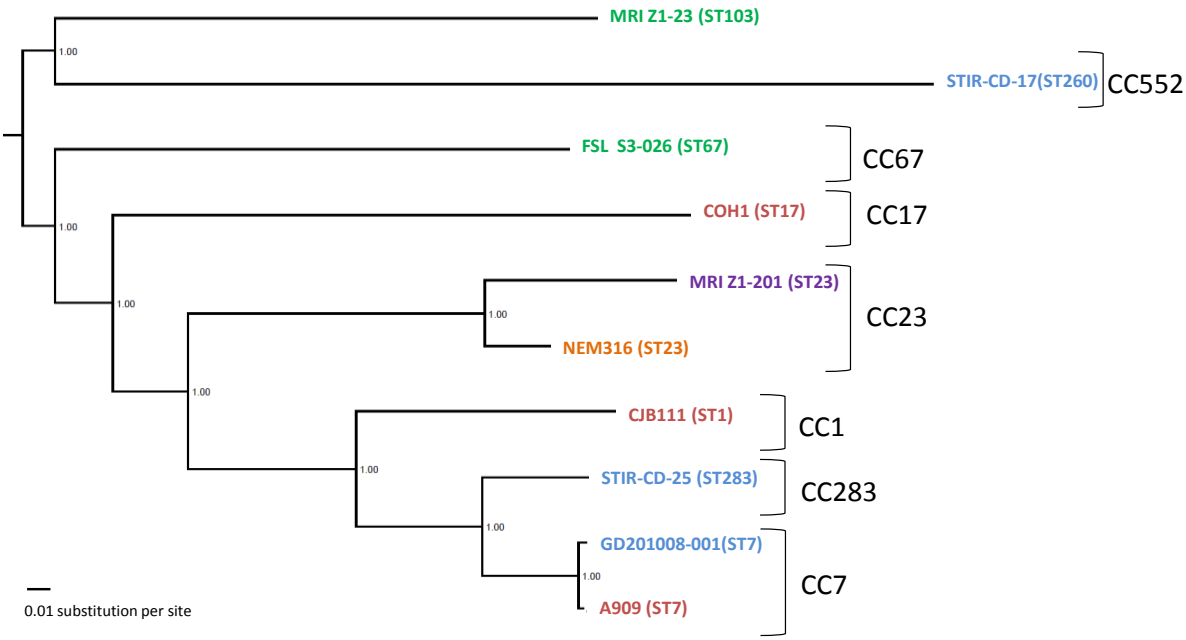


Figure 6.2 Bayesian phylogenetic tree constructed based on the core genome of 10 S. agalactiae isolates corresponding to sequence types (STs) belonging to clonal complexes (CCs) mentioned in Figure 6.1. The host of origin for each isolate is indicated by the colour of the isolate name: blue, fish; green, bovine; red, human; purple, sea mammal; orange, unknown. Posterior probabilities are shown at each node and the scale bar represents the substitutions per site**.**

6.3 Possible routes of exposure in aquatic animals

The fact that the same STs have been found in fish, other aquatic species and humans suggests that interspecies transmission may occur.

In fish, isolates belonging to CC7 (ST7 and ST500) or CC283 (ST283 and ST491) have been reported from clinical outbreaks in Kuwait (Chapter2; Evans et al., 2008), China (Ye et al., 2011), Thailand (Chapter 2) and Vietnam (Chapter 2). ST500 and ST491 have only been reported in fish thus far, but CC7 and CC283 are part of a large group of predominantly human-associated CCs (Chapter 2). ST7 has been associated with infection in humans in numerous countries, including countries in Asia (Evans et al., 2008), whereas the ST283 has been associated with invasive disease in human adults predominantly in Asia (Ip et al., 2006). ST7 also occurs as a carriage strain in humans, where *S. agalactiae* may be part of the natural flora of the gastro-intestinal and genito-urinary tracts (Dillon et al., 1982; Jones et al., 2003).

To our knowledge the STs of human-colonizing *S. agalactiae* strains have not been investigated among Asian populations. A Danish study further identified that *S. agalactiae* from human origin may survive up to 4 weeks in water (Jensen and Berg, 1982). These studies are not directly related to aquaculture but combination of this information suggests that human *S. agalactiae* may be present in aquaculture systems through effluents. In China and other Asian countries, where the bulk of worldwide tilapia production is concentrated, there is a long history of administering wastewater and human excreta to fish ponds. In extensive and semi-intensive aquaculture practise, latrines are often placed over the ponds to allow human excreta to enter the system (Lima dos Santos and Howgate, 2011), while wastewaters are used as a source of water for aquaculture, but also as a source of nutrients to support the growth of both fish and the organisms the fish feeds on (Little and Edwards, 1999). Moreover, in some rural regions, the use of untreated effluent in aquaculture ponds remains the only available method for wastewater treatment (World Health Organization,

2006a,b). These cultural practises enable high fish yields to be obtained with little or no additional feeding (Petersen and Dalsgaard, 2003). However, there are evident drawbacks in terms of food safety and disease transmission. Available evidence suggests that products from such systems will be contaminated by enteric pathogens on the skin and the gut, and that in some circumstances the organisms may invade tissues including muscles (Buras et al., 1987). Epidemiological evidence for transmission of pathogens between human and fish due to use of sewage water is, however, limited except when it comes to the transmission of certain trematode diseases, mainly those caused by *Clonorchis* (oriental liver fluke) and *Fasciolopsis* (giant intestinal fluke) (Lima dos Santos and Howgate, 2011). There is only one epidemiological study providing evidence for transmission of *S. agalactiae* between human and fish through wastewaters: a strain of *S. agalactiae* isolated from mullet during a massive fish kill in Kuwait Bay was found to have the same RAPD pattern as a *S. agalactiae* strain isolated from sewage outfalls adjacent to the infected site (Jafar et al., 2008). Challenge experiments have confirmed that some human-derived strains of *S. agalactiae* are pathogenic to fish. Using an ST7 from human origin, intra-peritoneal challenge of tilapia successfully induced disease and death in fish (Evans et al., 2009), while Pereira et al. (2010) demonstrated that out of the 8 *S. agalactiae* from human origin assessed, one induced clinical signs and mortality in both injection and immersion trials. In conclusion, the presence of human-associated STs affecting fish in Asian farms may be linked to aquaculture practise leading to contacts between *S. agalactiae* of human origin and fish through the water (**Figure 6.3**).

Other isolates identified in fish belong to CC552 (Chapter 2) and have been isolated from fish in North America, South America, Eurasia and Australia, suggesting a global spread of this clone (Chapter 2, Evans et al., 2008, Lusiastuti et al., 2012). Isolates belonging to CC552 are hypervirulent in fish and present characteristics of host-restricted bacteria (Chapter 3 and 4). Accordingly, isolates belonging to this CC have never been reported in mammals. It is therefore unlikely that these clones have been introduced through interaction

with humans or other hosts. It is unknown how these strains are introduced in aquaculture systems and it may be possible that fish carry the bacterium asymptomatically (**Figure 6.3**), as described in humans and bovines; however this possibility has never been investigated. Over the last years, recurrent epizootics of *S. agalactiae* in wild giant groupers and other wild fish have been reported in Australia (Bowater et al., 2012). The strain involved presents typical characteristics of the fish-adapted subgroup of *S. agalactiae*, *i.e.* non-haemolytic, thermosensitive and CAMP negative isolates. The origin or source of these infections is unknown but it may be suggested that a reservoir in the biotic or abiotic aquatic environment exists. No outbreak of *S. agalactiae* in fish had been reported prior the report from Bowater et al (2012). It is unknown how *S. agalactiae* was introduced in the wild population in Australia, but streptococcal disease may have been introduced through global trade of aquarium fish (Wittington and Chong, 2007). In chapter 2, we reported for the first time *S. agalactiae* CC552 from aquarium fish in Australia but originally imported from South-East Asia, indicating that the release of imported aquarium fish into the wild is a plausible source of introduction of the pathogen.

To date, it is unclear if or how the transmission of *S. agalactiae* between fish occurs for all *S. agalactiae* clones identified in fish. Using an ST283 recovered from a natural outbreak of streptococcosis in tilapia in Vietnam, a study was not successful in inducing the disease by immersion exposure (Wongsathein, 2012). There are no other reports of cohabitation or immersion challenges that used human-associated ST recovered from clinical infections in fish. It may be possible that high and sustained bacterial loads in the water, such as those likely to be achieved through aquaculture practices using wastewater, are required, possibly in combination with predisposing environmental or management factors such as high temperatures, low dissolved oxygen or high stocking density (Wongsathein, 2012). Mian et al. (2009) performed experimental cohabitation challenge studies using non-haemolytic Brazilian isolates likely to belong to CC552 and successfully reproduced the disease, indicating that horizontal transfer can occur with this clone under experimental conditions.

However numerous questions remain. It is has not been evaluated whether the bacterium is shed by infected fish and this could easily be done as for *S. iniae* (Baum et al., 2013). Moreover, the survival of fish-pathogenic *S. agalactiae* in seawater or freshwater microcosms has not been tested, unlike with other fish pathogens (Avendaño-Herrera et al. 2006; Duodu and Colquhoun, 2010). Similar experiments using isolates of CC7, CC283 and CC552 would help determine whether the water and currents may be an important factor for the survival of the bacterium and the transmission of the pathogen. Finally, the natural portal(s) of entry of this pathogen in fish have not been adequately investigated thus far and warrant further attention.

Streptococcus agalactiae has also been detected in aquatic or semi aquatic animals including seals, dolphins and crocodiles (Chapter 2, Bishop et al., 2007, Evans et al., 2006). *S. agalactiae* recovered from grey seals and crocodiles all belonged to ST23 (Bishop et al., 2007; Chapter 2). Within ST23, molecular serotypes Ia and III predominate, with serotype Ia linked to humans and serotype III primarily found in dairy cattle (Brochet et al., 2006; Sørensen et al., 2010). All seal and crocodile isolates had serotype Ia, suggesting a human origin. In humans, ST23 is a commonly recto-vaginal carried strain in adults although it may also cause neonatal invasive disease (Jones et al, 2003). Given the predominant niche of ST23 in humans, it is conceivable that its presence in these hosts is due to microbial contamination of surface water (**Figure 6.3**). *S. agalactiae* isolated from dolphins in Scotland and Kuwait in this study and elsewhere (Chapter2; Evans et al., 2006) were part of CC7 (ST7 and ST399). During the natural outbreak in Kuwait bay, ST7 was cultured from fish and a bottlenose dolphin (*Tursiops truncates)*. Authors hypothesised that the dolphin may have been infected through ingestion of infected fish (Evans et al., 2006), but it is equally likely that the dolphin became infected in the same manner as fish, *i.e.* due to immersion in or ingestion of sewage contaminated water (Evans et al., 2006; **Figure 6.3**). Transmission from fish to humans has not been reported but may theoretically occur (**Figure 6.3**). Consumption of fish has been associated with an increased risk of *S.*

agalactiae serotype Ia and Ib colonization in people (Foxman et al., 2007). In the case of *S. iniae*, it has been shown that people manipulating infected fish with wounds in their hands may be infected (Weinstein et al., 1997). It is possible that the emergence of ST283 in Asia has been driven by handling or consumption of fish, but molecular and epidemio-surveillance studies are required to investigate this possibility. Recently, isolates belonging to CC552 were identified from doctor fish used in spas in the UK and it was suggested that repeated contacts with humans may result in virulence of these strains in humans. In view of the genome reduction of this clone (Chapter 4), this seems unlikely, because it would probably require acquisition of large genomic fragments from human associated strains.

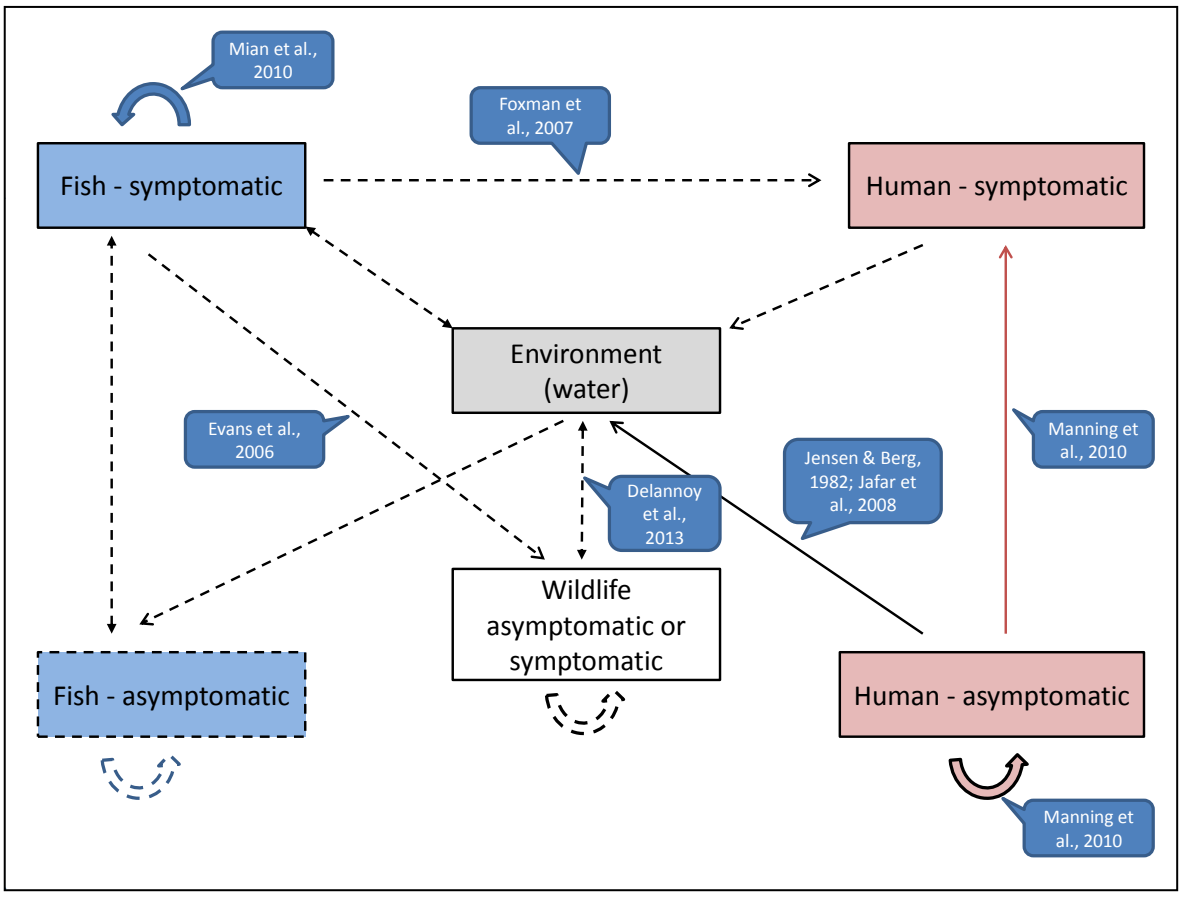


Figure 6.3 Possible routes of inter- and intra-species transmissions of *Streptococcus agalactiae* in humans and aquatic host species. Full arrows indicate proven routes, dashed arrows are hypothetical, references for routes of transmission are provided when available. Full boxes have been observed/described in the literature, dashed boxes are hypothetical.

6.4 *Streptococcus agalactiae* virulence and pathogenesis in fish

In this study, the clinical signs, the *in vivo* distribution of viable bacteria and bacterial antigens, and the gross and histopathological lesions that develop during the time course of *S. agalactiae* infection were investigated for the first time (Chapter 3). Moreover, 2 strains were evaluated, one belonging to a fish adapted subpopulation (ST260) and one belonging to a subpopulation never reported in fish (ST23). The ST260 strain was highly virulent with mortalities close to 100% with the lower dose (10^2 cfu per fish). Typical clinical signs including ataxia were induced, whereas no major clinical signs or mortalities occurred in the fish challenged with the ST23. After injection, both strains gained access to the bloodstream and viable bacteria were recovered from all organs under investigation. After 7 days post-challenge and for both strains, the formation of granulomata or other encapsulated structures and containment of bacteria within these structures appeared to be a major component of the fish response, which is in agreement with studies of natural outbreaks (Hernández et al., 2009). However, the load of viable bacteria remained high within organs and associated lesions were important in fish infected with ST260, suggesting that, unlike ST23, this strain is able to survive within macrophages and/or to evade the immune system. In chapter 3, it was demonstrated that the lack of report of ST23 strains in fish is probably not due to lack of exposure, but rather to a lack of virulence in fish, which provided an excellent basis to investigate genomic differences underlying these subpopulations.

In experimental challenges of tilapia using ST7 or ST283, a 50% mortality rate was obtained after intra-peritoneal injection of around $10^{3.3}$ and 10^7 colony forming units per fish respectively (Evans et al., 2002; Wongsathein, 2012), which correspond to doses that would induce about 100% mortalities with isolates belonging to the fish-adapted subpopulation (Chapter 3). Our challenge study therefore indicated that the fish-adapted ST260 is hypervirulent in fish, even though comparison of virulence of strains using standardised challenge models are required to confirm this. Based on the examples of comparative genomic studies of enterobacteria, staphylococci or even *S. agalactiae* from bovine origin, it

has been shown that gain in pathogenicity in bacteria may be driven by acquisition of genes through horizontal gene transfer (Fraser-Liggett, 2005; Raskin et al., 2006; Richards et al., 2011). However recent work has shed light on a much wider spectrum of virulence acquisition mechanisms in bacteria: it was shown that virulence in bacterial species like *Rickettsia* is not correlated with acquisition of foreign DNA but may result from regulatory changes of expression of virulence determinants due to genome decay (Fournier et al., 2009). In our study, the genome of the hypervirulent ST260 strain was sequenced and one of the most remarkable features of this genome was the massive genomic reduction resulting in the loss of almost all genomic islands described in *S. agalactiae* from human origin (Chapter 4). Comparative genomic analysis revealed the conserved presence of a limited number of genes encoding well-recognised virulence determinants in *S. agalactiae* from humans and including adhesins (*fbsA*, *pavA*, *srr1*), invasins (*cfb*, *hylB*) and immune evasins (*ponA*, *sodA*, *cps*) (Chapter 5). Comparison with the reference strain NEM316, which is of ST23 and known not to be virulent in tilapia (Pereira et al., 2010), similarly to the ST23 isolate used in our challenge study (Chapter 3), highlighted that these virulence genes were also present in ST23. This observation suggests that it is not the presence but the expression of these genes that may explain the ability of ST260 to cause disease in fish and this possibility requires further assessment. One important distinction between ST23 and ST260 was however identified and consisted of the presence of a distinct variant of the glycoprotein anchored protein gbs2018 in the fish-adapted group (Chapter 5). In view of the importance of gbs2018 to virulence in humans (Santi et al., 2007), it may be hypothesised that the existence of a specific variant among the fish-adapted clones reflects an adaptive evolution for virulence in fish. Finally, comparative genomic analysis enabled the identification of several genes that have possibly been acquired through horizontal gene transfer and apparently associated with strains occurring in fish (Chapter 4). With some exceptions, the majority of these genes were found to encode hypothetical proteins of unknown function and their role in virulence is unknown.

Numerous *S. agalactiae* genomes have been made publicly available over the end course of this project, including the genome of the ST23 isolate used in the challenge study (MRI Z1-201) and the non-haemolytic ST103 isolate used to construct **Figure 6.2**. Extension of the comparative genomic analysis performed in this study with these genomes would provide further insights into the genes possibly unique to piscine *S. agalactiae*. Furthermore, the approaches of the "OMICS" era have yet to be fully applied to piscine *S. agalactiae*. Further work including proteomic analyses and transcriptomic analyses to evaluate gene expression profile over the course of challenge studies would enable the identification of proteins of importance for pathogenesis in fish and could provide viable targets for vaccine development. Biological validation of the role of these genes in pathogenesis and virulence would necessitate the generation of mutants deficient or altered in the target genes, followed by *in vivo* studies to evaluate variations in virulence and/or *in vitro* studies to evaluate variations in adherence and invasion of cell lines, as well as survival within macrophages. Most studies of this sort in streptococci have been conducted using temperature-sensitive suicide plasmids such as pG^{+}host9 to generate knock-out mutants (Yamaoto et al., 2005). However, *S. agalactiae* strains belonging to the fish-adapted subgroup do not grow well at 37˚C or higher temperatures and other approaches will therefore have to be considered. An alternative for iditentification of genes important for virulence in fish may for example be achieved using a new method called TraDIS and standing for transposon directed insertion-site sequencing (Langridge et al., 2009).

6.5 Concluding remarks and future directions

The multidisciplinary approach used in this study has been successful in providing scientific information regarding the population structure, pathogenesis and virulence of *S. agalactiae* in fish. The major information researchers should take in account is that very distinct lineages of *S. agalactiae* occur in fish, with one hypervirulent lineage strictly adapted to poikilotherms through reductive evolution, and other lineages that are human-associated and apparently less virulent to fish. These lineages differ in term of transmission, pathogenesis

and virulence. This thesis laid a firm foundation for further studies; these should address the questions of epidemio-surveillance for assessment of transmission of *S. agalactiae* between human and fish and evaluation of the role of putative virulence determinants, with a view to effective control of the disease in fish through prevention or vaccine development.

6.6 References

Avendaño-Herrera, R., Irgang, R., Magariños, B., Romalde, J.L., Toranzo, A.E., 2006. Use of microcosms to determine the survival of the fish pathogen *Tenacibaculum maritimum* in seawater. Environ. Microbiol. 8, 921-928.

Baums, C.G., Hermeyer, K., Leimbach, S., Adamek, M., Czerny, C.-., Hörstgen-Schwark, G., Valentin-Weigand, P., Baumgärtner, W., Steinhagen, D., Establishment of a model of *Streptococcus iniae* meningoencephalitis in Nile tilapia (*Oreochromis niloticus*). J. Comp. Pathol. In press.

Bishop, E.J., Shilton, C., Benedict, S., Kong, F., Gilbert, G.L., Gal, D., Godoy, D., Spratt, B.G., Currie, B.J., 2007. Necrotizing fasciitis in captive juvenile *Crocodylus porosus* caused by *Streptococcus agalactiae*: An outbreak and review of the animal and human literature. Epidemiol. Infect. 135, 1248-1255.

Bowater, R.O., Forbes-Faulkner, J., Anderson, I.G., Condon, K., Robinson, B., Kong, F., Gilbert, G.L., Reynolds, A., Hyland, S., McPherson, G., Brien, J.O., Blyde, D., 2012. Natural outbreak of *Streptococcus agalactiae* (GBS) infection in wild giant Queensland grouper, *Epinephelus lanceolatus* (Bloch), and other wild fish in northern Queensland, Australia. J. Fish Dis. 35, 173-186.

Brochet, M., Couve, E., Zouine, M., Vallaeys, T., Rusniok, C., Lamy, M., Buchrieser, C., Trieu-Cuot, P., Kunst, F., Poyart, C., Glaser, P., 2006. Genomic diversity and evolution within the species *Streptococcus agalactiae*. Microbes Infect. 8, 1227-1243.

Buras, N., Duek, L., Niv, S., Hepher, B., Sandbank, E., 1987. Microbiological aspects of fish grown in treated wastewater. Water Res. 21, 1-10.

Cumley, N.J., Smith, L.M., Anthony, M., May, R.C., 2012. The CovS/CovR acid response regulator is required for intracellular survival of group B streptococcus in macrophages. Infection and Immunity 80, 1650-1661.

Dillon, H.C., Gray, E., Pass, M.A., Gray, B.M., 1982. Anorectal and vaginal carriage of group B streptococci during pregnancy. Journal of Infectious Diseases 145, 794-799.

Duodu, S. and Colquhoun, D., 2010. Monitoring the survival of fish-pathogenic *Francisella* in water microcosms. FEMS Microbiol. Ecol. 74, 534-541.

Evans, J.J., Klesius, P.H., Pasnik, D.J., Bohnsack, J.F., 2009. Human *Streptococcus agalactiae* isolate in Nile tilapia (*Oreochromis niloticus*). Emerging Infectious Diseases 15, 774.

Evans, J.J., Bohnsack, J.F., Klesius, P.H., Whiting, A.A., Garcia, J.C., Shoemaker, C.A., Takahashi, S., 2008. Phylogenetic relationships among *Streptococcus agalactiae* isolated from piscine, dolphin, bovine and human sources: A dolphin and piscine lineage associated with a fish epidemic in Kuwait is also associated with human neonatal infections in Japan. J. Med. Microbiol. 57, 1369-1376.

Evans, J.J., Pasnik, D.J., Klesius, P.H., Al-Ablani, S., 2006. First report of *Streptococcus agalactiae* and *Lactococccus garviae* from a wild bottlenose dolphin (*Tursiops truncatus*). J. Wildl. Dis. 42, 561-569.

Fettucciari, K., Fetriconi, I., Mannucci, R., Nicoletti, I., Bartoli, A., Coaccioli, S., Marconi, P., 2006. Group B streptococcus induces macrophage apoptosis by calpain activation. The Journal of Immunology 176, 7542-7556.

Fournier, P., El Karkouri, K., Leroy, Q., Robert, C., Giumelli, B., Renesto, P., Socolovschi, C., Parola, P., Audic, S., Raoult, D., 2009. Analysis of the rickettsia africae genome reveals that virulence acquisition in rickettsia species may be explained by genome reduction. BMC Genomics 10, 166.

Foxman, B., Gillespie, B.W., Manning, S.D., Marrs, C.F., 2007. Risk factors for group B streptococcal colonization: Potential for different transmission systems by capsular type. Ann. Epidemiol. 17, 854-862.

Fraser-Liggett, C.M., 2005. Insights on biology and evolution from microbial genome sequencing. Genome Research 15, 1603-1610.

Hernández, E., Figueroa, J., Iregui, C., 2009. Streptococcosis on a red tilapia, *Oreochromis* sp., farm: A case study. J. Fish Dis. 32, 247-252.

Ip, M., Cheuk, E.S.C., Tsui, M.H.Y., Kong, F., Leung, T.N., Gilbert, G.L., 2006. Identification of a *Streptococcus agalactiae* serotype III subtype 4 clone in association with adult invasive disease in Hong Kong. J. Clin. Microbiol. 44, 4252-4254.

Jafar, Q.A., Sameer, A.Z., Salwa, A.M., Samee, A.A., Ahmed, A.M., Al-Sharifi, F., 2008. Molecular investigation of *Streptococcus agalactiae* isolates from environmental samples and fish specimens during a massive fish kill in Kuwait bay. Pak. J. Biol. Sci. 11, 2500-2504.

Jensen, N. and Berg, B., 1982. Spildevand og akvatiske biotoper som potentielle gruppe B-streptokokreservoirer. Dansk Veterinaertidsskrift 65, 197-199.

Jones, N., Bohnsack, J.F., Takahashi, S., Oliver, K.A., Chan, M.S., Kunst, F., Glaser, P., Rusniok, C., Crook, D.W., Harding, R.M., Bisharat, N., Spratt, B.G., 2003. Multilocus sequence typing system for group B streptococcus. J. Clin. Microbiol. 41, 2530-2536.

Lartigue, M., Hery-Arnaud, G., Haguenoer, E., Domelier, A., Schmit, P., van der Mee-Marquet, N., Lanotte, P., Mereghetti, L., Kostrzewa, M., Quentin, R., 2009. Identification of *Streptococcus agalactiae* isolates from various phylogenetic lineages by matrix-assisted laser desorption ionization-time of flight mass spectrometry. J. Clin. Microbiol. 47, 2284-2287.

Langridge, G. C., Phan, M. D., Turner, D. J., Perkins, T. T., Parts, L., Haase, J et al., 2009. Simultaneous assay of every *Salmonella Typhi* gene using one million transposon mutants. Genome research 19.12, 2308-2316.

Lima dos Santos, C.A.M. and Howgate, P., 2011. Fishborne zoonotic parasites and aquaculture: A review. Aquaculture 318, 253-261.

Little, D.C. and Edwards, P, 1999. Alternative strategies for livestock–fish integration with emphasis on Asia. Ambio, 28 (2) (1999), pp. 118–124.

Lusiastuti, A.M., Textor, M., Seeger, H., Akineden, Ö., Zschöck, M., 2012. The occurrence of *Streptococcus agalactiae* sequence type 261 from fish disease outbreaks of tilapia *Oreochromis niloticus* in Indonesia. Aquacult. Res. In press.

Mian, G., Godoy, D., Leal, C., Yuhara, T., Costa, G., Figueiredo, H., 2009. Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. Vet. Microbiol. 136, 180-183.

Mian, G.F., Godoy, D.T., Leal, C.A.G., Yuhara, T.Y., Costa, G.M., Figueiredo, H.C.P., 2009. Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. Vet. Microbiol. 136, 180-183.

Pereira, U., Mian, G., Oliveira, I., Benchetrit, L., Costa, G., Figueiredo, H., 2010. Genotyping of *Streptococcus agalactiae* strains isolated from fish, human and cattle and their virulence potential in Nile tilapia. Vet. Microbiol. 140, 186-192.

Petersen, A. and Dalsgaard, A., 2003. Species composition and antimicrobial resistance genes of *Enterococcus* spp., isolated from integrated and traditional fish farms in Thailand. Environ. Microbiol. 5, 395-402.

Santi, I., Scarselli, M., Mariani, M., Pezzicoli, A., Masignani, V., Taddei, A., Grandi, G., Telford, J.L., Soriani, M., 2007. BibA: A novel immunogenic bacterial adhesin contributing to group B streptococcus survival in human blood. Mol. Microbiol. 63, 754-767.

Sørensen, U.B.S., Poulsen, K., Ghezzo, C., Margarit, I., Kilian, M., July/August 2010. Emergence and global dissemination of host-specific *Streptococcus agalactiae* clones. Mbio 1, e00178-10-e00178-18.

Spratt, B.G., 2012. The 2011 Garrod lecture: From penicillin-binding proteins to molecular epidemiology. Journal of Antimicrobial Chemotherapy 67, 1578-1588.

Raskin, D.M., Seshadri, R., Pukatzki, S.U., Mekalanos, J.J., 2006. Bacterial genomics and pathogen evolution. Cell 124, 703-714.

Richards, V.P., Lang, P., Pavinski Bitar, P.D., Lefébure, T., Schukken, Y.H., Zadoks, R.N., Stanhope, M.J., 2011. Comparative genomics and the role of lateral gene transfer in the evolution of bovine adapted streptococcus agalactiae. Infection, Genetics and Evolution 11, 1263-1275.

Urwin, R. and Maiden, M.C., 2003. Multi-locus sequence typing: A tool for global epidemiology. Trends Microbiol. 11, 479-487.

Weinstein, M.R., Litt, M., Kertesz, D.A., Wyper, P., Rose, D., Coulter, M., McGeer, A., Facklam, R., Ostach, C., Willey, B.M., Borczyk, A., Low, D.E., 1997. Invasive infections due to a fish pathogen, *Streptococcus iniae*. N. Engl. J. Med. 337, 589-594.

Whittington, R. and Chong, R., 2007. Global trade in ornamental fish from an Australian perspective: The case for revised import risk analysis and management strategies. Prev. Vet. Med. 81, 92.

Wongsathein, D., 2012. Factors affecting experimental *Streptococcus agalactiae* infection in tilapia, *Oreochromis niloticus*. Thesis from Stirling University.

World Health Organization, 2006a. World Health Organization, Guidelines for the safe use of wastewater, excreta and graywater, Wastewater and excreta use in aquaculture, Wastewater and Excreta Use in Aquaculture, vol. 3, Geneva, Switzerland (2006), pp. 1–140.

World Health Organization, 2006b. World Health Organization, Health guidelines for the use of wastewater in agriculture and aquaculture, Wastewater and excreta use in aquaculture, vol. 3Geneva, Switzerland (2006), pp. 1–140.

Yamamoto, Y., Poyart, C., Trieu-Cuot, P., Lamberet, G., Gruss, A., Gaudu, P., 2005. Respiration metabolism of group B streptococcus is activated by environmental haem and quinone and contributes to virulence. Mol. Microbiol. 56, 525-534.

Ye, X., Li, J., Lu, M., Deng, G., Jiang, X., Tian, Y., Quan, Y., Jian, Q., 2011. Identification and molecular typing of *Streptococcus agalactiae* isolated from pond-cultured tilapia in China. Fisheries Science 77, 623-632.

Appendix: Bacterial collection

A collection of 33 *S. agalactiae* from distinct geographical regions and host of origin was assembled with the help and generosity of a number of researchers worldwide. A list of the bacterial isolates used in this study is provided herby (NA, Non available).

Code №	Reference	Country	Date	Species	Organ	Additionnal information	Provided by
STIR-CD-01	K0101	Kuwait	Sep-01	*Liza klunzingeri* (mullet)	Brain	Natural outbreak in Kuwait bay	Dr Kim Thomson (University of Stirling, UK) and Dr Terutoyo Yoshida (University of Miyazaki, Japan)
STIR-CD-02	K0102	Kuwait	Sep-01	*Liza klunzingeri* (mullet)	Brain	Natural outbreak in Kuwait bay	
STIR-CD-03	K0103	Kuwait	Sep-01	*Liza klunzingeri* (mullet)	Brain	Natural outbreak in Kuwait bay	
STIR-CD-04	K0104	Kuwait	Sep-01	*Liza klunzingeri* (mullet)	Brain	Natural outbreak in Kuwait bay	
STIR-CD-05	K0105	Kuwait	Sep-01	*Liza klunzingeri* (mullet)	Brain	Natural outbreak in Kuwait bay	
STIR-CD-09	B09032 Sa Ti Be 08-18b	Colombia	Aug-08	*Oreochromis* sp. (tilapia)	Kidney	Cage farm in Lake Betania	Prof Carlos Iregui (Universidad Nacional de Colombia, Bogotá, Colombia)
STIR-CD-10	B09032 Sa Ti Be 08-21a	Colombia	Nov-08	*Oreochromis* sp. (tilapia)	Kidney	Cage farm in Lake Betania	
STIR-CD-11	B09032 Sa Ti Be 21b	Colombia	Nov-08	*Oreochromis* sp. (tilapia)	Kidney	Cage farm in Lake Betania	
STIR-CD-12	B09032 Sa Ti Be 08-18 a	Colombia	Aug-08	*Oreochromis* sp. (tilapia)	Kidney	Cage farm in Lake Betania	
STIR-CD-13	B09032 Sa Ti Cr 08-14b	Costa Rica	Jun-08	*Oreochromis* sp. (tilapia)	Eye	Farm CR1	
STIR-CD-07	B08059 E52 HEART	Honduras	Aug-08	*Oreochromis* sp. (tilapia)	Heart	Farm H1	Prof Hugh W Ferguson (St. George's University, Grenada, W. Indies)
STIR-CD-17	B08059 E28 heart	Honduras	Aug-08	*Oreochromis* sp. (tilapia)	Heart	Farm H1	
STIR-CD-18	B08059 E52 spleen	Honduras	Aug-08	*Oreochromis* sp. (tilapia)	Spleen	Farm H1	
STIR-CD-19	B08059 PE45 spleen	Honduras	Aug-08	*Oreochromis* sp. (tilapia)	Spleen	Farm H1	
STIR-CD-14	May 06	Vietnam	May-06	*Oreochromis* sp. (tilapia)	NA	NA	Margaret Crumlish (University of Stirling, UK)
STIR-CD-21	09011056/1L	Thailand	May-09	*Oreochromis* sp. (tilapia)	Liver	Saraburi Province	Dr Temdoung Somsiri (Kasetsart University Campus, Thailand)
STIR-CD-22	09011088/2L	Thailand	Oct-09	*Oreochromis* sp. (tilapia)	Liver	Nakhon Sawan Province	
STIR-CD-23	GQ338316	Thailand	NA	*Oreochromis* sp. (tilapia)	Kidney	Nakhon Pathom Province	Dr Janenuj Wongtavatchai (Chulalongkorn University, Thailand)
STIR-CD-24	GQ338317	Thailand	NA	*Oreochromis* sp. (tilapia)	Kidney	Kanchanaburi Province	
STIR-CD-25	GQ338318	Thailand	NA	*Oreochromis* sp. (tilapia)	Kidney	Chachoengsao Province	
STIR-CD-26	GQ169772	Thailand	NA	*Oreochromis* sp. (tilapia)	Kidney	Mukdahan Province	
STIR-CD-27	GQ169773	Thailand	NA	*Oreochromis* sp. (tilapia)	Kidney	Nakhon Phanom Province	
STIR-CD-28	GQ169774	Thailand	NA	*Oreochromis* sp. (tilapia)	Kidney	Phetchaburi Province	
STIR-CD-29	/	Belgium	Aug-07	*Oreochromis* sp. (tilapia)	Kidney /brain	Tilapia recirculating farm	Francois Lieffrig (Centre d' Economie Rurale Groupe, Belgium)
STIR-CD-30	TCFB 1876	Australia	2004	*Puntius conchonius* (rosy barb)	Kidney	Aquarium fish	Dr Jeremy Carson (University of Tasmania, Australia)
STIR-CD-31	TCFB2983, originally P97-1103(FD1631)	Australia	1997	NA	Heart	Aquarium fish, likely to have been imported from South-East Asia	Dr Nicky Buller, Department of Agriculture and Food Western Australia, Australia)
STIR-CD-32	TCFB 2984, originally P96-1986(FD1616)	Australia	1996	*Mikrogeophagus ramirez* i (golden ram)	Liver	Aquarium fish imported from Indonesia	
MRI Z1-198	/	UK	1995	*Tursiops truncatus* (bottlenose dolphin)	Lungs	Animal stranded around the coast of Scotland	Dr Geoff Foster (SAC Veterinary Services, UK) and Dr Mark Dagleish (MRI, UK)
MRI Z1-199	/	UK	Dec-02	*Halichoerus grypus* (grey seal)	Lungs	Animal stranded around the coast of Scotland	
MRI Z1-200	/	UK	Dec-02	*Halichoerus grypus* (grey seal)	Lungs	Animal stranded around the coast of Scotland	
MRI Z1-201	/	UK	Feb-03	*Halichoerus grypus* (grey seal)	Lungs	Animal stranded around the coast of Scotland	
MRI Z1-202	/	UK	Mar-03	*Halichoerus grypus* (grey seal)	Lungs	Animal stranded around the coast of Scotland	
MRI Z1-208	/	UK	2010	*Halichoerus grypus* (grey seal)	Lungs	Animal stranded around the coast of Scotland	