



Capsular Switching and ICE Transformation Occurred in Human *Streptococcus agalactiae* ST19 With High Pathogenicity to Fish

Rui Wang^{1,2†}, Liping Li^{1†}, Ting Huang¹, Weiyi Huang², Aiyi Lei¹ and Ming Chen^{1*}

¹ Guangxi Key Laboratory for Aquatic Genetic Breeding and Healthy Aquaculture, Guangxi Academy of Fishery Sciences, Nanning, China, ² Institute of Animal Science and Technology, Guangxi University, Nanning, China

OPEN ACCESS

Edited by:

Michael Kogut,
Agricultural Research Service (USDA),
United States

Reviewed by:

David Smith,
Heriot-Watt University,
United Kingdom
Viswas Konasagara Nagaleekar,
Indian Veterinary Research Institute
(IVRI), India

*Correspondence:

Ming Chen
cm990919@163.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Veterinary Infectious Diseases,
a section of the journal
Frontiers in Veterinary Science

Received: 29 June 2018

Accepted: 23 October 2018

Published: 13 November 2018

Citation:

Wang R, Li L, Huang T, Huang W,
Lei A and Chen M (2018) Capsular
Switching and ICE Transformation
Occurred in Human *Streptococcus*
agalactiae ST19 With High
Pathogenicity to Fish.
Front. Vet. Sci. 5:281.
doi: 10.3389/fvets.2018.00281

Although *Streptococcus agalactiae* (GBS) cross-infection between human and fish has been confirmed in experimental and clinical studies, the mechanisms underlying GBS cross-species infection remain largely unclear. We have found different human GBS ST19 strains exhibiting strong or weak pathogenic to fish (sGBS and wGBS). In this study, our objective was to identify the genetic elements responsible for GBS cross species infection based on genome sequence data and comparative genomics. The genomes of 11 sGBS strains and 11 wGBS strains were sequenced, and the genomic analysis was performed based on pan-genome, CRISPRs, phylogenetic reconstruction and genome comparison. The results from the pan-genome, CRISPRs analysis and phylogenetic reconstruction indicated that genomes between sGBS were more conservative than that of wGBS. The genomic differences between sGBS and wGBS were primarily in the Cps region (about 111 kb) and its adjacent ICE region (about 106 kb). The Cps region included the entire *cps* operon, and all sGBS were capsular polysaccharide (CPS) type V, while all wGBS were CPS type III. The ICE region of sGBS contained integrative and conjugative elements (ICE) with IQ element and *erm*(TR), and was very conserved, whereas the ICE region of wGBS contained ICE with mega elements and the variation was large. The capsular switching (III–V) and transformation of ICE adjacent to the Cps region occurred in human GBS ST19 with different pathogenicity to fish, which may be related to the capability of GBS cross-infection.

Keywords: GBS, ST19, capsular switching, ICE, cross-species infection

INTRODUCTION

Streptococcus agalactiae, also known as Group B *Streptococcus* (GBS), is considered as a colonizing agent of humans and the major cause of neonatal pneumonia, meningitis and sepsis, and the frequency of invasive infection in adults has also increased within the past two decades (1). Animal infection studies have demonstrated that human GBS can infect fish (2, 3), while the clonal complex (CC) 23 clinical isolates, which belong to primarily human isolates cluster, have been isolated from aquatic animals (4). In recent years, GBS has become an emerging pathogen in aquatic environments and infects many species of culture fish, causing septicemia and meningoencephalitis, especially on tilapia farms (5). Studies show that consumption of fish has been associated with an increased risk of GBS colonization in people, as well as the spread of fish GBS

in humans (6, 7). In 2015, an outbreak of infection with GBS associated with raw freshwater fish consumption was reported, which affected more than 200 patients in Singapore (8). GBS spread quickly through the fish and water, which can cause a large area of Streptococcal outbreak in aquaculture in a short time (9). Therefore, the cross-species infection of GBS between human and fish has become a serious threat to human health, but the mechanism of GBS cross-species infection remains largely unknown.

CC19, including its namesake sequence type (ST) 19, is at the center of the GBS population according to the extent of highly conserved regions between the different CCs (10), which have been linked to carriage or disease in people and cattle (11, 12). Nevertheless, none isolate of CC19 has been reported in fish. We used invasive human GBS ST19 (28 strains) to infect tilapia, and found that 17 strains were strong pathogenic to tilapia (sGBS, the mortality rate: 70–100%), while 11 strains showed weak pathogenicity to tilapia (wGBS, the mortality rate: 0–30%) (13). In the present study, the genomic difference analysis was performed between sGBS and wGBS to explore the mechanisms related to the capability of GBS cross-species infection.

MATERIALS AND METHODS

Genome Sequencing, Assembly, And Annotation

Twenty-nine invasive GBS ST19 strains were obtained from 29 patients in five hospitals of three cities in Guangxi Province, and the details of the strains information and the isolation procedure were reported in our previous article (13). The genomes of 22 strains including sGBS (11 strains) and wGBS (11 strains) were sequenced using the Illumina HiSeq2000 sequencing platform and the results were assembled using the ABySS program 1.9.0 (14). The assembly coverage was not $<250\times$, the number of contigs was 44–118, average N50 value was 94,316, and average number of coding sequences was 2,233. The whole genome shotgun (WGS) sequences of 22 ST19 GBS strains have been deposited in GenBank. The genome contigs of sGBS strain LZFF001 and wGBS strain LZFF006 were rearranged using the genome of GBS 2603V/R strain as the reference genome by Mauve (15). The CPS type, ST, CC, host and GenBank accession number of the sequenced strains and 2603V/R were listed in Additional file 1. The assembled genome was uploaded to the RAST website and was annotated by using the RASTtk program on line (16).

Pan-Genome Analysis of ST19 GBS

The pan genome analysis was performed for 22 ST19 GBS strains using Roary with a blast identity cutoff of 97% (17). The visualization of the pan-genome was performed by Anvi'o 2.4.0 (18).

CRISPRs Analysis, and Phylogenetic Reconstruction

The clustered regularly interspaced short palindromic repeats (CRISPRs) of the strains were identified by CRISPRs finder and CRISPR recognition tool (CRT) (19, 20), each unique

spacer of the CRISPRs was numbered manually, and further optimized based on the report from Lier et al. (21). The orthologous protein searches were performed using OrthoMCL 2.0.9 (22). Multiple sequence alignment of single copy homologous protein sequences was performed by MAFFT and the poorly aligned positions and divergent regions of the alignment were removed. The maximum likelihood (ML) estimation of phylogenetic trees and model parameters were performed by PhyML program, and the optimal amino acid substitution model was obtained by comparing AIC and BIC scores (23). RaxML software was used to construct the ML based phylogenetic tree with 1,000 bootstrap replications (24). The visualization of the phylogenetic tree was performed by FigTree 1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Genomic Comparison

The genome comparative analysis was performed for 12 ST19 GBS representative strains, including 6 sGBS strains and 6 wGBS strains. The common and unique genes from annotated genomes of strains were search and compared by using the sequence based genomic comparison tool provided by SEED viewer (25). Briefly, wGBS LZFF006 and sGBS LZFF001 were used as the reference genome, respectively, and the genomes of other 10 strains were aligned to the reference genome. The result listed the genes of the reference organism in chromosomal order and displayed hits on the comparison organisms accordingly. The evolutionary analysis for homologous proteins from GBS *cps* operon was performed by using the above-mentioned method. The CPS type, ST, CC, host and GenBank accession number of strains involved in evolutionary analysis were listed in **Supplementary Table 1**.

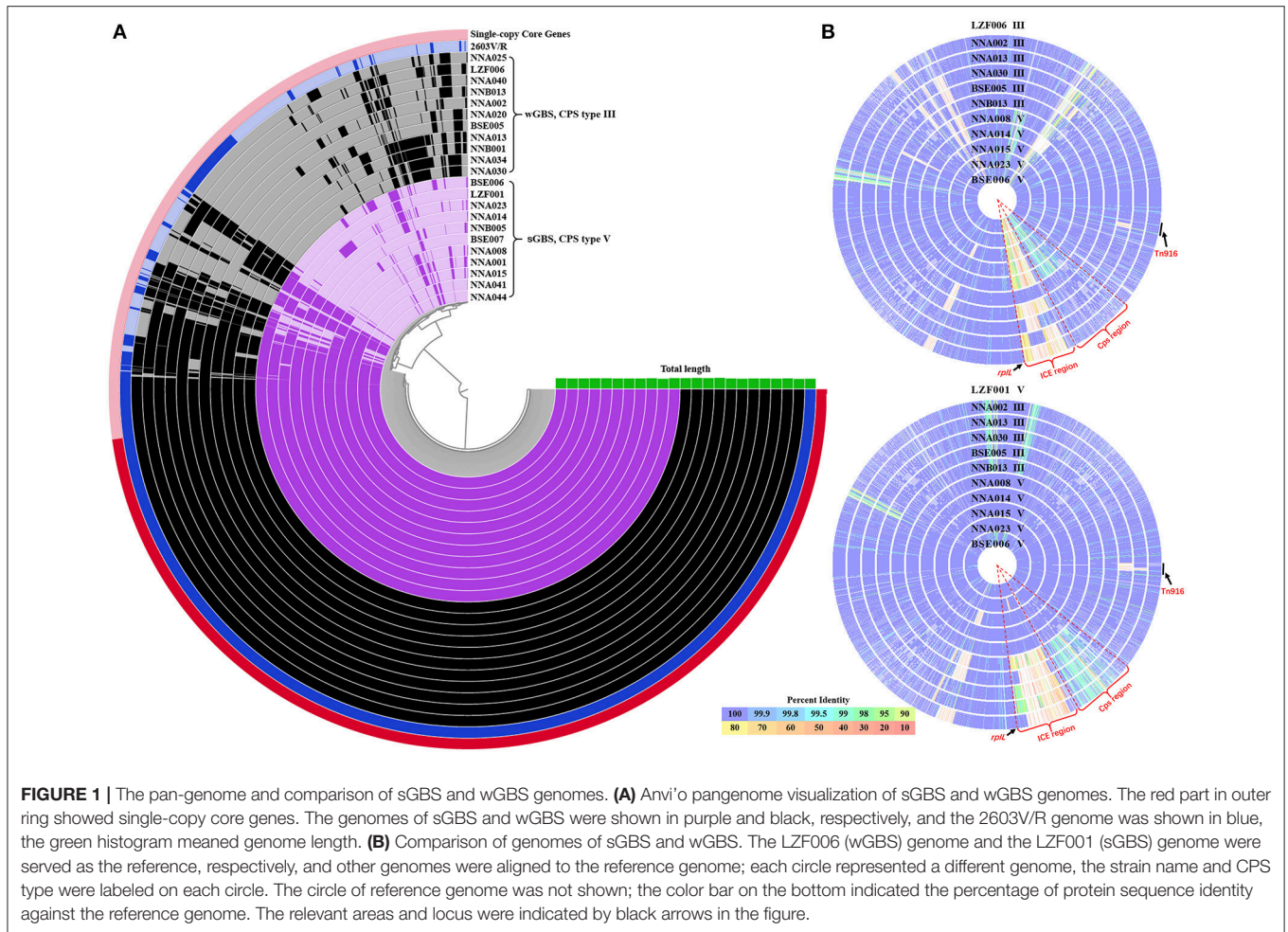
RESULTS

Analysis of the Pan-Genome

The average genome size of 22 GBS ST19 strains was 2.24 Mb, with an average gene density of 1 gene per kb. The GC content of genome from 11 wGBS strains was 35.70%, and the average number of genes was 2,247 with an average length of 884 bp for each gene. The GC content of genome from 11 sGBS strains was 35.66%, and the average number of genes was 2,208 with an average length of 902 bp for each gene. The total number of genes from 11 wGBS strains was 2,753, and the proportion of core genes and accessory genes was 67.09% (1,847/2,753) and 32.91% (906/2,753), respectively. The total number of genes from 11 sGBS strains was 2,713, and the proportion of core genes and accessory genes was 70.59% (1,915/2,713) and 29.41% (798/2,713), respectively. The total number of genes from 22 GBS strains was 3,223, among which 1,758 were core genes (54.55%) and 1,465 were accessory genes (45.45%). Visualizations of the results was shown in **Figure 1A**.

CRISPRs Analysis and Phylogenetic Reconstruction

CRISPRs analysis and phylogenetic reconstruction of ST19 GBS were shown in **Figure 2**. The spacers numbered 11–16 and 53–56 were more conserved spacers in wGBS strains. Generally, the number and type of spacers in wGBS strains were more than



that in sGBS strains. The phylogenetic analysis showed that the evolution distance between wGBS strains was farther than that between sGBS strains. The sGBS strain BSE006 was clustered with the wGBS strains, and the CRISPRs structure of BSE006 was also similar to the wGBS strains.

Genome Comparison

The genome comparison between the wGBS strains and the sGBS strains was shown in **Figure 1B**. The regions with common variation in genome were mainly located at the Cps region containing complete *cps* operon and ICE region.

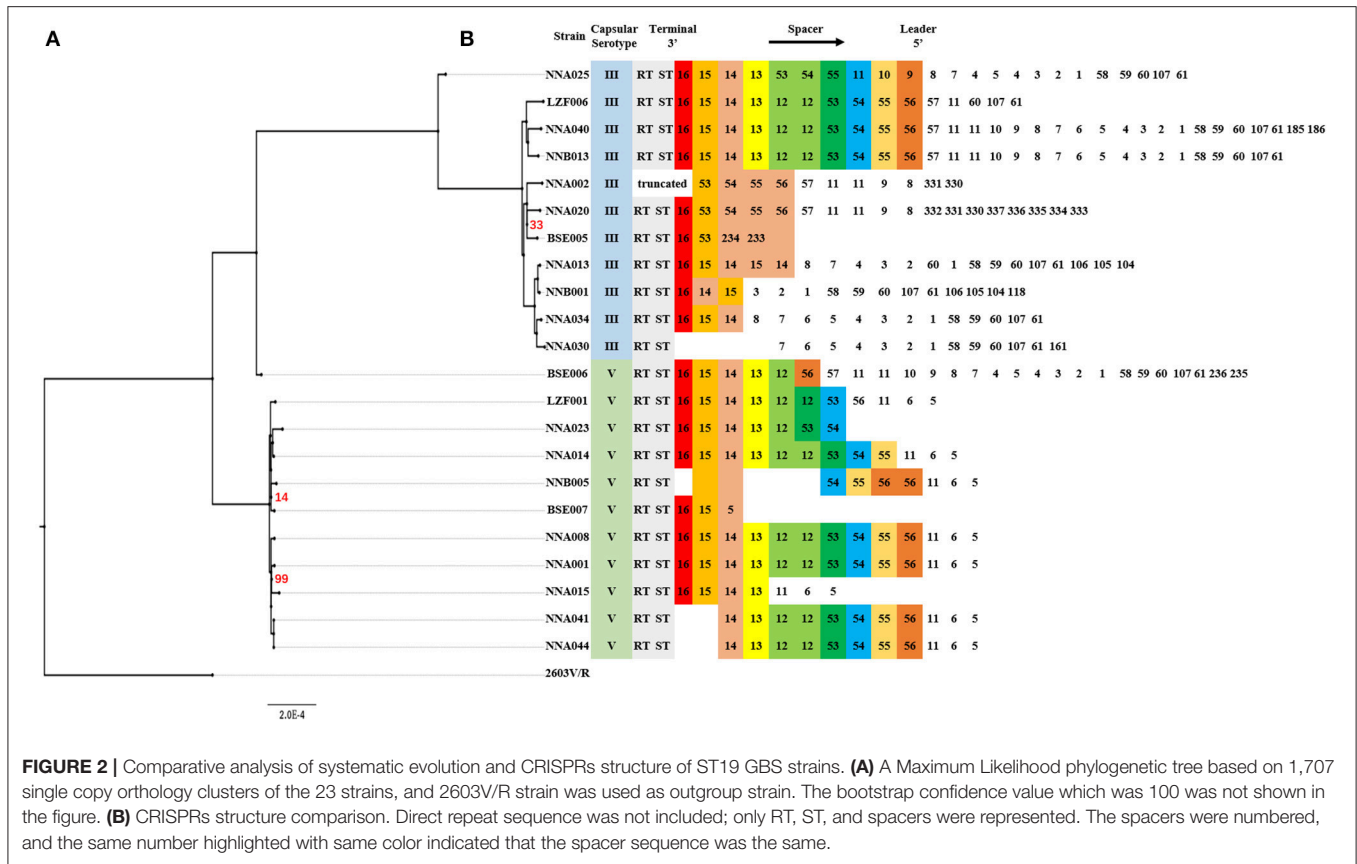
Cps Region

The results of the genome comparison showed that a large fragment of about 111 kb was involved in the variation of capsule locus region (**Figure 1B**). This large fragment was flanked by two transposons, involving 109 genes including 19 genes on *cps* operon. Except the deletion mutation in the *cps* operon, the percentage of protein sequence identity of the other genes was 97.04–100 against the reference genome. The analysis of *cps* operon showed that all sGBS were CPS type V, while all wGBS were CPS type III. The *cps* operon structures of CPS types Ia, Ib, III, and V were shown in **Figure 3**. All *cps* operons from Ia,

Ib, and III did not contain *cpsM* and *cpsN* genes. There were different degree of variation in the *cps* operons in Ia, Ib, III, and V, while the *cpsB* gene sequence of CPS type V was exactly same as Ia and Ib. The phylogenetic trees for the homologous proteins of *cps* operon from different CPS types were further constructed and it was shown that the different ST strains from the same CPS type were clustered into one branch. The human Ia strain and the fish Ia strain, as well as the human Ib strain and the fish Ib strain were clustered into one branch.

ICE Region

A total of 127 genes were involved in the ICE region of the sGBS strains and were identified with characteristic of two putative ICE, including ICESagLZF001IQ with IQ element and ICESagLZF001erm with *erm* fragment. IQ element carried two associated genes *mef(I)* and *catQ*, which were macrolides resistant gene, and chloramphenicol resistance gene of *Streptococcus*, respectively (26–28). *erm* fragment carried *Erm(TR)*, which was an *erm(A)* subclass gene discovered by Seppala et al. (29) and has become an increasingly popular macrolides resistance gene of *Streptococcus* (30).



ICESagLZF001IQ was 28,763 bp in size, had a G + C content of 36.12% and consisted of 54 putative open reading frames (ORFs). Its genetic organization was characterized using ICESpn529IQ (GenBank accession: HG965092) (31) as the genetic reference (Figure 4A). As in ICESpn529IQ, the recombination related proteins, such as integrase, relaxase, and recombinase were encoded, an IQ elements and a TA pair were contained in ICESpn529IQ. The products of *mef(I)* and *catQ* displayed 99 and 100% protein identity to the corresponding amino acid sequences of ICESpn529IQ, respectively. The TA pair coding for a toxin-antitoxin (TA) system includes the toxin gene *pezT* encoding a protein that is homologous to the zeta toxin, and the antitoxin gene *PezA*. The TA pair of ICESagLZF001IQ showed 100% protein identity to the corresponding TA pair of ICESpn529IQ. The details of the ICE region comparison between ICESag5H001IQ and ICESpn529IQ were shown in Supplementary Table 2.

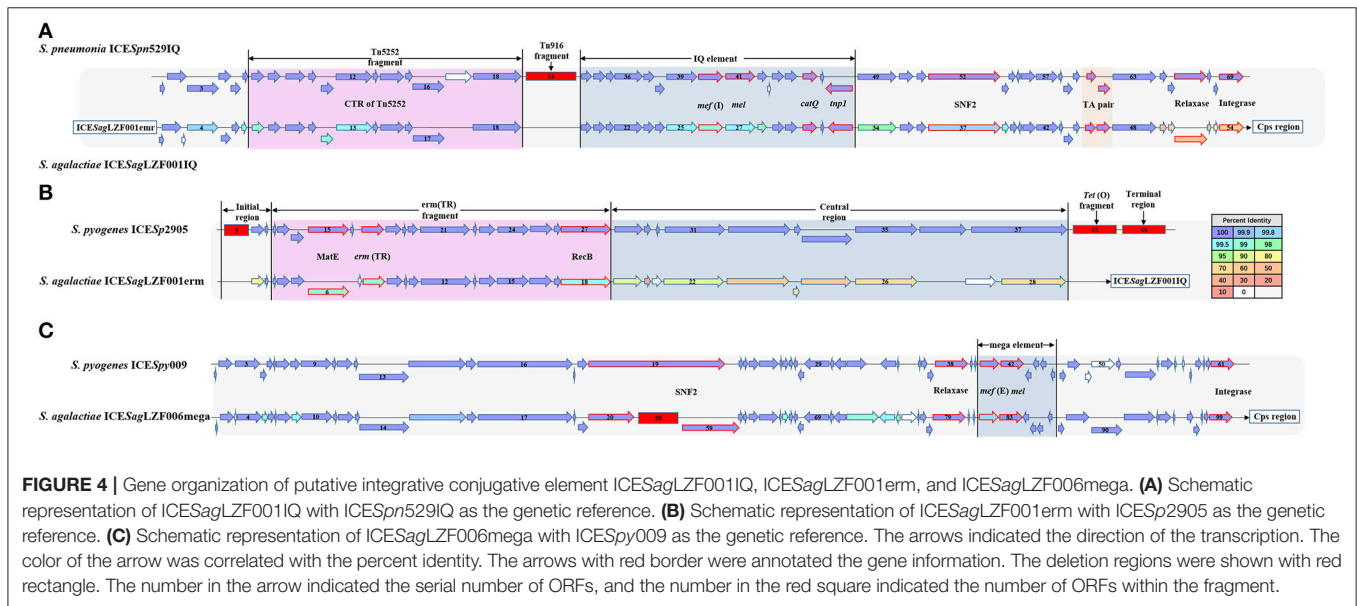
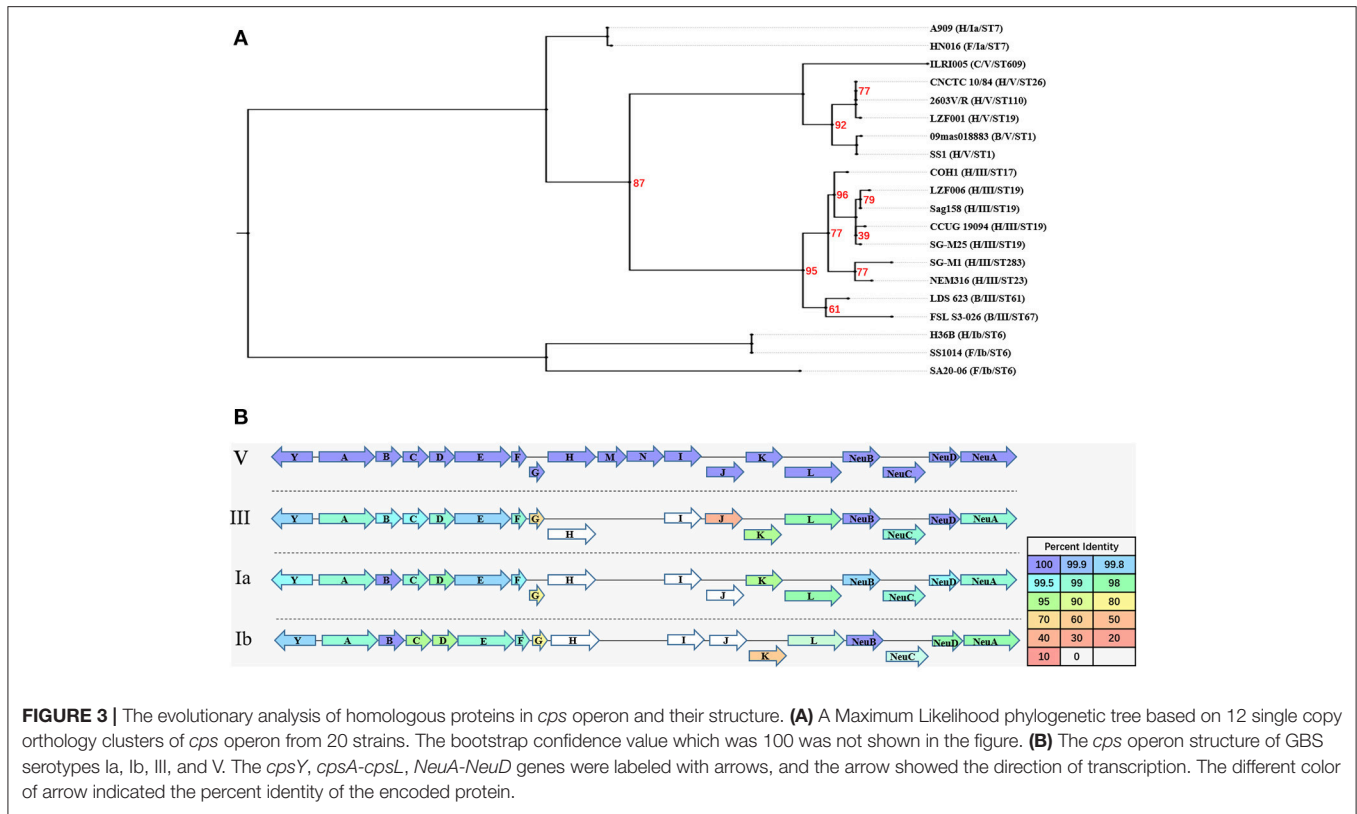
ICESagLZF001erm was 36,531 bp in size, had a G + C content of 32.20% and consisted of 28 putative ORFs. Its genetic organization was characterized using ICESp2905 (GenBank accession: FR691055) (30) as the genetic reference (Figure 4B). As in ICESp2905, an initial region, entire *erm(TR)* fragment, and central region were contained in ICESagLZF001erm. The *erm(TR)* fragment contained *emr* (TR) encoding erythromycin resistance methylase and *MatE* gene encoding multi antimicrobial extrusion protein. Except *MatE*,

erm(TR), and recombinase (*RecB*), the products of the other 12 genes in *erm(TR)* fragment showed 100% protein identity to the corresponding amino acid sequences of ICESp2905. The details of the ICE region comparison between ICESagLZF001erm and ICESp2905 were shown in Supplementary Table 3.

The ICE region of the wGBS strains included 102 genes and was identified with the characteristic of one putative ICE, ICESagLZF006mega, with mega element. The mega element carried macrolides resistant gene *mef* (E) (26, 27). ICESagLZF006mega was 106,400 bp in size, had a G + C content of 36.09% and consisted of 99 putative ORFs. Its genetic organization was characterized using ICESpy009 (GenBank accession: KU056701) (26) as the genetic reference (Figure 4C). As in ICESpy009, integrase, relaxase, SNF2, and other ICE recombination related proteins were encoded, and a mega element containing *mef(I)* and *mel* was carried in ICESagLZF006mega. In addition, the SNF2 gene of ICESagLZF006mega was truncated and a fragment containing 38 ORFs was inserted. The details of the ICE region comparison between ICESagLZF006mega and ICESpy009 were shown in Supplementary Table 4.

DISCUSSION

The serotype of the strains was confirmed by molecular serotyping in our other study (13). In this study, we have found



that the capsular switching (III–V) and transformation of ICE adjacent to the *Cps* region occurred in human GBS ST19 with different pathogenicity to fish, which may be related to the capability of GBS cross-species infection.

Although a variety of DNA-based GBS population diversity analysis have clearly indicated the existence of GBS capsular switching (32–35), there was no report of capsular switching

analysis based on comparative genomics and the relationship with cross infection. In the present study, the results of the pan-genome, CRISPRs analysis and phylogenetic reconstruction of ST19 GBS all suggested that CPS type V sGBS emerged through capsular switching from CPS type III and BSE006 seemed be an intermediate strain in the switching. The whole genome comparative analysis showed that the high variations

were observed not only in *cps* operon between CPS type V and III strains, but also in the flanking regions of *cps* operon, suggesting that this region likely underwent large fragment recombination. In addition, a capsular switching from CPS type III–IV within CC17 GBS also occurred due to the exchange of a 35.5 Kb DNA fragment containing the entire *cps* operon (32), indicating that the recombination of large DNA fragments including *cps* operon is an important way of GBS capsular switching. The major CPS types of GBS that currently infect fish are CPS types Ia and Ib (36). In this study, the results showed the *cpsB* gene in the *cps* operon of the ST19 CPS type V was identical to that of the Ia and Ib. The capsules can affect the adherence of GBS to fish intestinal epithelial cells (37), and CpsB participates in determining the length and adhesion of GBS capsules (38). In summary, the capsules, especially CpsB, may play an important role in GBS virulence and cross-species infection.

Clustered regularly interspaced short palindromic repeats (CRISPRs) are a bacterial adaptive immune defense mechanism for defending against foreign nucleic acids, such as ICE, plasmids, and phages (39). CRISPR arrays are composed of a peculiar family of DNA repeats which are usually multiple, non-contiguous, direct DNA repeats interspaced by non-repetitive nucleotides (spacer) (40). When foreign nucleic acids invades bacteria, the CRISPRs integrate and save the intruding gene fragment (spacer). Under the re-invasion of the same genes, mediated by specific RNA, CRISPRs, and CRISPR-associated proteins (Cas proteins) will cut and destroy the invading foreign nucleic acids (41). The repeats are highly conserved within a given CRISPR array, and the spacers are correspond to segments of captured viral and plasmid sequences in most cases (42). The differences of the CRISPRs structure implied different capabilities between wGBS and sGBS to defend against foreign nucleic acids, which may affect their adaptive capacity in different host. The strains of sGBS derived from a recent common ancestor as the CRISPRs structure of sGBS was more consistent than that of wGBS. The mechanism of the recombination of large DNA fragments, capsular switching and differences of CRISPRs structure are not completely understood. Capsular switching and different CRISPRs structure are probably driven by the host immune response and the living conditions of the bacteria. Therefore, study of characteristics of immunity and *in vivo* environment of the hosts, which sGBS and wGBS were isolated from, would help to clarify the relevant mechanism.

ICE are widely distributed mobile DNA elements that can be vertically propagated by integrating into the host genome or can be transferred horizontally by excision and transfer of elements to new hosts. We identified an ICE variant region associated with GBS ST19 capsular switching through the whole genome

comparison. With the transformation of ST19 GBS CPS type from III to V, the ICE in this region was also transformed from the original ICE to the new one. Similarly to the capsular switching, it is possible that the ICE transformation is a result from the adaptation of bacteria to the host immunity and environment. In CPS type III strains, the region is located at the 3' terminus of ST19 GBS Cps region, and contains an ICE carrying mega element. In contrast, this region in CPS type V contains two ICEs, carrying IQ element and the *erm* (TR) fragment, respectively. The resistant genes in mega element, IQ element, and *erm* (TR) fragment are all macrolides resistance genes (26, 27, 29, 30). The effect of different macrolides resistance genes on GBS adaptability remains to be further studied.

The large-scale aquaculture of fish and the high carrying rate of GBS in human result in an increased frequency of genome exchange intra- and inter-species with different hosts, which increases the risk of producing cross-species infection strains or highly adaptable strains. Our data indicated that capsular switching and ICE transformation might improve the ability of GBS cross-species infection, which resulted in an increased difficulty for preventing GBS diseases. It should be emphasized that great attention should be paid to monitor the colonization and disease of GBS in practitioners from fish farms, and study the associations among GBS strains with different hosts, which will be important for the prevention and control of GBS disease.

AUTHOR CONTRIBUTIONS

RW and LL analyzed data and wrote the manuscript; TH, WH, and AL performed experiments and analyzed data; MC conceived and designed the study. All authors reviewed and approved the manuscript.

FUNDING

This work was supported by National Natural Science Foundation of China (31460695), Guangxi innovation-driven development special funds (Grant no. AA17204081-3), Guangxi Natural Science Foundation (2016GXNSFDA380020), the funds of Guangxi Key Laboratory for Aquatic Genetic Breeding and Healthy Aquaculture (2016-2018).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fvets.2018.00281/full#supplementary-material>

REFERENCES

- Lambertsen L, Ekelund K, Skovsted IC, Liboriussen A, Slotved HC. Characterisation of invasive group B streptococci from adults in Denmark 1999 to 2004. *Eur J Clin Microbiol Infect Dis.* (2010) 29:1071–7. doi: 10.1007/s10096-010-0941-z
- Evans JJ, Klesius PH, Pasnik DJ, Bohnsack JF. Human *Streptococcus agalactiae* isolate in Nile tilapia (*Oreochromis niloticus*). *Emerg Infect Dis.* (2009) 15:774–6. doi: 10.3201/eid1505.080222
- Chen M, Wang R, Luo FG, Huang Y, Liang WW, Huang T, et al. *Streptococcus agalactiae* isolates of serotypes Ia, III and V from human and cow are able to infect tilapia. *Vet Microbiol.* (2015) 180:129–35. doi: 10.1016/j.vetmic.2015.07.033

4. Delannoy CM, Crumlish M, Fontaine MC, Pollock J, Foster G, Dagleish MP, et al. Human *Streptococcus agalactiae* strains in aquatic mammals and fish. *BMC Microbiol.* (2013) 13:41. doi: 10.1186/1471-2180-13-41
5. Ma K, Cao Q, Luo S, Wang Z, Liu G, Lu C, et al. Cas9 enhances bacterial virulence by repressing the *regR* transcriptional regulator in *Streptococcus agalactiae*. *Infect Immun.* (2017) 86:e00552-17. doi: 10.1128/IAI.00552-17
6. Rajendram P, Mar Kyaw W, Leo YS, Ho H, Chen WK, Lin R, et al. Group B streptococcus sequence type 283 disease linked to consumption of raw fish, Singapore. *Emerg Infect Dis.* (2016) 22:1974-7. doi: 10.3201/eid2211.160252
7. Foxman B, Gillespie BW, Manning SD, Marrs CF. Risk factors for group B streptococcal colonization: potential for different transmission systems by capsular type. *Ann Epidemiol.* (2007) 17:854-62. doi: 10.1016/j.annepidem.2007.05.014
8. Tan K, Wijaya L, Chiew HJ, Sitoh YY, Shafi H, Chen RC, et al. Diffusion-weighted MRI abnormalities in an outbreak of *Streptococcus agalactiae* Serotype III, multilocus sequence type 283 meningitis. *J Magn Reson Imaging* (2017) 45:507-14. doi: 10.1002/jmri.25373
9. Mian GF, Godoy DT, Leal CA, Yuhara TY, Costa GM, Figueiredo HC. Aspects of the natural history and virulence of *S. agalactiae* infection in Nile tilapia. *Vet Microbiol.* (2009) 136:180-3. doi: 10.1016/j.vetmic.2008.10.016
10. Brochet M, Rusniok C, Couvé E, Dramsi S, Poyart C, Trieu-Cuot P, et al. Shaping a bacterial genome by large chromosomal replacements, the evolutionary history of *Streptococcus agalactiae*. *Proc Natl Acad Sci USA.* (2008) 105:15961-6. doi: 10.1073/pnas.0803654105
11. Wang X, Cao X, Li S, Ou Q, Lin D, Yao Z, et al. Phenotypic and molecular characterization of *Streptococcus agalactiae* colonized in Chinese pregnant women: predominance of ST19/III and ST17/III. *Res Microbiol.* (2018) 169:101-7. doi: 10.1016/j.resmic.2017.12.004
12. Zadoks RN, Middleton JR, McDougall S, Katholm J, Schukken YH. Molecular epidemiology of mastitis pathogens of dairy cattle and comparative relevance to humans. *J Mammary Gland Biol Neoplasia* (2011) 16:357-72. doi: 10.1007/s10911-011-9236-y
13. Li L, Wang R, Huang Y, Huang T, Luo F, Huang W, Yang X, et al. High incidence of pathogenic *Streptococcus agalactiae* ST485 strain in pregnant/puerperal women and isolation of hyper-virulent human CC67 strain. *Front Microbiol.* (2018) 9:50. doi: 10.3389/fmicb.2018.00050
14. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. ABySS: a parallel assembler for short read sequence data. *Genome Res.* (2009) 19:1117-23. doi: 10.1101/gr.089532.108
15. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* (2004) 14:1394-403. doi: 10.1101/gr.2289704
16. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep.* (2015) 5:8365. doi: 10.1038/srep08365
17. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* (2015) 31:3691-3. doi: 10.1093/bioinformatics/btv421
18. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* (2015) 3:e1319. doi: 10.7717/peerj.1319
19. Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* (2007) 35:W52-7. doi: 10.1093/nar/gkm360
20. Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpides NC, et al. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* (2007) 8:209. doi: 10.1186/1471-2105-8-209
21. Lier C, Baticle E, Horvath P, Haguenoer E, Valentin AS, Glaser P, et al. Analysis of the type II-A CRISPR-Cas system of *Streptococcus agalactiae* reveals distinctive features according to genetic lineages. *Front Genet.* (2015) 6:214. doi: 10.3389/fgene.2015.00214
22. Li L, Stoeckert CJ Jr., Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* (2003) 13:2178-89. doi: 10.1101/gr.1224503
23. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* (2011) 27:1164-5. doi: 10.1093/bioinformatics/btr088
24. Stamatakis A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* (2014) 30:1312-3. doi: 10.1093/bioinformatics/btu033
25. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* (2014) 42:D206-14. doi: 10.1093/nar/gkt1226
26. Del Grosso M, Camilli R, Rizzi E, Pietrelli A, De Bellis G, Pantosti A. ICESpy009, a conjugative genetic element carrying *mef(E)* in *Streptococcus pyogenes*. *Antimicrob Agents Chemother.* (2016) 60:3906-12. doi: 10.1128/AAC.03082-15
27. Mingoia M, Vecchi M, Cochetti I, Tili E, Vitali LA, Manzini A, et al. Composite structure of *Streptococcus pneumoniae* containing the erythromycin efflux resistance gene *mefI* and the chloramphenicol resistance gene *catQ*. *Antimicrob Agents Chemother.* (2007) 51:3983-7. doi: 10.1128/AAC.00790-07
28. Gay K, Stephens DS. Structure and dissemination of a chromosomal insertion element encoding macrolide efflux in *Streptococcus pneumoniae*. *J Infect Dis.* (2001) 184:56-65. doi: 10.1086/321001
29. Seppala H, Skurnik M, Soini H, Roberts MC, Huovinen P. A novel erythromycin resistance methylase gene (*ermTR*) in *Streptococcus pyogenes*. *Antimicrob Agents Chemother.* (1998) 42:257-62.
30. Giovanetti E, Brenciani A, Tiberi E, Bacciaglia A, Varaldo PE. ICESp2905, the *erm(TR)-tet(O)* element of *Streptococcus pyogenes*, is formed by two independent integrative and conjugative elements. *Antimicrob Agents Chemother.* (2012) 56:591-4. doi: 10.1128/AAC.05352-11
31. Mingoia M, Morici E, Morrioni G, Giovanetti E, Del Grosso M, Pantosti A, et al. Tn5253 family integrative and conjugative elements carrying *mef(I)* and *catQ* determinants in *Streptococcus pneumoniae* and *Streptococcus pyogenes*. *Antimicrob Agents Chemother.* (2014) 58:5886-93. doi: 10.1128/AAC.03638-14
32. Bellais S, Six A, Fouet A, Longo M, Dmytruk N, Glaser P, et al. Capsular switching in group B *Streptococcus* CC17 hypervirulent clone: a future challenge for polysaccharide vaccine development. *J Infect Dis.* (2012) 206:1745-52. doi: 10.1093/infdis/jis605
33. Brochet M, Couvé E, Zouine M, Vallaey T, Rusniok C, Lamy MC, et al. Genomic diversity and evolution within the species *Streptococcus agalactiae*. *Microbes Infect.* (2006) 8:1227-43. doi: 10.1016/j.micinf.2005.11.010
34. Martins ER, Melo-Cristino J, Ramirez M. Evidence for rare capsular switching in *Streptococcus agalactiae*. *J Bacteriol.* (2010) 192:1361-9. doi: 10.1128/JB.01130-09
35. Luan SL, Granlund M, Sellin M, Lagergard T, Spratt BG, Norgren M. Multilocus sequence typing of Swedish invasive group B streptococcus isolates indicates a neonatally associated genetic lineage and capsule switching. *J Clin Microbiol.* (2005) 43:3727-33. doi: 10.1128/JCM.43.8.3727-3733.2005
36. Delannoy CM, Zadoks RN, Crumlish M, Rodgers D, Lainson FA, Ferguson HW, et al. Genomic comparison of virulent and non-virulent *Streptococcus agalactiae* in fish. *J Fish Dis.* (2016) 39:13-29. doi: 10.1111/jfd.12319
37. Barato P, Martins ER, Vasquez GM, Ramirez M, Melo-Cristino J, Martinez N, et al. Capsule impairs efficient adherence of *Streptococcus agalactiae* to intestinal epithelium in tilapias *Oreochromis sp.* *Microb Pathog.* (2016) 100:30-6. doi: 10.1016/j.micpath.2016.08.040
38. Toniolo C, Balducci E, Romano MR, Proietti D, Ferlenghi I, Grandi G, et al. *Streptococcus agalactiae* capsule polymer length and attachment is determined by the proteins CpsABCD. *J Biol Chem.* (2015) 290:9521-32. doi: 10.1074/jbc.M114.631499
39. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* (2010) 327:167-70. doi: 10.1126/science.1179555

40. Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet.* (2011) 45:273–97. doi: 10.1146/annurev-genet-110410-132430
41. Deveau H, Garneau JE, Moineau S. CRISPR/Cas system and its role in phage-bacteria interactions. *Annu Rev Microbiol.* (2010) 64:475–93. doi: 10.1146/annurev.micro.112408.134123
42. Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, Moineau S, et al. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol.* (2008) 190:1401–12. doi: 10.1128/JB.01415-07

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Wang, Li, Huang, Huang, Lei and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.