Application of Omics and Bioinformatics Tools in *Streptococcus* Research

Min Liao[1,2#], Ting Tong[1,2#], Yawen Zong[1,2#], Xuedong Zhou[1], Lei Cheng[1,2], Ruijie Huang[1,3*], Biao Ren[1*], Gil Alterovitz[4]

1. State Key Laboratory of Oral Diseases and National Clinical Research Center for Oral Diseases, West China Hospital of Stomatology, Sichuan University
2. Department of Operative Dentistry and Endodontics, West China Hospital of Stomatology, Sichuan University
3. Department of Pediatric Dentistry, West China School of Stomatology, Sichuan University.
4. Harvard Medical School, Boston Children's Hospital, Boston, MA, USA

#: These authors contributed equally to this work.
*Corresponding authors
Ruijie Huang: hwangrj@163.com
Biao Ren: renbiao@scu.edu.cn

Abstract

Researchers used to focus on analyzing single gene or protein expression of the microbes. But recently, genome, transcriptome, proteome and metabolome have gained more and more attention. Based on technologies of omics, including genomics, transcriptomics and metabolomics, a large quantity of information about cells, microbes and human, such as the information about phylogeny, virulence, antibiotic resistance and other aspects, has been revealed. Genus *Streptococcus* is one of the most invasive groups of bacteria that cause both human and animal diseases, threatening public health. In this review, we summarize the application of omics to analyze this genus-*Streptococcus.*

Introduction

Over the past century, biotechnology has developed rapidly and provided possibilities to determine the function of micro composition of life and describe life phenomena in a micro or macro perspective. Omics, which is a multiple biotechnology based on the genomics, focuses on global and high-throughput analytical methods to investigate the mechanisms of life evolution and the role of genes (Yan et al., 2015). Saving genomics, omics includes proteomics, transcriptomics, and metabolomics. Meanwhile, a large quantity of information, such as the information about phylogeny, virulence, antibiotic resistance and other aspects of microbes, has been revealed about cells, microbes, and human in the past two decades. With abundant data and increasing accuracy, omics is applied in many life science aspects including microorganism. Genus *Streptococcus* is one of the most invasive groups of bacteria that causes both human and animal's disease, threatening human's health, especially *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Streptococcus mutans* and *Streptococcus suis*. Taking this genus as the research object, omics and bioinformatics tools provide a systematic path to cognize the law of biological activities of microorganisms.

In 1995, the first complete genomic sequence, as also the first completed bacterial genome sequence, was obtained from *Haemophilus influenzae* Rd (Fleischmann et al., 1995). Subsequently, researchers have completed sequencing of hundreds of microbes. With the improvement of sequencing technology and accuracy of computer model analysis, omics enables us to understand the mechanism of molecular function and interaction. As for *Streptococcus*, the known genome of Streptococcal strains varies in size from about 0.36 to 9.62 Mbp. The genus *Streptococcus* possesses relatively low GC content, which may attribute to mutation and various selections (Gao et al., 2014). Related to genomics, proteomics mainly focuses on the bacterial surface proteome, a major source for identifying vaccine targets and relating to serological assays. For example, Jimenez-Munguia et al. (2015) found that a kind of cell wall surface anchor family protein was significantly lower in pneumococcal pneumonia

children compared to those in controls. However, Kreth et al. (2011) found that the balance of transcriptome expression of *Streptococcus* was regulated by special genes.

Overall, omics technology has been applied to many fields of microbe investigation, such as molecular function, identification and immune response to drugs. Simultaneously, the development of omics has advanced the development of researches on the origin of life, evolution, disease occurrence and drug industry. However, there are still plentiful challenges, such as the need of standard database and convenient analysis software, systematic research processes and macroscopic application. In this review, we summarize the systematic application of omics technology to analyze genus-*Streptococcus* (Figure 1).

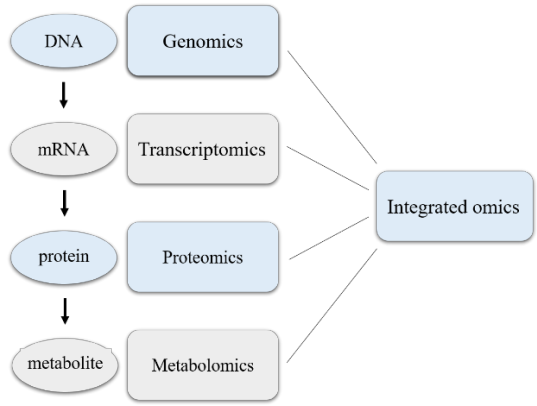

Figure 1. Schematic diagram of different omics tools targeting different molecular layers of *Streptococcus*.

Genomics

Genomics, a branch of genetics, is a popular subject based on whole gene sequencing (WGS) technology and bioinformatic tools. Since the first complete genome of free-living organism, *Haemophilus influenzae* Rd, was sequenced in 1995 (Fleischmann et al., 1995), DNA sequencing technology has experienced a rapid development over the next few decades. In 2005, a profound revolution had been brought about by a new technology in this filed, namely “next-generation sequencing (NGS)”, also called high-throughput sequencing, which made genomic sequencing much more accessible with relatively lower expense and

higher efficiency since then. Nowadays, many new companies and equipments in this field have sprung up, including Illumina, Oxford Nanopore, Pacific Biosciences, and etc. (Levy and Myers, 2016). They can not only provide genome organization of a selected subject, but also make a comprehensive analysis based on the sequencing data. Thanks to it, genomics can be widely implemented to explore the secret of different lives on earth, including microorganisms, human, plants, and etc.

Genome Characteristics of genus Streptococcus

With the development of sequencing, new whole genomes are sequenced completely and reported almost weekly with higher speed and lower cost. Today, the completed genomes of nearly 12,500 Streptococcal strains can be found at the National Center for Biotechnology Information (NCBI) website, with frequent data update. The genome of Streptococcal strains varies in size from about 0.36 to 9.62 Mbp. The genus *Streptococcus* possesses relatively low GC content, which may attribute to mutation and various selection (Gao et al., 2014). The pan-genome (all genes present in the genomes of the genus) size increases with the addition of newly sequenced strains, indicating that this genus has an open pan-genome. Conversely, the core-gene size (common genes shared by all Streptococcal strains) size tends to reduce (Gao et al., 2014; Liu et al., 2013b).

Comparative genomics

With an increasing number of sequenced Streptococcal whole genomes, comparative genomics is widely conducted to illustrate the functions of genes by comparing the genome of interested isolates with the reference genomes available in public database and using bioinformatic tools. Through comparing intraspecies variations or closely related species variations (Gao et al., 2014), it can provide us with comprehensive understanding of this genus and its connections between distinct phenotype and genetic characteristics, including the evolutionary relationship of *Streptococcus*, virulence determinants, as well as the genetic basis underlying adaptive changes to selective pressure and different hosts. Analysis of genome sequences reveal the dynamic nature of genetic

material, and existence of multiple genetic events, such as genome decay, genome rearrangement and acquisition of new genes by horizontal gene transfer (HGT) (Fraser-Liggett, 2005; Lefébure and Stanhope, 2007; Muzzi and Donati, 2011), which have a profound influence on the evolutionary process and decide the phenotype to some extent.

Environmental adaptation

The environment that bacteria live in has a strong impact on the bacterial evolution over time, and genomics tools can help to investigate the genetic basis of their adaptation to habitat (Papadimitriou et al., 2014; Prajapati, 2013). A representative example of environmental adaptation is about the dairy starter strain and the nutritious dairy they live in. For instance, *Streptococcus thermophilus*, a common dairy starter with outstanding fermentation capacity, is widely used in dairy industry. The mechanism of fermentation property and food safety attracts significant attention and has been explored genetically. *S. thermophilus*, as a member of salivarius group, shows certain differences in genomic characteristics compared with other two genetically similar group members, *Streptococcus salivarius* and *Streptococcus vestibularis*. The *S. thermophilus* has a smaller genome size of about 1.7Mb and a relatively low level of nucleotide variability (Delorme et al., 2015). Notably, most Streptococcal virulence-associated genes are either presented as pseudogenes or missing in this species (Bolotin et al., 2004; Prajapati, 2013), suggesting that it is nonpathogenic. For instance, antibiotic resistance is an importance character of pathogenicity, but there is no antibiotic resistance genes detected in *S. thermophilus* genome. Similarly, genome decay and absence of classical Streptococcal virulent factors are detected in *Streptococcus infantarius* subsp. *infantarius* and the *Streptococcus macedonicus* isolated from fermented dairy (Jans et al., 2013; Papadimitriou et al., 2014). It seems that narrow and nutritious dairy habitat induces Streptococcal genome to experience a reductive evolution (Delorme et al., 2015; Papadimitriou et al., 2014). Meanwhile, *S. thermophilus* possesses some functional genes that seem to be beneficial to human when the bacteria adapt to milk. This species is reported to possess encoding genes for

probiotic features, like adhesion, acid resistance, producing bacteriocin and utilizing lactose (Prajapati, 2013). Recently, an isolate that has a novel eps gene cluster can produce the highest level of exopolysaccharide (EPS) among this species and produce both capsular and ropy EPS, making the viscosity and texture of dairy better (Wu et al., 2014). Four separate clustered regularly interspaced short palindromic repeats (CRISPR)/Cas loci found in its genome enable it to resist unwanted bacteriophage infection. These researches support the use of *S. thermophilus* in manufacture of yogurt and cheese and probably wider use in biotechnological applications (Wu et al., 2014).

Conversely, the adaption to environment of pathogens always involves in complex regulation of virulence rather than merely attenuation of virulence. Lefebure et al. (2012) compared genomes of *S. pyogenes*, a common human pathogen, with several closely related species, *Streptococcus dysgalactiae*, *Streptococcus equi*, and *Streptococcus canis*. It turned out that the *S. pyogenes* gained 113 genes during evolution and about 14 genes were related to virulence, including *sal* antibiotic locus, *speB* gene, *rgg* gene, and etc. The *rgg* can regulate the expression of *speB* and interact with it, and this interaction may be beneficial to *S. pyogenes* colonization in human pharynx without dispersing to cause invasive diseases (Dmitriev et al., 2006). Both the acquisition of virulence genes and the establishment of regulation network contribute to the bacterial adaption to human host (Lefebure et al., 2012).

Virulence factors

Many species of genus *Streptococcus* are critical pathogens in both human and animal diseases, and their virulence plays an important role in infection. Through comparing the genome of interested strains with the annotated sequences in public database by using bioinformatics tools, a number of virulence genes may be detected in the interested strains. The identified conserved Streptococcal virulence determinates are mainly involved in adhesion, invasion, spreading, and other related regulatory genes, which can be inferred in virulence-related public database, such as Virulence Factors of the Pathogenic Bacteria Database (Gao

et al., 2014; Olson et al., 2013). Some other genetic elements derived from HGT also belong to Streptococcal virulence. For instance, 12 characterized prophages in *S. suis* probably contribute to the pathogenicity of *S. suis* for carrying putative virulence-related genes, like toxin–antitoxin system, Clp protease and DNA methyltransferase (Tang et al., 2013). A virulent *Streptococcus zooepidemicus* strain, Sz35246, is reported to acquire four pathogenicity islands (PAIs) through HGT, including three toxin-antitoxin (TA) systems PAIs and a restriction modification system (RM system) PAI (Ma et al., 2013).

A recent research compared genomes of 3615 M1 Group A Streptococcus (GAS) strains isolated within about the last 100 years all around the world, revealing that the contemporary epidemic strain experienced a series genetic evolutionary events to form a high virulent strain. The precursor cell just contained one phage carrying an extracellular DNase gene, and then it acquired a phage encoding SpeA1. After that, it experienced a single-nucleotide change to form SpeA2 variant. Next, the species obtained a 36-kb region encoding secreted toxins NAD+-glycohydrolase and streptolysin O through HGT, shaping the high-virulent strain, MGAS5005-like M1 strains that spread worldwide recently (Nasser, 2014). However, Fiebig et al. (2015) found that the hypervirulence of *S. pyogenes* M1 is not restricted to the MGAS5005-like genotype. They compared a MGAS5005-like strain, 5884, with AP1. Although the AP1 is not MGAS5005-like strain, it can cause systemic infection and a quick death in infectious mouse (Oehmcke et al., 2009; Oehmcke et al., 2013). Compared to the 5448 strain, the higher virulent strain AP1 had its CRISPR-Cas system destroyed by the phage insertion, which attenuated its defense capacity against intrusive genetic element and caused an enrichment of phages. Besides, the AP1 genome has an additional protein H encoding gene and a mutation of *rofA* gene, which may influence its regulatory networks and virulence, but whether these characteristics contribute to its virulence deserves further researches (Fiebig et al., 2015). All these results suggest that the virulence determinants are dynamics and influenced extensively by the gene gain, gene loss and genetic recombination events.

Genomics analysis can be used to compare genome sequence focusing on different virulence determinants and may explain the genetic basis of distinct virulence and pathogenesis of different strains to some extent. However, this does not work every time. Sometimes, different clinical outcomes cannot be explained by the bacterial genetic differences associated with virulence. For instance, no clear difference of gene content was detected from The Streptococcus Anginosus Group (SAG) strains isolated from both respiratory and invasive infections (Olson et al., 2013). Likewise, Argimon et al. (2014) compared the genome of *S. mutans* from severe early childhood caries (S-ECC) and caries-free (CF) child using a new genome-scale in silico subtractive hybridization. Although they quickly identified the genetic loci unique to strains associated with S-ECC, they failed to find any specific genes related to caries, suggesting that there is no virulence determinants difference of *S. mutans* from S-ECC or SF children responsible for the distinct clinical outcomes. Moreover, Hossain et al. (2015) compared genomes of virulent and avirulent *Streptococcus uberis* strains and found that the potential virulence genes are detected in both strains and highly variable with high level of SNPs, indicating that the presence of virulence genes merely is hardly to explain the virulence capacity and infection status. There is a more complex dynamic that determines the infection status and clinical outcomes, and this is more than one simple genetic component. More comprehensive analysis of multiple factors need to be carried out, including genetic transcription, protein translation, environmental effect, interaction of various microorganisms, and etc.

Antibiotic resistance

With the increasing use of antibiotics, many strains that are susceptible to antibiotics primitively become resistant and widespread, making it difficult to control of the epidemics of infection. A bacterial resistance surveillance in USA shows that the *S. pneumoniae* isolated from patient infection site shows increased rates of insusceptibility to several antibiotics, including amoxicillin/clavulanate, erythromycin, and levofloxacin (Pfaller et al., 2012). Analysis of bacterial genome sequences enables us to identify the putative antibiotic resistance genes and may provide guidance for clinical treatment. Several antibiotic resistance gene

databases have been established, including ARGO, MvirDB, ARDB, Resfinder, CARD and ARG-ANNOT, of which the last three are relatively detailed and comprehensive. The powerful and versatile ARG-ANNOT, released in 2013, contains more than 1600 antibiotic resistance genes and is widely used in bacterial drug resistance researches (Gupta et al., 2014). Metcalf et al. (2017) used short-read whole genome sequencing combined with ARG-ANNOT and ResFinder databases to detect multiple resistance determinants in invasive group B Streptococci (iGBS), and this was highly accurate, economical and practical for large-scale analysis compared to phenotypic testing.

Bacteria can obtain resistance genes mainly through mutations and HGT from other microbes over the course of evolution, especially under the selection pressure of antibiotics (Blair et al., 2015; Huddleston, 2014). The main types of HGT include conjugation, natural transformation and transduction (Huddleston, 2014). *S. pneumoniae* is a common pathogen that shows high rate of recombination. The Molecular Epidemiology Network clone 1 (PMEN1), a multi-drug resistant *S. pneumoniae* lineage, is reported to possess multiple genetic elements related to antibiotic resistance, including integrative and conjugative element (ICE), plasmids and other element with resistance genes (Croucher et al., 2009). Croucher et al. (2011) analyzed 240 sequenced strains of this lineage from all over the world and found distinguished different ways of resistance genes acquisition. The fluoroquinolone and rifampicin resistance mainly resulted from gene mutation, while resistance to non-β-lactam derived from recombination. Three different genetic elements were integrated into the Tn916 transposon, including Tn917 transposon (containing an *ermB* gene), mega element (containing a *mef*/*mel* gene) and omega element (containing an *ermB* gene and an aminoglycoside phosphotransferase). Those elements encode products involved in methylation of target ribosomal RNA, drug efflux pump system, and enabling the PMEN1 insensitive to macrolide. These elements are acquired by the strains across the established phylogeny more than once, and the macrolide-sensitive strains are replaced by the insusceptible strains, suggesting a genetic plasticity of this lineage during adaption to the selection pressure of clinical intervention.

Phylogeny

Phylogenetic analysis uses substantial genomic data and various analysis software to build a phylogenetic tree, describing the phylogenetic relationships of target species visually. In general, phylogenetic tree is built based on certain genetic loci, such as housekeeping genes or some genes of interest (Olson et al., 2013; Teng et al., 2014) as well as multiple genetic loci or even the whole genome sequences. The tree built on single gene, such as virulence genes, and this can throw light on the evolutionary history of the specific factor at species or genus level (Argimon et al., 2013; Gao et al., 2014). Glucosyltransferases (Gtfs) is an important enzyme in *Streptococcus* to catalyze the synthesis of glucans. The GtfB catalysate, a type of water-insoluble glucans (WIG), is beneficial for the formation of dental plaque and the adhesion of bacteria. Thus, it influences the development of dental caries and periodontitis (Mattos-Graner et al., 2004; Munro et al., 1991). Researchers conducted phylogenetic analysis among the *gtfs* genes of strains isolated from oral cavity of human, chimpanzees and macaque monkeys. They found that the gene duplication of *gtfB* and *gtfC* and the divergency of primates took place almost at the same period. Besides, the divergency of viviparous occurred before the expression of *S. mutans gtfB* and *gtfC*, which regulates the WIG synthesis. This phenomenon indicates that the g*tfs* diversification precedes human, and it is unrelated to the changes in human diet (Argimon et al., 2013).

16S rRNA is a popular marker used in phylogenetic analysis, and the majority of Streptococcal species are clustered into six major clades based on the 16S rRNA gene phylogenetic tree accordingly. Those species are classified into six “species group”, named: “Pyogenic”, “Mitis”, “Anginosus”, “Bovis”, “Mutans” and “Salivarius” (Gao et al., 2014; Richards et al., 2014; Stackebrandt et al., 2002). Yet, there are limitations in this analysis and classification. For instance, sometimes different species may share identical 16S rRNA sequence, like the *Streptococcus mitis* and *Streptococcus oralis* (Suzuki et al., 2005). This identical 16S rRNA sequence makes it difficult to distinguish different species, and besides, the exist of

horizontal transfer of segments of the 16S rRNA genes among species of the anginosus group makes the result inaccurate (Schouls et al., 2003). Hence, an established database based on larger volume of molecular data is needed. Phylogenomic analysis, a preferable method, is widely adopted by scientists. Phylogenomic analysis is based on multiple locus and possesses advantages of discriminate different species over other single locus analysis. Olson et al. (2013) compared the phylogenetic trees based on single locus 16S rRNA, housekeeping genes (*rpoB*, *cpn60*) and in-house core-SNP pipeline, respectively. Although these trees have similar overall topology, the multilocus approaches acquired longer branch lengths and a greater degree of statistical certainty; conversely, the 16S rRNA possesses a lower discriminatory power. Moreover, the multilocus analysis is particularly suitable for defining location of some ambiguous species. For instance, *Streptococcus sinensis* turned out to be equally close to both the anginosus and mitis groups in the phylogenetic tree built on 16S rRNA gene (Teng et al., 2014). The accurate location of *S. sinensis* on the phylogenetic trees remains ambiguous merely according to the single locus analysis. Teng et al. (2014) used phylogenomic analysis based on concatenated genes and whole genomes, and at the same time, they also used matrix-assisted laser desorption ionization-time of flight mass spectrometry to analyze the phylogenetic clade of *S. sinensis*, *Streptococcus oligofermentans* and *Streptococcus cristatus*. By these approaches, they found those three species formed a distinct phylogenetic clade, thus they proposed a new group of the three species, named “sinensis group”. These results indicate that the multilocus phylogenomic analysis plays an important role in the taxonomy and identification of *Streptococcus,* and the multilocus approach possesses distinct advantages over the single genetic locus approach.

The phylogenomic tools and the genomic data have also been extensively used in the genetic and evolutionary relationship investigation at both strain and species levels. Watanabe et al. (2013) used BEAST with a post-probabilistic approach to construct a phylogenetic tree based on whole genomes of *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) strains and estimated the

ages of the subclade, and the result revealed that the most recent common ancestor of the selected SDSE strains appeared 446 years ago. Meanwhile, the results of analysis based on a core set of genes after omitting the phage regions exhibit a remarkable consistency. A recently published global phylogenomic analysis of core genome identified distinct lineages among nonencapsulated *S. pneumoniae* (non-Ec-*Sp*) and encapsulated *S. pneumoniae* (Ec-*Sp*) strains. Most non-Ec-*Sp* isolates clustered to form a separate (deep-branching classic) lineage and few non-Ec-*Sp* strains sporadically joined in the lineage of Ec-*Sp* on the phylogenetic tree. Most isolates from this separate non-Ec-*Sp* lineage showed highly insensitive to antibiotics and possessed more mobile elements that might contribute to the increased ability of adherence, revealing that non-Ec-*Sp* isolates react differently under the selection pressure of clinical intervention, such as the use of vaccines and antimicrobial (Hilty et al., 2014).

Genome editing

CRISPR/Cas system was first characterized over a decade ago as an adaptive immune system in genomes of various bacteria and archaea to eliminate extrinsic plasmid or phage (Barrangou and Horvath, 2017). The host is able to integrate the invading genes on its CRISPR loci as “spacer”, although the mechanism of incorporation remains unclear (Marraffini, 2016). Afterwards, this sequence is transcribed to RNA, namely small-interfering CRISPR RNA (crRNA). The crRNA hybridizes with the transactivating CRISPR RNA (tracrRNA) to form a guide RNA (gRNA), guiding the Cas nuclease to specifically cleave invading cognate sequences with the help of protospacer-adjacent motif (PAM) located downstream of the target sequence (Makarova et al., 2015; Steinert et al., 2015; Xu et al., 2015). To date, it has developed into a versatile and powerful tool widely applied to manipulate almost any genomes in biology, including yeast, human cells, plant, mice, drosophila, zebrafish, and etc. (DiCarlo et al., 2013; Gratz et al., 2013; Hwang et al., 2013; Li et al., 2013b; Muller et al., 2016; Sander and Joung, 2014). Realization of gene editing including specific gene knock-out as well as integration of desired genes requires introducing or expressing the Cas9 and engineered gRNA in the target cells or organisms. (Sander and Joung, 2014). The

most frequently studied and used type of this system is the *S. pyogenes* CRISPR-Cas9 (SpCas9). It can generate sequence-specific double stand DNA break (DSB) at multiple sites more precisely, compared to other editing tools such as zinc-finger or TALE nucleases (Kleinstiver et al., 2015; Marraffini, 2016). CRISPR/Cas9 systems from different bacteria possess different Cas proteins of distinct sizes and various PAM requirements (Steinert et al., 2015). The other widely used Cas9 system is from *S. thermophilus*, which recognizes longer PAM sequences and shows relatively lower off-target mutagenesis compared to the SpCas9 system (Muller et al., 2016).

Although the CRISPR/Cas system has been widely and successfully used, the system still has drawbacks that need to be optimized, particularly, the high frequency off-target effect that may lead to DNA cleavage at wrong sites. A recently published research used T7 endonuclease I (T7EI) assay to unveil multiple off-target sites of the *VEGFA*, *EMX1* genes and its high frequency at genomic level in human U2OS.EGFP cells. The off-target cleavage can lead to various mutations, including insertions, deletion and point mutations, and the rate of mutations is reported to reach up to 125% (Cradick et al., 2013; Fu et al., 2013). Therefore, several approaches of reducing the off-target effect have been designed, including using cas9 nickase (Shen et al., 2014), truncating the sgRNA (Fu et al., 2014), constructing fusions of catalytically inactive Cas9 and FokI nuclease (fCas9) (Wishart, 2016), and etc. However, these methods can only solve part of the problem. Recently, researches constructed a SpCas9-HF1 variant, of which four Spcas9 residues (N497A, R661A, Q695A, and Q926A) were substituted by alanine. The SpCas9-HF1 was able to significantly minimized the genome-wide off-target effects to undetectable levels with the on-target activities remaining high compared to wide type, and this might provide a general approach to optimize other CRISPR-RNA-guide nucleases (Kleinstiver et al., 2016).

Metagenomics

Oral microenvironment is a pool of more than 700 species of microbes. It not only affects the oral health and diseases, but also affects systematic diseases, such

as infective endocarditis, bacterial pneumonia, and etc. (Brown, 2007; Paik et al., 2005; Xu et al., 2007; Xu and Gunsolley, 2014). To gain insight in the microbial related diseases requires comprehensive study of entirety of microbes in a niche. Metagenomics has emerged in recent years and has been applied in analyzing the total genomes in a selected microenvironment (Figure 2). It is a preferred method of investigating the microbial community and its diversity, without isolating and cultivating the enormous organisms (Martínez-Porchas and Vargas-Albores, 2017). Belda-Ferre et al. (2012) analyzed the oral metagenome under different oral health conditions by 454 pyrosequencing. Interestingly, they found that *Streptococcus sanguinis* and *Aggregatibacter* were abundant in the caries-free individuals, while a plenty of *Streptococcus gordonii* and *Leptotrichia buccalis* were detected in the caries affected individuals. Besides, *S. mutans* was not prominent in the carious cavities, although it was thought to be the principle pathogenic factor of caries. Instead, there was an intricate bacterial community, indicating that the caries is a polymicrobial disease, and the *S. mutans* may trigger the carious and cooperate with other bacteria to promote the development of caries.

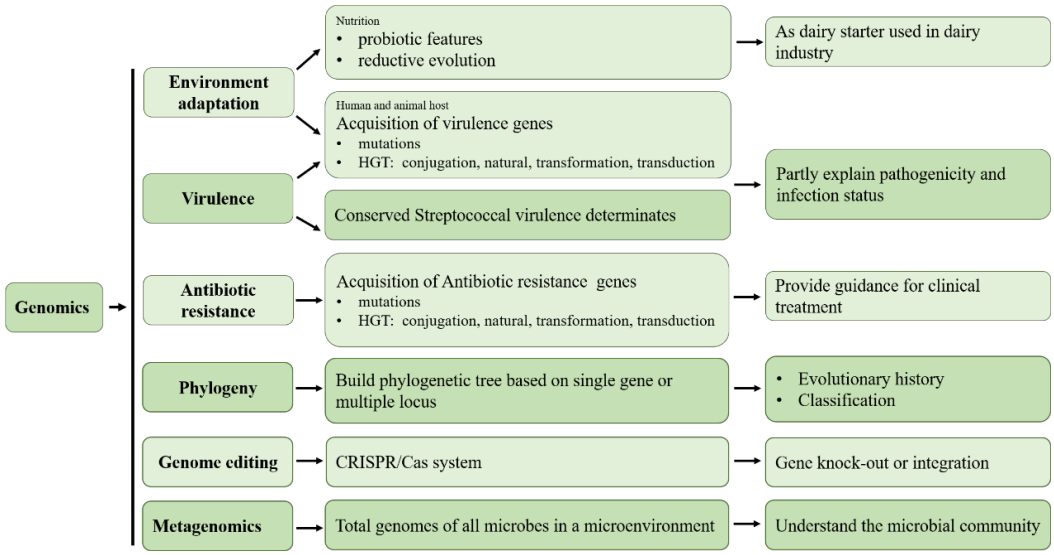

Figure 2. Schematic view of application of genomics in *Streptococcus*. By comparing with the genome of genetically similar group members in database, *S. thermophilus* shows a lack of most Streptococcal virulence-associated genes. Besides, it possesses encoding genes for probiotic features, enabling it to serve as diary starter. Genomics can also unveil the conserved virulence

genes and the putative virulence-associated genes acquired over the course of evolution, which may be attribute to the environment adaptation and partly explain the pathogenicity and infection status of host. Some genes acquired is related to antibiotic resistance, which might provide guidance for clinical treatment. Through phylogenetic analysis, a phylogenetic tree based on single gene or multiple genes can be made to illustrate the evolutionary history of the *Streptococcus* and contribute to the classification of some species. Moreover, the CRISPR/Cas system of *Streptococcus* is a powerful genome editing tool to realize gene knock-out or integration. As for a niche, the metagenomics can be used to analyze the total genomes of all microorganisms in the microenvironment and gain insight in the microbial community and its diversity.

Transcriptomics

Transcriptomics is a discipline that analyzes the transcription of genes in cells and the laws of transcription regulation on the whole. Briefly, transcriptomics studies the gene expression at the RNA level. Transcriptomics plays an important role in bridging genomics and proteomics. The transcriptome is defined as the sum total of all mRNA molecules of an organism (McGettigan, 2013). Since a human EST (expression sequence tag) database that retained 609 cDNA clones was firstly reported by a group from National Institutes of Health (NIH) in 1991 (Adams et al., 1991), transcriptomics, which has become an important method in the field of cell phenotype and function research, enables us to unveil the gene expression under a certain condition and provides us information about the strength of the gene expression relative to a reference (Afzal et al., 2015a).

For many years, researchers thought that the structure of bacterial transcriptome was less complicated compared to the structure of a eukaryote (Sharma and Vogel, 2014; Stazic and Voss, 2016). However, with the development of sequencing technologies, the interests in bacterial transcriptomics has increased recently due to the discovery of dense transcriptions on the bacterial chromosome, abundant small regulatory RNAs (sRNAs) and antisense transcripts (Sorek and Cossart, 2010; Thomason and Storz, 2010). Based on the next-generation sequencing technologies, the process of bacterial transcriptomics includes total RNA extraction, target RNA enrichment, library preparation and sequencing, reads mapping, analysis, and etc. It is acknowledged that microarray, containing

abundant probes that can perform multiple genetic tests at the same time, is a reliable, rapid and comparatively economical method in bacterial transcriptomics (Afzal et al., 2015a; Conway and Schoolnik, 2003). However, compared with microarrays, the advantages of high-throughput RNA sequencing (RNA-seq) includes single base pair resolution, low background signal, a large dynamic range of expression levels, higher levels of reproducibility, smaller sample requirements for starting RNA and no limitation in detecting. Those advantages make it a powerful approach for assaying bacterial transcriptomes (McClure et al., 2013; Wang et al., 2009). Although it is still a developing technology, transcriptomics provides a novel approach of studying bacterial diversification at the genetic level.

Advances in Streptococcus transcriptomics

Virulence and Pathogenicity

In transcriptomic studies, a large number of genes have been shown to be involved in the expression of *Streptococcus* virulence and pathogenicity. For instance, in the absence of the two-component control of virulence regulator-sensor (CovRS) of Group A Streptococcus, an increased expression of genes associated with antiphagocytosis, toxin gene abrogation and destruction of some gene products was detected, unlike previous studies. Moreover, a genome-scale metabolic model was established and 16 nonredundant metabolic gene modules were found to be essential for pathogen growth (Yun-Juan et al., 2015). Meanwhile, Kreth et al. (2011) found that the global regulator CodY could control the regulation behavior of CovRS according to nutritional status, which indicated that there was a counteractive balance of transcriptome expression. Kajfasz et al. (2011) studied the mechanism of how the ClpXP proteolytic complex controlled the key virulence properties of *S. mutans,* and they demonstrated that genes encoding enzymes were required for the biosynthesis of intracellular polysaccharides and malolactic fermentation was upregulated. In addition, environment plays a key role in pathogenicity of *Streptococcus* as well. Pettigrew et al. (2014) pointed out that genes, associated with carbohydrate metabolism and production of bacteriocins, were upregulated. Besides, genes with

colonization were downregulated. Those gene expression changes resulted in higher virulence when *S. pneumoniae* exposed to influenza A virus (IAV) infection. In the entire pathogenic process, Virtaneva et al. (2005) reported that there were some temporal changes in the Group A Streptococcus transcriptome. In particular, prophage virulence gene expression and prophage induction occurred predominantly when high pathogen cell densities and acute inflammation presented. As for the interaction with the host microenvironment, Geng et al. (2012) found that genes related to overall pathogenic capacity of *Streptococcus parasanguinis* expressed differently to enable bacteria to avoid host immune clearance, to colonize in host tissues and to survive within oral biofilms.

The main content of *Streptococcus* transcriptomics is centered on its virulence and pathogenicity, which are comprehensively affected by physiological characteristics such as competence and physiological development. Genes associated with those characteristics may influence the final characterization of bacteria, and usually, virulence and pathogenicity are mostly affected. The advances of these researches are as follows.

Competence

The competence is important for *Streptococcus* to survive and behave in the society. Referring to the regulatory links in the competence of *S. mutans*, Khan et al. (2016) found that the XIP-encoding gene *comS* determined the entire late response to competence-stimulating peptides (CSP) and the response regulator ComE determined the immediate transcriptional response to CSP. They also compared their data with published transcriptome data, permitting the identification of all of the operons in each regulon and identified a core set of 27 pan streptococcal competence genes within the SigX regulon. However, combined with the bacterial flow-cytometric sorting of cells, Lemme et al. (2011) concluded that the activation of bacteriocin-related genes was the main approach of cells' response to CSP, and then, some of these cells would be separated into two subpopulations of which only one would become competent.

Physiological development and biofilms

The different expression of transcriptome of *Streptococcus* in its different growth phases might be associated with its host infection capacity. Geng et al. (2012) indicated that genomic islands (GIs) and various open reading frames outside the GIs expressed differently in early exponential and early stationary growth phases, contributing to its pathogenic capacity. As for biofilm cells, transcriptome of *Streptococcus* affects its phenotype and response to microenvironment. Klein et al. (2010) reported that the expression of genes involved in sugar metabolism was mostly significantly affected in the *S. mutans* biofilm formation. And different expression of other related genes, such as genes associated with two-component systems, fermentation/glycolysis and iron transport, may lead to its pathogenicity. Liu et al. (2013a) created a hyperosmotic condition for *S. mutans* as an environmental stress and found that biofilm was dispersed due to *gtfB* and *comC* genes down-regulation and genes involved in carbohydrate metabolism up-regulation. In addition, the upregulation of some specific genes related to other environmental stress may suggest that there was a potential cross-talk between these stimuli. Wen et al. (2011) analyzed the growth of LuxS-deficient *S. mutans* in biofilm and found that some genes, such as those related to sugar-specific enzymes II of the phosphotransferase (PTS) system and the ATP operon, were upregulated and the ability of confronting environmental stress was compromised. These results may reveal the relationship between the select gene, LuxS, and biofilm formation and virulence of *S. mutans*. Polymicrobial biofilm is a more reality biofilm mode compared to monomicrobial biofilm in *Streptococcus* research. Transcriptomics plays an important role in interspecies interactions for their physiology and virulence. Sztajer et al. (2011) co-cultivated the *S. mutans* and *Candida albicans* and found the communication between the two pathogens. They used gas chromatography–mass spectrometry and transcriptome analysis and revealed a relatively lower level of EPS expression of *S. mutans* in dual-species biofilm. Activation of *sigX*, the alternative sigma factor of *S. mutans* induced by quorum sensing signals, was detected only when cultured in mixed biofilms or treated with conditioned media from co-cultivated-biofilms, not in the single-species biofilms or with culture supernatants from single-species biofilms.

These results demonstrate that *C. albicans* is able to stimulate the quorum sensing system of *S. mutans*, leading to a change of pathogenicity of *S. mutans*.

Novel gene function of *Streptococcus*

Transcriptomics makes it possible to discover new gene functions. For example, Yuzenkova et al. (2014) suggested that gene expression in *S. pneumoniae* may regulate transcription elongation. Afzal et al. (2015b) found that the transcriptional regulator NanR acted as a transcriptional activator of nan operon I and the *nanA* gene regulated the presence of sialic acid in *S. pneumoniae*.

Identification

In addition, transcriptomics, as an important part of bacterial omics, can also help to identify the genetic distance and reveal the extent of genetic relatedness between closely related strains. Park et al. (2012) proved this point of view by comparing transcriptomes of *Streptococcus pseudopneumoniae* with *S. pneumoniae*, *S. mitis* and *S. oralis*. With genes up- or down-regulation, this novel species was demonstrated to be closest related to *S. pneumoniae*.

Other aspects

Transcriptomics is also used in exploring the effect of some specific factors on *Streptococcus* transcriptome, such as metal ions (Manzoor et al., 2015), biologically active substances (Feng et al., 2013; Haase et al., 2015; Zeng and Burne, 2015), drug ingredients (Feng et al., 2011; Ferrándiz and de, 2014), and special environmental conditions (Kalpana et al., 2016; Richards et al., 2013; Sitkiewicz et al., 2009). In addition, transcriptomics was used in studying transcriptome changes of host immune organs after the infection of *Streptococcus* in animal models (Gaur et al., 2014; Sitkiewicz et al., 2009; Wu et al., 2010).

Macro transcriptomics of Streptococcus

In a microenvironment, transcriptomics of microbiome becomes a novel research direction. Taking oral microenvironment as an example, Peterson et al. (2014) used RNA-seq to analyze dental plaque microbiota and they found that 45%

transcription was derived from *Streptococcus*. And Benitez-Paez et al. (2014) found that expression of genes, which were associated with bacterial adhesion to tooth surface, changed significantly at the the early biofilm stage compared to the mature biofilm stage.

Transcriptomics enables us to understand bacterial physiological activities, competence, interaction with the host and other characteristics of *Streptococcus* at gene level. And it also provides novel clues of the relationship of bacterial genus and the microenvironment that bacteria live. RNA-seq becomes the most frequently used method in transcriptomics, although there are still problems in analyzing excessive data of bacterial genes. The future directions of new sequencing technologies are on the premise of ensuring safety to determine the transcriptome of a certain bacteria quickly and simply.

Proteomics

With the development of omics, proteomics technology is widely used and it plays an important role in microbiology research. Based on proteome, proteomics is applied in the studies of the composition and protein expressions of cells, tissues and organisms. Compared to genomics, proteomics provides more information about proteins, such as protein localization, interactions among proteins and the occurrence of post-translational modifications. Genes and transcripts are both important and useful in characterizing the activity of microorganisms, because they control the protein expression. However, protein studies provide more direct information of the functional activity of the cell, organism or community (Franzosa et al., 2015). Microbes are special organisms that they live in human, animal and the environment around us. Global analysis of microbial proteome has been widely used in investigating the pathogenesis of important human bacterial pathogens (Bittaye and Cash, 2015). Analyzing the proteome expressed by microbes is essential to demonstrate the relationship of microbes and human diseases, and the pathogenic mechanism of microbial virulence. Here, the proteomics researches of *Streptococcus* are summarized as below.

Streptococcus genus is one of the most invasive bacterial groups and has been identified as sources of invasive infections in human. *S. pneumoniae*, *S. pyogenes*, *S. agalactiae*, *S. mutans* and *S. suis* are considered to cause infections. Therefore, revealing the virulence of *Streptococcus* genus and the mutation of *Streptococcus* is a worldwide object for human's health. This requires more information about the pathogen, especially the proteome data (Bittaye and Cash, 2015; Olaya-Abril et al., 2014). As we know, penicillin and other antibiotics are widely used in treating infections of *Streptococcus*, meanwhile, the antibiotic resistance emerges worldwide and becomes a serious problem that threatens public health. The efficiency of vaccines and antibiotics are compromised because of the super virulent strains and multidrug resistant strains, such as *S. pneumoniae* mutant. Protein vaccines have many advantages compared to the current polysaccharide-based vaccines, which are serotype-dependent and subjected to losing efficacy because of serotype replacement and high manufacturing complexity (Olaya-Abril et al., 2013). Proteomics plays a vital role in the researches about microbes, such as *Streptococcus.*

Virulence of Streptococcus

Proteins are closely related to the virulence of microbes. Through analyzing different expression of proteins especially through proteomics, researchers could provide more information to treat or avoid the infection of microbes. For example, protein vaccines are higher immunogenicity and more cost-effective than polysaccharide-based vaccines (Donnarumma et al., 2016). Therefore, current researches focus on developing vaccine candidates of several surface-associated proteins that elicit protection against more *Streptococcus* strains, especially *Pneumoniae*. Searching for potential microbial therapeutic targets and providing microbial protein-makers are also the objectives. Through proteomics researches, such as functional and comparative proteomics, more and more proteins are revealed (Bogaert et al., 2004).

Functional proteomics related to *Streptococcus*

Streptococcus pneumoniae

As one of the most common kinds of pneumonia, bacterial pneumonia usually happens in children and the elderly, and it is responsible for high morbidity and mortality rates in both developed and developing countries. *S. pneumoniae* is a leading pathogen that causes invasive diseases such as pneumonia, even sepsis, meningitis, and etc. Current controlling strategies include the use of antibiotics for treating ongoing infections together with pneumococcal vaccines for the prevention of further pneumococcal infection (Bittaye and Cash, 2015; Bogaert et al., 2004; Olaya-Abril et al., 2014).

Two-dimensional gel electrophoresis (2DGE) has been frequently used in many researches to investigate the proteome of *S. pneumoniae*. Recent development in global proteome analysis by label-free liquid chromatography-tandem mass spectrometry (LC-MS/MS) -based technologies also helps to provide more information about many microbial proteomes. Studies on the global proteome of *S. pneumoniae* were very limited before, but with the effort of numerous scientists, extensive information about the bacterial surface proteome has been collected, providing a major source for identifying vaccine targets and relating to serological assays. On the bases of proteomic experimental identification followed by LC-MS/MS analysis, Jimenez-Munguia et al. (2015) selected 64 recombined pneumococcal proteins and finally found that acute serum IgG levels against RrgB, a kind of cell wall surface anchor family protein, were significantly lower in less than 4 years old children with pneumococcal pneumonia than those in controls. This result helps the scientists to design a multiplex bead-based platform to assess natural IgG antibodies against pneumococcal protein antigens in children. In addition, Olaya-Abril et al. (2013) identified potential new protein vaccines through pan-surfomic analysis of pneumococcal clinical isolates from adults. In this study, four novel surface proteins (Spr0012, Spr0328, Spr0561 and SP670_2141) are selected for vaccine test, because of their capacity to raise immune response in infected patients and induce an IgM response. IgM should be a priori and good indicator of antigen exposure.

Recently, extracellular vesicles have gained more attention. They are produced

by pathogens with various microbial virulence factors. Therefore, synthesizing and releasing of extracellular vesicles derived from the outer surface are in favor of bacterial pathogens to combat the defense responses of invaded hosts. Based on proteomics and fatty acid analysis of *S. pneumoniae*, it has been demonstrated that these vesicles and the plasma membrane are similar in essential aspects, but there are some differences. The vesicles are more enriched in lipoproteins and short-chain fatty acids, and they act as vectors of surface proteins and virulence factors (Olaya-Abril et al., 2014). Studies on the composition of extracellular vesicles are essential to reveal the virulence of microbes and to provide information to explore vaccine candidates.

In order to cause infection, pathogens such as *S. pneumoniae* need to capture iron from the host environment. In Streptococcal species, there are three known iron-uptake systems responsible for iron acquisition from the host, the ABC transporters PiaABC, PiuABC and PitABC. Because of further analysis by the integrative transcriptomics and proteomics, rich information and insightful clues have been provided for the investigations of iron-transporting mechanism in bacteria, the interplay between Streptococcal iron availability and the biological metabolic pathways (Yang et al., 2016a). By using immuno-proteomic, N-acetylglucosamine-6-phosphate deacetylase (NagA) was identified in a protein mixture secreted by *S. pneumoniae* and its strong immunogenicity was confirmed. On account of high immunogenicity and immunospecificity, NagA may be a novel diagnostic marker for *S. pneumoniae* (Choi et al., 2013).

Streptococcus suis

S. suis is an emerging zoonotic pathogen that causes severe human infections, such as meningitis, septicemia, endocarditis and pneumonia. Among these strains, *S. suis* serotype 2 (SS2) infection is a leading cause of septicemia and meningitis in pigs and humans. And it is often accompanied by bacteremia, the infection of which has become increasingly potent, especially in the southeast Asian countries like Thailand, Vietnam and China. Fibrinogen-binding surface proteins of *S. suis* (SsFBPs) can bind to human fibrinogen (hFg) to help the

microbe to survive in the host, to evade host immunity and to survive and disseminate in blood. By combining the results of proteomic and immune blot, novel FBPs of SS2 were identified, especially, two SsFBPs (MRP and Enolase) were identified as important FBPs of SS2 (Pian et al., 2015). Haas and Grenier (Haas and Grenier, 2015) applied analyzed proteomics of membrane vesicles (MVs), which range in diameter from 13 to 130 nm and appear to be coated by capsular material. They revealed that in the MVs there were 46 proteins, 9 of which are considered as proven or suspected virulence factors. The results implied that *S. suis* MVs may play a role as a virulence factor in the pathogenesis of *S. suis* infections. Zhou et al. (2015) demonstrated that substrate binding protein (Sbp) might play an important role in the pathogenesis of SS2. Identified by immunoproteomic method, four SS2 membrane associated proteins, including L-lactate dehydrogenase (Ldh), Dihydrolipoamide dehydrogenase (Dldh), Pyruvate dehydrogenase E1 component (Pec) and amino acid ABC Sbp, were cloned and expressed as recombinant proteins with Histag. These findings may provide foundations for future researches of SS2 pathogenic mechanisms and targets, provide foundations for excellent candidates for vaccine development, and are important for novel therapies for targeting SS2-induced sepsis.

Streptococcus pyogenes

S. pyogenes, a kind of human pathogens, can cause various infectious diseases, ranging from mild pharyngitis and impetigo to severe necrotizing fasciitis and autoimmune sequelae. Acidic stimulus and growth phase influence the secretion of *S. pyogenes*-associated virulence factors, which are ultimately associated with *S. pyogenes* invasive capability. Proteomics, electron microscopy and Fluorescence Activated Cell Sorting (FACS) analysis demonstrate that *S. pyogenes* can release lipoproteins (Lpps) in lipoprotein-rich membrane vesicles (LMVs), a vesicle-like structure, and reveal that Lpps are the major components of the LMVs. LMVs represent as an excellent vector for antigen delivery via their potential intrinsic adjuvanticity (Biagini et al., 2015). The “surface interactome” of *S. pyogenes* was characterized by protein array technology (Galeotti et al., 2012). The proteins were proved to get involved in protein folding and transportation.

The proteins are OppA, DppA, PrsA and TlpA, all of which are localized at the septum. Besides, scientists have detected interactions among the metal transporters AdcA, Lmb and SpeI, a major player in *S. pyogenes* pathogenesis. Those findings might indicate the way SpeI obtains the metal from the environment. Through applying selected reaction monitoring mass spectrometry (SRM-MS), a method to identify and quantify proteins with high sensitivity, specificity and high reproducibility, Karlsson et al. (2012) provided a proteome-wide selected reaction monitoring (SRM) assay repository resource for *S. pyogenes* to facilitate SRM-MS analysis for this bacterium.

Other *Streptococcus*

S. mutans is one of the major pathogens that causes dental caries. *S. mutans* and sucrose are responsible for the development of virulent-cariogenic biofilms. Klein et al. (2012) used high-throughput quantitative proteomics to examine relevant proteins that were produced by *S. mutans* to facilitate its growth and optimal survival during mixed-species biofilm development induced by sucrose. They found that the protein synthesis and gene expression were usually augmented when *S. mutans* grew in mixed-species biofilms compared to in single-species biofilms. The specific genes, which were upregulated, were associated with glucan synthesis and remodeling (gtfBC, dexA) and glucan binding (gbpB). This result indicated that the presence of other organisms, which were fundamentally different in the matrix assembly, survival and biofilm maintenance, increased the virulence of *S. mutans*.

S. thermophiles, as a starter for dairy productions, is widely used in food industry. In order to contribute to future investigations on protein complex assembly and composition or search for bacterial strains with specific biotechnological applications, Salzano et al. (2013) applied proteomic procedures to characterize the heteromultimeric and homomeric protein complexes from the membrane fraction of *S. thermophilus*. Moreover, researchers analyzed proteome of extracted cytoplasmic proteins from *S. sanguinis* and revealed O-Glycosylation of the N-terminal region of the serine-rich adhesin Srr1 of *S. agalactiae* by Mass Spectrometry (Chaze et al., 2014; El-Rami et al., 2017).

Comparative proteomics

Isobaric tagging for relative and absolute quantification (iTRAQ), a high-throughput gel-free proteomic approach, allows high resolution quantitative comparisons of protein profiles between multiple phenotypes. By iTRAQ, Allan et al. (2014) studied the growth and metabolic proteins of *S. pneumoniae*, a clinical serotype 14 strain, in planktonic and biofilm status to represent target candidates for future vaccine development. They found several ABC transporter system proteins were significantly up-regulated in pneumococcal biofilms. In one study, iTRAQ-based quantitative proteomics was also applied to analyze the protein alterations in *S. pneumoniae* in response to Sodium new houttuyfonate (SNH) treatment. It was proved that SNH killed *S. pneumoniae* in a dose-dependent manner accompanied by the increase of H_2O_2 (Yang et al., 2016b). Yang et al. (2015) performed 2-DGE-based proteomic analysis to characterize the protein alterations in *S. pneumoniae* after being treated with polypyridylruthenium (II) (Ru (II)) complex X-03, a metal compound. They found that it could obstruct bacterial fatty acid synthesis and oxidation-reduction process in order to suppress the growth of *S. pneumoniae* and interfere iron-acquisition pathway in the bacterium. Vancomycin is one of the most important drugs currently used in the treatment of gram-positive bacterial infections. Yang et al. (2017) applied 2-DGE and LM-MS/MS to analyze differentially expressed proteins (DEPs) of whole cellular proteins extracted from *S. pneumoniae* D39 with or without vancomycin treatment. They found that 27 proteins were upregulated and four proteins were downregulated in vancomycin-treated *S. pneumoniae*. These findings provide important information for future investigations of vancomycin tolerance mechanisms of *S. pneumoniae*. Through the iTRAQ strategy, Wang et al. (2016a) compared the protein expression profiles of *S. suis* growing in sub-MIC tylosin versus no tylosin, and finally 1501 proteins were identified. These proteins were associated with *S. suis* metabolic regulation and virulence, and appear to be important in biofilm growth at sub-MIC tylosin level. Through analyzing acid-influenced secretome of *S. pyogenes* in acidic and neutral conditions, it was demonstrated that acidic stimuli and growth-phase cues were crucial for classical

protein secretion in *S. pyogenes*. Histidine triad protein A (HtpA), as a novel acid-induced virulence factor, was proved to improve the anti-phagocytosis ability for causing necrotizing fasciitis (Wen et al., 2014). By using immunoproteomics to analyze and compare the transcriptomes and proteomes of *S. pyogenes* biofilms, Freiberg et al. (2016) found that the *S. pyogenes in vivo* expressed proteome matches the proteome *in vitro* biofilm better, in compared with *in vitro* planktonic proteome.

With the help of different proteomics analysis, it was demonstrated that virulence factors of *S. mutant* expressed differently. By a 2DGE-based proteome reference map of the cytoplasmic and extracellular proteins from *S. mutans*, Li et al. (2013a) identified 239 protein spots that represented 192 different cytoplasmic proteins, which were regulated by the carolacton isolated from the myxobacterium *Sorangium cellulosum*. Finally, carolacton was found to inhibit *S. mutans* by disturbing the peptidoglycan biosynthesis and degradation, thus causing damages to the integrity of the cell envelope and disturbing *S. mutans* biofilm viability. Thereby this function makes carolacton a potential anti-biofilm drug.

Streptococcus coaggregates with *Porphyromonas gingivalis*, a major periodontal pathogen, and assists it with the establishment of subgingival biofilm. *S. gordonii*, *S. oralis*, *S. mutans* and *S. sanguinis* have been proved to interact with *P. gingivalis*. By using shotgun proteomics, Maeda et al. (2015) examined the molecular basis of mixed-culture biofilm formation by *P. gingivalis* with *S. oralis*. Compared to the profile in the respective monoculture biofilm, 31 proteins of *P. gingivalis* and two proteins of *S. oralis* were down-regulated in the mixed-culture biofilm. The Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) of *S. oralis* (GapC) as well as the FimA of *P. gingivalis* were up-regulated in mixed-culture, which indicated that the GapC might interacted with FimA to promote the formation of mixed-culture biofilm. The housekeeping proteins GyrB and RpoD increased in the mixed-culture biofilm and reflected the activity or viability of bacterial cells.

Interaction between Streptococcus and other microbes/human

A recent study based on LC/Q-TOF mass spectrometry proteomics analyzed the changing protein profiles of *Helicobacter pylori* and *S. mitis* in mixed-culture biofilm, and the results indicated that the interaction between *H. pylori* and *S. mitis* may enhance bacterial survival rate and reduce carcinogenesis. The *S. mitis* in mixed-culture biofilm has a high-level expression of posphoglycerate kinase (PGK), a major enzyme involved in glycolysis, which is beneficial to the survival of the bacterium. Meanwhile, some oxidative stress related proteins of *H. pylori* decreased in the multi-species setting, including thioredoxin (TrxA) and other redox-regulating enzymes, and this was thought to be protective and less pathogenic (Khosravi et al., 2016). *S. gordonii* is a pioneer colonizer of the tooth surface. By 2-DE, Yoshida et al. (2015) analyzed the protein expression profiles of *S. mutans* in the biofilm, with or without the presence of *S. gordonii*. It was suggested that Dps-like peroxide resistant protein (Dpr) might be one of the essential proteins for *S. mutans* survival on teeth in the presence of early colonizing oral *streptococcus*.

Sjöholm et al. (2014) used a combination of MS based techniques to comprehensively quantify the components of the *S. pyogenes*-plasma protein interaction network. They subdivided the interacting protein list by using a SRM depending on the level of enrichment and protein concentration on the bacterial surface. And their results provided important information about preventing or treating the infection of *Streptococcus*.

Metabolomics

Since the metabolite is the ultimate output of the biological hierarchy and is more stable than mRNA or proteins (Patel and Ahmed, 2015), the process of metabolism and the metabolite is always directly related to the bacterial growth, pathogenicity, evading host immune system, interaction with other organisms, research of antibiotics, vaccine development, and etc., therefore, metabolite seems to be important for disease prevention and control (Fei et al., 2016; Schulz

and Hammerschmidt, 2013; Takahashi et al., 2012; Wang et al., 2016b). For example, the acid produced by *S. mutans* and other bacteria in oral cavity plays a vital role in the development of caries (Takahashi and Nyvad, 2011). Metabolomics, serving as a complement of other omics, provide us with an integral view of the metabolism of the living cells and can be used to explore some key products and potential mechanisms by comparing metabolic differences under different circumstance (Schulz and Hammerschmidt, 2013; Tang, 2011).

Streptococcus iniae is an important zoonotic pathogen that causes infections in various fish including tilapia (Lau et al., 2006). Recently, a series of gas chromatography-mass spectrometer (GC/MS) based metabolomic researches revealed a regulatory mechanism of tilapia against *S. iniae* infection. The L-leucine in the survival fish infected by *S. iniae* was relatively higher than that in the dead fish, indicating the potential modulator role of L-leucine in the *S. iniae* infection. Furthermore, the addition of exogenous L-leucine can raise the survival capacity of infected tilapia, and this validated the result of metabolomic analysis (Ma et al., 2015). The GC/MS-based metabolomic analysis revealed the mechanism of how L-leucine works. Elevated level of L-serine triggers further up-regulation of the *IL-1β* and *IL-8* mRNA expression (Du et al., 2017). The metabolic profile of bacteria can also reflect the bacterial adaption changes by comparing the final outputs. *S. agalactiae* is a common pathogen that causes systemic infections in human and fish, and it is serum-resistant and capable to escape from the host immune system. By using GC/MS-based metabolomics analysis, a recent research illuminated a metabolic trick of *S. agalactiae* and explained the key pathways in response to plasma of yellow grouper, including an increase of adenosine and a decrease of malic acid that decreased the sensitivity to plasma (Wang et al., 2016b).

Drug discovery

Genomics, proteomics and metabolomics are the three fields conventionally and effectively applied in drug development and clinical application, especially when confronted with the increased antibiotic resistance. Genomics has been applied

in multiple areas, including identification of biomarker (osteoarthritis), expression profiling of tissue, toxicogenomics, and etc. (Meulenbelt et al., 2011; Russell et al., 2013). The proteomics owns its unique advantages over genomics, because the proteomics reflected the cellular events after gene transcriptions. The employment of it can be found nearly throughout the pipeline of drug development, from target identification to clinical trial of drugs (Burbaum and Tobal, 2002; Russell et al., 2013). Metabolomics make it clear about the metabolic pathways and all of the products of both pathogens and the host, revealing how the pathogenic factors act and host defense responses, helping us to understand the mechanisms of diseases at the molecular level and enabling us to fight with them. There are multiple approaches implemented in the development of antibiotics against Streptococcal infections. The comparative genomics can be used to identify the essential genes for S. pneumoniae growth as potential target (Song et al., 2005). A more valid approach is the computational subtractive genomics analysis. The complete proteome of S. pneumonia is downscaling gradually to a small number, which are related to pathogenicity and involved in pathogen metabolic pathways that are essential for the survival of the bacteria. These selected proteins may be potential drug targets. Simultaneously, this approach also encompasses the use of BLAST tool to recognize the non-homologous against human genomes. Eventually, two candidates can guide new drug discovery and reduce the side effect on human (Wadood et al., 2018). A recent published research used an integrated proteomic and metabolomic analysis has evaluated the mechanism of how rhodomyrtone antagonizes S. pneumoniae. A significant down-regulate of two enzymes (glycosyltransferase and UTP-glucose-1-phosphate Uridylyltransferase) and three metabolites (uridine 5′-diphosphoglucuronic acid (UDP-glucuronic acid), UDP-glucose, and UDP-N-acetyl-D-Galactosamine) related to the pneumococcal capsule synthesis was observed in the S. pneumoniae stimulated by rhodomyrtone. This result, confirmed the antimicrobial activity and potential use of rhodomyrtone as antibiotic, thus, providing an alternative antibiotic for treatment against S. pneumoniae (Mitsuwan et al., 2017).

Integrated omics

Omics is a powerful tool used in covering genomics, transcriptomics, proteomics and metabolomics to gain insight of the biological features of *Streptococcus* as well as other creatures. Each of these omics tools can provide a relatively comprehensive understanding of the given object, but sometimes it is not enough to explain more complex phenomena based on a single method. For example, Argimon et al. (2014) used genomics tools and failed to find critical virulence genes of *S. mutans* to explain caries status. But a combination of multiple omics can provide ample evidence of genes, mRNA, proteins and other metabolites at the molecular level. For example, a transcriptomic analysis revealed an up regulation of salvage and de novo nucleotide synthesis pathway of *Streptococcus intermedius* in the aerobic culture compared to anaerobic environment, which was validated by the metabolomic results (Fei et al., 2016). The integration of different omics can imply the “missing link” between gene transcription, translation of protein, the following modification and other process, revealing the complicated regulatory mechanisms and networks behind certain phenotype. For example, the matrix metalloproteinases 2 (MMP2) mRNA expression of oral epithelia stimulated by *C. albicans* is up-regulated, but the MMP2 protein expression is down-regulated, which is remained confused (Claveau et al., 2004). An integrated omics may help to solve these problems.

A recent published research used an integrated proteomic and metabolomic analysis has revealed the mechanism of rhodomyrtone antibacterial activity against *S. pneumoniae*. A significant decrease of two enzymes (glycosyltransferase and UTP-glucose-1-phosphate Uridylyltransferase) and three metabolites (uridine 5′-diphosphoglucuronic acid (UDP-glucuronic acid), UDP-glucose, and UDP-N-acetyl-D-Galactosamine) related to the pneumococcal capsule synthesis was observed in the *S. pneumoniae* stimulated by rhodomyrtone. This result confirmed the antimicrobial activity and potential use of rhodomyrtone as an antibiotic at both transcription and protein levels (Mitsuwan et al., 2017).

The putative genes that regulate iron-uptake of *S. pneumoniae* are ATP-binding cassette (ABC) transporters, PiaABC, PiuABC, and PitABC. Those genes are critical to the growth and pathogenicity of the bacteria. Recently, SPD_0090 and SPD_1609 were identified as potential novel iron-bunding proteins by using translatomics and proteomics, and they were upregulated as compensation in the known ABC transporters mutant *S. pneumoniae* (Yang et al., 2016a). The integration of different omics can provide more detailed and comprehensive evidence to throw light on the potential mechanisms and pathways of certain phenotype. The progress in bioinformation and sequencing technology makes it more practical and economic to carry out integrated omics.

References

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F.*, et al.* (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science 252*, 1651-1656.

Afzal, M., Manzoor, I., and Kuipers, O.P. (2015a). A Fast and Reliable Pipeline for Bacterial Transcriptome Analysis Case study: Serine-dependent Gene Regulation in Streptococcus pneumoniae. *J Vis Exp*. doi:10.3791/52649

Afzal, M., Shafeeq, S., Ahmed, H., and Kuipers, O.P. (2015b). Sialic acid-mediated gene expression in Streptococcus pneumoniae and role of NanR as a transcriptional activator of the nan gene cluster. *Applied and environmental microbiology 81*, 3121-3131. doi:10.1128/AEM.00499-15

Allan, R.N., Skipp, P., Jefferies, J., Clarke, S.C., Faust, S.N., Hall-Stoodley, L., and Webb, J. (2014). Pronounced metabolic changes in adaptation to biofilm growth by Streptococcus pneumoniae. *PLoS ONE 9*, e107015. doi:10.1371/journal.pone.0107015

Argimon, S., Alekseyenko, A.V., DeSalle, R., and Caufield, P.W. (2013). Phylogenetic analysis of glucosyltransferases and implications for the coevolution of mutans streptococci with their mammalian hosts. *PLoS One 8*, e56305. doi:10.1371/journal.pone.0056305

Argimon, S., Konganti, K., Chen, H., Alekseyenko, A.V., Brown, S., and Caufield, P.W. (2014). Comparative genomics of oral isolates of Streptococcus mutans by in silico genome subtraction does not reveal accessory DNA associated with severe early childhood caries. *Infect Genet Evol 21*, 269-278. doi:10.1016/j.meegid.2013.11.003

Barrangou, R., and Horvath, P. (2017). A decade of discovery: CRISPR functions and applications. *Nat Microbiol 2*, 17092. doi:10.1038/nmicrobiol.2017.92

Belda-Ferre, P., Alcaraz, L.D., Cabrera-Rubio, R., Romero, H., Simon-Soro, A., Pignatelli, M., and Mira, A. (2012). The oral metagenome in health and disease. *ISME J 6*, 46-56. doi:10.1038/ismej.2011.85

Benitez-Paez, A., Belda-Ferre, P., Simon-Soro, A., and Mira, A. (2014). Microbiota diversity and gene expression dynamics in human oral biofilms. *BMC Genomics*

15, 311. doi:10.1186/1471-2164-15-311

Biagini, M., Garibaldi, M., Aprea, S., Pezzicoli, A., Doro, F., Becherelli, M., Taddei, A.R., Tani, C., Tavarini, S., Mora, M.*, et al.* (2015). The Human Pathogen Streptococcus pyogenes Releases Lipoproteins as Lipoprotein-rich Membrane Vesicles. *Molecular & Cellular Proteomics : MCP 14*, 2138-2149. doi:10.1074/mcp.M114.045880

Bittaye, M., and Cash, P. (2015). Streptococcus pneumoniae proteomics: determinants of pathogenesis and vaccine development. *Expert review of proteomics 12*, 607-621. doi:10.1586/14789450.2015.1108844

Blair, J.M., Webber, M.A., Baylay, A.J., Ogbolu, D.O., and Piddock, L.J. (2015). Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol 13*, 42-51. doi:10.1038/nrmicro3380

Bogaert, D., De Groot, R., and Hermans, P.W. (2004). Streptococcus pneumoniae colonisation: the key to pneumococcal disease. *The Lancet Infectious diseases 4*, 144-154. doi:10.1016/s1473-3099(04)00938-7

Bolotin, A., Quinquis, B., Renault, P., Sorokin, A., Ehrlich, S.D., Kulakauskas, S., Lapidus, A., Goltsman, E., Mazur, M., Pusch, G.D.*, et al.* (2004). Complete sequence and comparative genome analysis of the dairy bacterium Streptococcus thermophilus. *Nat Biotechnol 22*, 1554-1558. doi:10.1038/nbt1034

Brown, J.S. (2007). Oral biofilms, periodontitis and pulmonary infections. *Oral Diseases 13*, 513-514. doi:10.1111/j.1601-0825.2007.01410b.x

Burbaum, J., and Tobal, G.M. (2002). Proteomics in drug discovery. *Current Opinion in Chemical Biology 6*, 427-433. doi:https://doi.org/10.1016/S1367-5931(02)00337-X

Chaze, T., Guillot, A., Valot, B., Langella, O., Chamot-Rooke, J., Di Guilmi, A.M., Trieu-Cuot, P., Dramsi, S., and Mistou, M.Y. (2014). O-Glycosylation of the N-terminal region of the serine-rich adhesin Srr1 of Streptococcus agalactiae explored by mass spectrometry. *Molecular & Cellular Proteomics : MCP 13*, 2168-2182. doi:10.1074/mcp.M114.038075

Choi, C.W., An, H.Y., Lee, Y.J., Lee, Y.G., Yun, S.H., Park, E.C., Hong, Y., Kim, G.H., Park, J.E., Baek, S.J.*, et al.* (2013). Characterization of Streptococcus pneumoniae N-acetylglucosamine-6-phosphate deacetylase as a novel

diagnostic marker. *Journal of microbiology (Seoul, Korea) 51*, 659-664. doi:10.1007/s12275-013-3451-8

Claveau, I., Mostefaoui, Y., and Rouabhia, M. (2004). Basement membrane protein and matrix metalloproteinase deregulation in engineered human oral mucosa following infection with Candida albicans. *Matrix Biol 23*, 477-486. doi:10.1016/j.matbio.2004.08.006

Conway, T., and Schoolnik, G.K. (2003). Microarray expression profiling: capturing a genome-wide portrait of the transcriptome. *Mol Microbiol 47*, 879-889.

Cradick, T.J., Fine, E.J., Antico, C.J., and Bao, G. (2013). CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res 41*, 9584-9592. doi:10.1093/nar/gkt714

Croucher, N.J., Harris, S.R., Fraser, C., Quail, M.A., Burton, J., van der Linden, M., McGee, L., von Gottberg, A., Song, J.H., Ko, K.S.*, et al.* (2011). Rapid Pneumococcal Evolution in Response to Clinical Interventions. *Science 331*, 430.

Croucher, N.J., Walker, D., Romero, P., Lennard, N., Paterson, G.K., Bason, N.C., Mitchell, A.M., Quail, M.A., Andrew, P.W., Parkhill, J.*, et al.* (2009). Role of Conjugative Elements in the Evolution of the Multidrug-Resistant Pandemic Clone Streptococcus pneumoniae(Spain23F) ST81. *Journal of Bacteriology 191*, 1480-1489. doi:10.1128/JB.01343-08

Delorme, C., Abraham, A.L., Renault, P., and Guedon, E. (2015). Genomics of Streptococcus salivarius, a major human commensal. *Infect Genet Evol 33*, 381-392. doi:10.1016/j.meegid.2014.10.001

DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res 41*, 4336-4343. doi:10.1093/nar/gkt135

Dmitriev, A.V., McDowell, E.J., Kappeler, K.V., Chaussee, M.A., Rieck, L.D., and Chaussee, M.S. (2006). The Rgg regulator of Streptococcus pyogenes influences utilization of nonglucose carbohydrates, prophage induction, and expression of the NAD-glycohydrolase virulence operon. *Journal of bacteriology 188*, 7230-7241. doi:10.1128/jb.00877-06

Donnarumma, D., Faleri, A., Costantino, P., Rappuoli, R., and Norais, N. (2016). The role of structural proteomics in vaccine development: recent advances and

future prospects. *Expert review of proteomics 13*, 55-68. doi:10.1586/14789450.2016.1121113

Du, C.C., Yang, M.J., Li, M.Y., Yang, J., Peng, B., Li, H., and Peng, X.X. (2017). Metabolic Mechanism for l-Leucine-Induced Metabolome To Eliminate Streptococcus iniae. *J Proteome Res 16*, 1880-1889. doi:10.1021/acs.jproteome.6b00944

El-Rami, F., Nelson, K., and Xu, P. (2017). Proteomic Approach for Extracting Cytoplasmic Proteins from Streptococcus sanguinis using Mass Spectrometry. *Journal of molecular biology research 7*, 50-57. doi:10.5539/jmbr.v7n1p50

Fei, F., Mendonca, M.L., McCarry, B.E., Bowdish, D.M.E., and Surette, M.G. (2016). Metabolic and transcriptomic profiling of Streptococcus intermedius during aerobic and anaerobic growth. *Metabolomics 12*. doi:10.1007/s11306-016-0966-0

Feng, G., Klein, M., Gregoire, S., Singh, A., Vorsa, N., and Koo, H. (2013). The specific degree-of-polymerization of A-type proanthocyanidin oligomers impacts Streptococcus mutans glucan-mediated adhesion and transcriptome responses within biofilms. *Biofouling 29*, 629-640.

Feng, J., Billal, D.S., Lupien, A., Racine, G., Winstall, E., Legare, D., Leprohon, P., and Ouellette, M. (2011). Proteomic and transcriptomic analysis of linezolid resistance in Streptococcus pneumoniae. *J Proteome Res 10*, 4439-4452. doi:10.1021/pr200221s

Ferrándiz, M., and de, l., Campa, AG. (2014). The fluoroquinolone levofloxacin triggers the transcriptional activation of iron transport genes that contribute to cell death in Streptococcus pneumoniae. *Antimicrob Agents Chemother 58*, 247-257.

Fiebig, A., Loof, T.G., Babbar, A., Itzek, A., Koehorst, J.J., Schaap, P.J., and Nitsche-Schmitz, D.P. (2015). Comparative genomics of Streptococcus pyogenes M1 isolates differing in virulence and propensity to cause systemic infection in mice. *Int J Med Microbiol 305*, 532-543. doi:10.1016/j.ijmm.2015.06.002

Fleischmann, R., Adams, M., White, O., Clayton, R., Kirkness, E., Kerlavage, A., Bult, C., Tomb, J., Dougherty, B., Merrick, J.*, et al.* (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science 269*, 496-512.

Franzosa, E.A., Hsu, T., Sirota-Madi, A., Shafquat, A., Abu-Ali, G., Morgan, X.C.,

and Huttenhower, C. (2015). Sequencing and beyond: integrating molecular 'omics' for microbial community profiling. *Nature reviews Microbiology 13*, 360-372. doi:10.1038/nrmicro3451

Fraser-Liggett, C.M. (2005). Insights on biology and evolution from microbial genome sequencing. *Genome research 15*, 1603-1610. doi:10.1101/gr.3724205

Freiberg, J.A., Le Breton, Y., Tran, B.Q., Scott, A.J., Harro, J.M., Ernst, R.K., Goo, Y.A., Mongodin, E.F., Goodlett, D.R., McIver, K.S.*, et al.* (2016). Global Analysis and Comparison of the Transcriptomes and Proteomes of Group A Streptococcus Biofilms. *mSystems 1*. doi:10.1128/mSystems.00149-16

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol 31*, 822-826. doi:10.1038/nbt.2623

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology 32*, 279. doi:10.1038/nbt.2808

https://www.nature.com/articles/nbt.2808#supplementary-information

Galeotti, C.L., Bove, E., Pezzicoli, A., Nogarotto, R., Norais, N., Pileri, S., Lelli, B., Falugi, F., Balloni, S., Tedde, V.*, et al.* (2012). Surface Interactome in Streptococcus pyogenes*. *Molecular & Cellular Proteomics : MCP 11*. doi:10.1074/mcp.M111.015206

Gao, X.Y., Zhi, X.Y., Li, H.W., Klenk, H.P., and Li, W.J. (2014). Comparative genomics of the bacterial genus Streptococcus illuminates evolutionary implications of species groups. *PLoS One 9*, e101229. doi:10.1371/journal.pone.0101229

Gaur, U., Xiong, Y.Y., Luo, Q.P., Yuan, F.Y., Wu, H.Y., Qiao, M., Wimmers, K., Li, K., Mei, S.Q., and Liu, G.S. (2014). Breed-specific transcriptome response of spleen from six to eight week old piglet after infection with Streptococcus suis type 2. *Mol Biol Rep 41*, 7865-7873. doi:10.1007/s11033-014-3680-x

Geng, J., Chiu, C.H., Tang, P., Chen, Y., Shieh, H.R., Hu, S., and Chen, Y.Y. (2012). Complete genome and transcriptomes of Streptococcus parasanguinis FW213: phylogenic relations and potential virulence mechanisms. *PLoS One 7*, e34769. doi:10.1371/journal.pone.0034769

Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., and O'Connor-Giles, K.M. (2013). Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics 194*, 1029-1035. doi:10.1534/genetics.113.152710

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

Haas, B., and Grenier, D. (2015). Isolation, Characterization and Biological Properties of Membrane Vesicles Produced by the Swine Pathogen Streptococcus suis. *PLoS ONE 10*. doi:10.1371/journal.pone.0130528

Haase, E., Feng, X., Pan, J., Miecznikowski, J., and Scannapieco, F. (2015). Dynamics of the Streptococcus gordonii Transcriptome in Response to Medium, Salivary α-Amylase, and Starch. *Appl Environ Microbiol 81*, 5363-5374.

Hilty, M., Wuthrich, D., Salter, S.J., Engel, H., Campbell, S., Sa-Leao, R., de Lencastre, H., Hermans, P., Sadowy, E., Turner, P.*, et al.* (2014). Global phylogenomic analysis of nonencapsulated Streptococcus pneumoniae reveals a deep-branching classic lineage that is distinct from multiple sporadic lineages. *Genome Biol Evol 6*, 3281-3294. doi:10.1093/gbe/evu263

Hossain, M., Egan, S.A., Coffey, T., Ward, P.N., Wilson, R., Leigh, J.A., and Emes, R.D. (2015). Virulence related sequences; insights provided by comparative genomics of Streptococcus uberis of differing virulence. *BMC Genomics 16*, 334. doi:10.1186/s12864-015-1512-6

Huddleston, J.R. (2014). Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. *Infect Drug Resist 7*, 167-176. doi:10.2147/IDR.S48820

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R.J., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology 31*, 227. doi:10.1038/nbt.2501

https://www.nature.com/articles/nbt.2501#supplementary-information

Jans, C., Follador, R., Hochstrasser, M., Lacroix, C., Meile, L., and Stevens, M.J.A.

(2013). Comparative genome analysis of Streptococcus infantarius subsp. infantariusCJ18, an African fermented camel milk isolate with adaptations to dairy environment. *BMC Genomics 14*, 200. doi:10.1186/1471-2164-14-200

Jimenez-Munguia, I., van Wamel, W.J., Olaya-Abril, A., Garcia-Cabrera, E., Rodriguez-Ortega, M.J., and Obando, I. (2015). Proteomics-driven design of a multiplex bead-based platform to assess natural IgG antibodies to pneumococcal protein antigens in children. *Journal of proteomics 126*, 228-233. doi:10.1016/j.jprot.2015.06.011

Kajfasz, J., Abranches, J., and Lemos, J. (2011). Transcriptome analysis reveals that ClpXP proteolysis controls key virulence properties of Streptococcus mutans. *Microbiology 157*, 2800-2890.

Kalpana, D., Im, C., and Lee, Y.S. (2016). Comparative growth, cross stress resistance, transcriptomics of Streptococcus pyogenes cultured under low shear modeled microgravity and normal gravity. *Saudi J Biol Sci 23*, 24-33. doi:10.1016/j.sjbs.2015.02.004

Karlsson, C., Malmström, L., Aebersold, R., and Malmström, J. (2012). Proteome-wide selected reaction monitoring assays for the human pathogen Streptococcus pyogenes. *Nature Communications 3*, 1301-. doi:10.1038/ncomms2297

Khan, R., Rukke, H.V., Hovik, H., Amdal, H.A., Chen, T., Morrison, D.A., and Petersen, F.C. (2016). Comprehensive Transcriptome Profiles of Streptococcus mutans UA159 Map Core Streptococcal Competence Genes. *mSystems 1*. doi:10.1128/mSystems.00038-15

Khosravi, Y., Loke, M.F., Goh, K.L., and Vadivelu, J. (2016). Proteomics Analysis Revealed that Crosstalk between Helicobacter pylori and Streptococcus mitis May Enhance Bacterial Survival and Reduces Carcinogenesis. *Frontiers in microbiology 7*, 1462. doi:10.3389/fmicb.2016.01462

Klein, M.I., DeBaz, L., Agidi, S., Lee, H., Xie, G., Lin, A.H., Hamaker, B.R., Lemos, J.A., and Koo, H. (2010). Dynamics of Streptococcus mutans transcriptome in response to starch and sucrose during biofilm development. *PLoS One 5*, e13478. doi:10.1371/journal.pone.0013478

Klein, M.I., Xiao, J., Lu, B., Delahunty, C.M., Yates, J.R., 3rd, and Koo, H. (2012). Streptococcus mutans protein synthesis during mixed-species biofilm

development by high-throughput quantitative proteomics. *PLoS ONE 7*, e45795. doi:10.1371/journal.pone.0045795

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature 529*, 490-495. doi:10.1038/nature16526

Kleinstiver, B.P., Prew, M.S., Topkar, V.V., Tsai, S.Q., and Joung, J.K. (2015). 58. Engineered Cas9 Variants with Novel PAM Specificities Expand the Targeting Range of CRISPR/Cas Nucleases. *Molecular Therapy 23*, S26. doi:10.1016/S1525-0016(16)33663-2

Kreth, J., Chen, Z., Ferretti, J., and Malke, H. (2011). Counteractive balancing of transcriptome expression involving CodY and CovRS in Streptococcus pyogenes. *J Bacteriol 193*, 4153-4165. doi:10.1128/JB.00061-11

Lau, S.K.P., Woo, P.C.Y., Luk, W.-k., Fung, A.M.Y., Hui, W.-t., Fong, A.H.C., Chow, C.-w., Wong, S.S.Y., and Yuen, K.-y. (2006). Clinical isolates of Streptococcus iniae from Asia are more mucoid and β-hemolytic than those from North America. *Diagnostic Microbiology and Infectious Disease 54*, 177-181. doi:https://doi.org/10.1016/j.diagmicrobio.2005.09.012

Lefebure, T., Richards, V.P., Lang, P., Pavinski-Bitar, P., and Stanhope, M.J. (2012). Gene repertoire evolution of Streptococcus pyogenes inferred from phylogenomic analysis with Streptococcus canis and Streptococcus dysgalactiae. *PLoS One 7*, e37607. doi:10.1371/journal.pone.0037607

Lefébure, T., and Stanhope, M.J. (2007). Evolution of the core and pan-genome of Streptococcus: positive selection, recombination, and genome composition. *Genome Biology 8*, R71. doi:10.1186/gb-2007-8-5-r71

Lemme, A., Grobe, L., Reck, M., Tomasch, J., and Wagner-Dobler, I. (2011). Subpopulation-specific transcriptome analysis of competence-stimulating-peptide-induced Streptococcus mutans. *J Bacteriol 193*, 1863-1877. doi:10.1128/JB.01363-10

Levy, S.E., and Myers, R.M. (2016). Advancements in Next-Generation Sequencing. *Annu Rev Genomics Hum Genet 17*, 95-115. doi:10.1146/annurev-genom-083115-022413

Li, J., Wang, W., Wang, Y., and Zeng, A.P. (2013a). Two-dimensional gel-based

proteomic of the caries causative bacterium Streptococcus mutans UA159 and insight into the inhibitory effect of carolacton. *Proteomics 13*, 3470-3477. doi:10.1002/pmic.201300077

Li, W., Teng, F., Li, T., and Zhou, Q. (2013b). Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotechnol 31*, 684-686. doi:10.1038/nbt.2652

Liu, C., Niu, Y., Zhou, X., Zhang, K., Cheng, L., Li, M., Li, Y., Wang, R., Yang, Y., and Xu, X. (2013a). Hyperosmotic response of streptococcus mutans: from microscopic physiology to transcriptomic profile. *BMC Microbiol 13*, 275. doi:10.1186/1471-2180-13-275

Liu, G., Zhang, W., and Lu, C. (2013b). Comparative genomics analysis of Streptococcus agalactiae reveals that isolates from cultured tilapia in China are closely related to the human strain A909. *BMC Genomics 14*, 775. doi:10.1186/1471-2164-14-775

Ma, Y.M., Yang, M.J., Wang, S., Li, H., and Peng, X.X. (2015). Liver functional metabolomics discloses an action of L-leucine against Streptococcus iniae infection in tilapias. *Fish Shellfish Immunol 45*, 414-421. doi:10.1016/j.fsi.2015.04.037

Ma, Z., Geng, J., Yi, L., Xu, B., Jia, R., Li, Y., Meng, Q., Fan, H., and Hu, S. (2013). Insight into the specific virulence related genes and toxin-antitoxin virulent pathogenicity islands in swine streptococcosis pathogen Streptococcus equi ssp. zooepidemicus strain ATCC35246. *BMC Genomics 14*, 377. doi:10.1186/1471-2164-14-377

Maeda, K., Nagata, H., Ojima, M., and Amano, A. (2015). Proteomic and transcriptional analysis of interaction between oral microbiota Porphyromonas gingivalis and Streptococcus oralis. *Journal of proteome research 14*, 82-94. doi:10.1021/pr500848e

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H.*, et al.* (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol 13*, 722-736. doi:10.1038/nrmicro3569

Manzoor, I., Shafeeq, S., and Kuipers, O.P. (2015). Transcriptome analysis of

Streptococcus pneumoniae D39 in the presence of cobalt. *Genom Data 6*, 151-153. doi:10.1016/j.gdata.2015.08.033

Marraffini, L.A. (2016). The CRISPR-Cas system of Streptococcus pyogenes: function and applications (University of Oklahoma Health Sciences Center, Oklahoma City (OK)).

Martínez-Porchas, M., and Vargas-Albores, F. (2017). Microbial metagenomics in aquaculture: a potential tool for a deeper insight into the activity. *Reviews in Aquaculture 9*, 42-56. doi:10.1111/raq.12102

Mattos-Graner, R.O., Napimoga, M.H., Fukushima, K., Duncan, M.J., and Smith, D.J. (2004). Comparative Analysis of Gtf Isozyme Production and Diversity in Isolates of Streptococcus mutans with Different Biofilm Growth Phenotypes. *Journal of Clinical Microbiology 42*, 4586-4592. doi:10.1128/JCM.42.10.4586-4592.2004

McClure, R., Balasubramanian, D., Sun, Y., Bobrovskyy, M., Sumby, P., Genco, C.A., Vanderpool, C.K., and Tjaden, B. (2013). Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res 41*, e140. doi:10.1093/nar/gkt444

McGettigan, P.A. (2013). Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol 17*, 4-11. doi:10.1016/j.cbpa.2012.12.008

Metcalf, B.J., Chochua, S., Gertz, R.E., Jr., Hawkins, P.A., Ricaldi, J., Li, Z., Walker, H., Tran, T., Rivers, J., Mathis, S.*, et al.* (2017). Short-read whole genome sequencing for determination of antimicrobial resistance mechanisms and capsular serotypes of current invasive Streptococcus agalactiae recovered in the USA. *Clin Microbiol Infect 23*, 574 e577-574 e514. doi:10.1016/j.cmi.2017.02.021

Meulenbelt, I., Kraus, V.B., Sandell, L.J., and Loughlin, J. (2011). Summary of the OA biomarkers workshop 2010 – genetics and genomics: new targets in OA. *Osteoarthritis and Cartilage 19*, 1091-1094. doi:https://doi.org/10.1016/j.joca.2011.06.002

Mitsuwan, W., Olaya-Abril, A., Calderon-Santiago, M., Jimenez-Munguia, I., Gonzalez-Reyes, J.A., Priego-Capote, F., Voravuthikunchai, S.P., and Rodriguez-Ortega, M.J. (2017). Integrated proteomic and metabolomic analysis reveals that rhodomyrtone reduces the capsule in Streptococcus pneumoniae. *Sci Rep 7*, 2715. doi:10.1038/s41598-017-02996-3

Muller, M., Lee, C.M., Gasiunas, G., Davis, T.H., Cradick, T.J., Siksnys, V., Bao, G., Cathomen, T., and Mussolino, C. (2016). Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. *Mol Ther 24*, 636-644. doi:10.1038/mt.2015.218

Munro, C., Michalek, S.M., and Macrina, F.L. (1991). Cariogenicity of Streptococcus mutans V403 glucosyltransferase and fructosyltransferase mutants constructed by allelic exchange. *Infection and immunity 59*, 2316-2323.

Muzzi, A., and Donati, C. (2011). Population genetics and evolution of the pan-genome of Streptococcus pneumoniae. *Int J Med Microbiol 301*, 619-622. doi:10.1016/j.ijmm.2011.09.008

Nasser, W. (2014). Evolutionary pathway to increased virulence and epidemic group A Streptococcus disease derived from 3,615 genome sequences. *Proceedings of the National Academy of Sciences of the United States of America v. 111*, pp. E1768-E1776-2014 v.1111 no.1717. doi:10.1073/pnas.1403138111

Oehmcke, S., Shannon, O., von Köckritz-Blickwede, M., Mörgelin, M., Linder, A., Olin, A.I., Björck, L., and Herwald, H. (2009). Treatment of invasive streptococcal infection with a peptide derived from human high-molecular weight kininogen. *Blood 114*, 444-451. doi:10.1182/blood-2008-10-182527

Oehmcke, S., Westman, J., Malmström, J., Mörgelin, M., Olin, A.I., Kreikemeyer, B., and Herwald, H. (2013). A Novel Role for Pro-Coagulant Microvesicles in the Early Host Defense against Streptococcus pyogenes. *PLOS Pathogens 9*, e1003529. doi:10.1371/journal.ppat.1003529

Olaya-Abril, A., Jimenez-Munguia, I., Gomez-Gascon, L., Obando, I., and Rodriguez-Ortega, M.J. (2013). Identification of potential new protein vaccine candidates through pan-surfomic analysis of pneumococcal clinical isolates from adults. *PLoS ONE 8*, e70365. doi:10.1371/journal.pone.0070365

Olaya-Abril, A., Prados-Rosales, R., McConnell, M.J., Martin-Pena, R., Gonzalez-Reyes, J.A., Jimenez-Munguia, I., Gomez-Gascon, L., Fernandez, J., Luque-Garcia, J.L., Garcia-Lidon, C*., et al.* (2014). Characterization of protective extracellular membrane-derived vesicles produced by Streptococcus pneumoniae. *Journal of proteomics 106*, 46-60. doi:10.1016/j.jprot.2014.04.023

Olson, A.B., Kent, H., Sibley, C.D., Grinwis, M.E., Mabon, P., Ouellette, C., Tyson,

S., Graham, M., Tyler, S.D., Van Domselaar, G., *et al.* (2013). Phylogenetic relationship and virulence inference of Streptococcus Anginosus Group: curated annotation and whole-genome comparative analysis support distinct species designation. *BMC Genomics 14*, 895. doi:10.1186/1471-2164-14-895

Paik, S., Senty, L., Das, S., Noe, J.C., Munro, C.L., and Kitten, T. (2005). Identification of Virulence Determinants for Endocarditis in Streptococcus sanguinis by Signature-Tagged Mutagenesis. *Infection and Immunity 73*, 6064-6074. doi:10.1128/IAI.73.9.6064-6074.2005

Papadimitriou, K., Anastasiou, R., Mavrogonatou, E., Blom, J., Papandreou, N.C., Hamodrakas, S.J., Ferreira, S., Renault, P., Supply, P., Pot, B., *et al.* (2014). Comparative genomics of the dairy isolate Streptococcus macedonicus ACA-DC 198 against related members of the Streptococcus bovis/Streptococcus equinus complex. *BMC Genomics 15*, 272. doi:10.1186/1471-2164-15-272

Park, H.K., Myung, S.C., and Kim, W. (2012). Comparative transcriptomic analysis of streptococcus pseudopneumoniae with viridans group streptococci. *BMC Microbiol 12*, 77. doi:10.1186/1471-2180-12-77

Patel, S., and Ahmed, S. (2015). Emerging field of metabolomics: Big promise for cancer biomarker identification and drug discovery. *Journal of Pharmaceutical and Biomedical Analysis 107*, 63-74. doi:https://doi.org/10.1016/j.jpba.2014.12.020

Peterson, S.N., Meissner, T., Su, A.I., Snesrud, E., Ong, A.C., Schork, N.J., and Bretz, W.A. (2014). Functional expression of dental plaque microbiota. *Front Cell Infect Microbiol 4*, 108. doi:10.3389/fcimb.2014.00108

Pettigrew, M., Marks, L., Kong, Y., Gent, J., Roche-Hakansson, H., and Hakansson, A. (2014). Dynamic changes in the Streptococcus pneumoniae transcriptome during transition from biofilm formation to invasive disease upon influenza A virus infection. *Infect Immun 82*, 4607-4619.

Pfaller, M.A., Farrell, D.J., Sader, H.S., and Jones, R.N. (2012). AWARE Ceftaroline Surveillance Program (2008-2010): trends in resistance patterns among Streptococcus pneumoniae, Haemophilus influenzae, and Moraxella catarrhalis in the United States. *Clin Infect Dis 55 Suppl 3*, S187-193. doi:10.1093/cid/cis561

Pian, Y., Wang, P., Liu, P., Zheng, Y., Zhu, L., Wang, H., Xu, B., Yuan, Y., and Jiang, Y. (2015). Proteomics identification of novel fibrinogen-binding proteins of Streptococcus suis contributing to antiphagocytosis. *Frontiers in cellular and infection microbiology 5*, 19. doi:10.3389/fcimb.2015.00019

Prajapati, J.B. (2013). Genomic Analysis of Dairy Starter Culture Streptococcus thermophilus MTCC 5461. *Journal of Microbiology and Biotechnology 23*, 459-466. doi:10.4014/jmb.1210.10030

Richards, V.P., Choi, S.C., Pavinski Bitar, P.D., Gurjar, A.A., and Stanhope, M.J. (2013). Transcriptomic and genomic evidence for Streptococcus agalactiae adaptation to the bovine environment. *BMC Genomics 14*, 920. doi:10.1186/1471-2164-14-920

Richards, V.P., Palmer, S.R., Pavinski Bitar, P.D., Qin, X., Weinstock, G.M., Highlander, S.K., Town, C.D., Burne, R.A., and Stanhope, M.J. (2014). Phylogenomics and the dynamic genome evolution of the genus Streptococcus. *Genome Biol Evol 6*, 741-753. doi:10.1093/gbe/evu048

Russell, C., Rahman, A., and Mohammed, A.R. (2013). Application of genomics, proteomics and metabolomics in drug discovery, development and clinic. *Therapeutic Delivery 4*, 395-413. doi:10.4155/tde.13.4

Salzano, A.M., Novi, G., Arioli, S., Corona, S., Mora, D., and Scaloni, A. (2013). Mono-dimensional blue native-PAGE and bi-dimensional blue native/urea-PAGE or/SDS-PAGE combined with nLC-ESI-LIT-MS/MS unveil membrane protein heteromeric and homomeric complexes in Streptococcus thermophilus. *Journal of proteomics 94*, 240-261. doi:10.1016/j.jprot.2013.09.007

Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol 32*, 347-355. doi:10.1038/nbt.2842

Schouls, L.M., Schot, C.S., and Jacobs, J.A. (2003). Horizontal Transfer of Segments of the 16S rRNA Genes between Species of the Streptococcus anginosus Group. *Journal of Bacteriology 185*, 7241-7246. doi:10.1128/JB.185.24.7241-7246.2003

Schulz, C., and Hammerschmidt, S. (2013). Exploitation of physiology and metabolomics to identify pneumococcal vaccine candidates. *Expert Rev Vaccines 12*, 1061-1075. doi:10.1586/14760584.2013.824708

Sharma, C.M., and Vogel, J. (2014). Differential RNA-seq: the approach behind and the biological insight gained. *Curr Opin Microbiol 19*, 97-105. doi:10.1016/j.mib.2014.06.010

Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X.*, et al.* (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods 11*, 399-402. doi:10.1038/nmeth.2857

Sitkiewicz, I., Green, N.M., Guo, N., Bongiovanni, A.M., Witkin, S.S., and Musser, J.M. (2009). Transcriptome adaptation of group B Streptococcus to growth in human amniotic fluid. *PLoS One 4*, e6114. doi:10.1371/journal.pone.0006114

Sjöholm, K., Karlsson, C., Linder, A., and Malmström, J. (2014). A comprehensive analysis of the Streptococcus pyogenes and human plasma protein interaction network. *Molecular bioSystems 10*, 1698-1708. doi:10.1039/c3mb70555b

Song, J.-H., Ko, K.S., Lee, J.-Y., Yang Baek, J., Sup Oh, W., Sik Yoon, H., Jeong, J.-Y., and Chun, J. (2005). Identification of essential genes in Streptococcus pneumoniae by allelic replacement mutagenesis, Vol 19.

Sorek, R., and Cossart, P. (2010). Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet 11*, 9-16. doi:10.1038/nrg2695

Stackebrandt, E., Frederiksen, W., Garrity, G.M., Grimont, P.A.D., Kämpfer, P., Maiden, M.C.J., Nesme, X., Rosselló-Mora, R., Swings, J., Trüper, H.G.*, et al.* (2002). Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *International journal of systematic and evolutionary microbiology 52*, 1043-1047. doi:10.1099/ijs.0.02360-0

Stazic, D., and Voss, B. (2016). The complexity of bacterial transcriptomes. *J Biotechnol 232*, 69-78. doi:10.1016/j.jbiotec.2015.09.041

Steinert, J., Schiml, S., Fauser, F., and Puchta, H. (2015). Highly efficient heritable plant genome engineering using Cas9 orthologues from Streptococcus thermophilus and Staphylococcus aureus. *Plant J 84*, 1295-1305. doi:10.1111/tpj.13078

Suzuki, N., Seki, M., Nakano, Y., Kiyoura, Y., Maeno, M., and Yamashita, Y. (2005). Discrimination of Streptococcus pneumoniae from Viridans Group Streptococci

by Genomic Subtractive Hybridization. *Journal of Clinical Microbiology 43*, 4528-4534. doi:10.1128/JCM.43.9.4528-4534.2005

Sztajer, H., Szafranski, S.P., Tomasch, J., Reck, M., Nimtz, M., Rohde, M., and Wagner-Dobler, I. (2014). Cross-feeding and interkingdom communication in dual-species biofilms of Streptococcus mutans and Candida albicans. *ISME J 8*, 2256-2271. doi:10.1038/ismej.2014.73

Takahashi, N., and Nyvad, B. (2011). The role of bacteria in the caries process: ecological perspectives. *J Dent Res 90*, 294-303. doi:10.1177/0022034510379602

Takahashi, N., Washio, J., and Mayanagi, G. (2012). Metabolomic approach to oral biofilm characterization—A future direction of biofilm research. *Journal of Oral Biosciences 54*, 138-143. doi:10.1016/j.job.2012.02.005

Tang, F., Bossers, A., Harders, F., Lu, C., and Smith, H. (2013). Comparative genomic analysis of twelve Streptococcus suis (pro)phages. *Genomics 101*, 336-344. doi:10.1016/j.ygeno.2013.04.005

Tang, J. (2011). Microbial Metabolomics. *Current Genomics 12*, 391-403. doi:10.2174/138920211797248619

Teng, J.L., Huang, Y., Tse, H., Chen, J.H., Tang, Y., Lau, S.K., and Woo, P.C. (2014). Phylogenomic and MALDI-TOF MS analysis of Streptococcus sinensis HKU4T reveals a distinct phylogenetic clade in the genus Streptococcus. *Genome Biol Evol 6*, 2930-2943. doi:10.1093/gbe/evu232

Thomason, M., and Storz, G. (2010). Bacterial antisense RNAs: how many are there, and what are they doing? *Annul Rev Genet 44*, 167-188.

Virtaneva, K., Porcella, S.F., Graham, M.R., Ireland, R.M., Johnson, C.A., Ricklefs, S.M., Babar, I., Parkins, L.D., Romero, R.A., Corn, G.J.*, et al.* (2005). Longitudinal analysis of the group A Streptococcus transcriptome in experimental pharyngitis in cynomolgus macaques. *Proc Natl Acad Sci U S A 102*, 9014-9019. doi:10.1073/pnas.0503671102

Wadood, A., Jamal, A., Riaz, M., Khan, A., Uddin, R., Jelani, M., and Azam, S.S. (2018). Subtractive genome analysis for in silico identification and characterization of novel drug targets in Streptococcus pneumonia strain JJA. *Microbial Pathogenesis 115*, 194-198.

doi:https://doi.org/10.1016/j.micpath.2017.12.063

Wang, S., Yang, Y., Zhao, Y., Zhao, H., Bai, J., Chen, J., Zhou, Y., Wang, C., and Li, Y. (2016a). Sub-MIC Tylosin Inhibits Streptococcus suis Biofilm Formation and Results in Differential Protein Expression. *Frontiers in microbiology 7*, 384. doi:10.3389/fmicb.2016.00384

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet 10*, 57-63. doi:10.1038/nrg2484

Wang, Z., Li, M.Y., Peng, B., Cheng, Z.X., Li, H., and Peng, X.X. (2016b). GC-MS-Based Metabolome and Metabolite Regulation in Serum-Resistant Streptococcus agalactiae. *J Proteome Res 15*, 2246-2253. doi:10.1021/acs.jproteome.6b00215

Watanabe, S., Kirikae, T., and Miyoshi-Akiyama, T. (2013). Complete genome sequence of Streptococcus dysgalactiae subsp. equisimilis 167 carrying Lancefield group C antigen and comparative genomics of S. dysgalactiae subsp. equisimilis strains. *Genome Biol Evol 5*, 1644-1651. doi:10.1093/gbe/evt117

Wen, Y.T., Wang, J.S., Tsai, S.H., Chuan, C.N., Wu, J.J., and Liao, P.C. (2014). Label-free proteomic analysis of environmental acidification-influenced Streptococcus pyogenes secretome reveals a novel acid-induced protein histidine triad protein A (HtpA) involved in necrotizing fasciitis. *Journal of proteomics 109*, 90-103. doi:10.1016/j.jprot.2014.06.026

Wen, Z.T., Nguyen, A.H., Bitoun, J.P., Abranches, J., Baker, H.V., and Burne, R.A. (2011). Transcriptome analysis of LuxS-deficient Streptococcus mutans grown in biofilms. *Mol Oral Microbiol 26*, 2-18. doi:10.1111/j.2041-1014.2010.00581.x

Wishart, D.S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Discov 15*, 473-484. doi:10.1038/nrd.2016.32

Wu, Q., Tun, H.M., Leung, F.C., and Shah, N.P. (2014). Genomic insights into high exopolysaccharide-producing dairy starter bacterium Streptococcus thermophilus ASCC 1275. *Sci Rep 4*, 4974. doi:10.1038/srep04974

Wu, Z., Zhang, W., Lu, Y., and Lu, C. (2010). Transcriptome profiling of zebrafish infected with Streptococcus suis. *Microb Pathog 48*, 178-187. doi:10.1016/j.micpath.2010.02.007

Xu, K., Ren, C., Liu, Z., Zhang, T., Zhang, T., Li, D., Wang, L., Yan, Q., Guo, L., Shen, J., *et al.* (2015). Efficient genome engineering in eukaryotes using Cas9 from Streptococcus thermophilus. *Cell Mol Life Sci 72*, 383-399. doi:10.1007/s00018-014-1679-z

Xu, P., Alves, J.M., Kitten, T., Brown, A., Chen, Z., Ozaki, L.S., Manque, P., Ge, X., Serrano, M.G., Puiu, D., *et al.* (2007). Genome of the opportunistic pathogen Streptococcus sanguinis. *Journal of bacteriology 189*, 3166-3175. doi:10.1128/jb.01808-06

Xu, P., and Gunsolley, J. (2014). Application of metagenomics in understanding oral health and disease. *Virulence 5*, 424-432. doi:10.4161/viru.28532

Yan, S.K., Liu, R.H., Jin, H.Z., Liu, X.R., Ye, J., Shan, L., and Zhang, W.D. (2015). "Omics" in pharmaceutical research: overview, applications, challenges, and future perspectives. *Chin J Nat Med 13*, 3-21. doi:10.1016/S1875-5364(15)60002-4

Yang, X.Y., He, K., Du, G., Wu, X., Yu, G., Pan, Y., Zhang, G., Sun, X., and He, Q.Y. (2016a). Integrated Translatomics with Proteomics to Identify Novel Iron-Transporting Proteins in Streptococcus pneumoniae. *Front Microbiol 7*, 78. doi:10.3389/fmicb.2016.00078

Yang, X.Y., Shi, T., Du, G., Liu, W., Yin, X.F., Sun, X., Pan, Y., and He, Q.Y. (2016b). iTRAQ-Based Proteomics Revealed the Bactericidal Mechanism of Sodium New Houttuyfonate against Streptococcus pneumoniae. *Journal of agricultural and food chemistry 64*, 6375-6382. doi:10.1021/acs.jafc.6b02147

Yang, X.Y., Xu, J.Y., Wei, Q.X., Sun, X., and He, Q.Y. (2017). Comparative Proteomics of Streptococcus pneumoniae Response to Vancomycin Treatment. *Omics : a journal of integrative biology 21*, 531-539.

Yang, X.Y., Zhang, L., Liu, J., Li, N., Yu, G., Cao, K., Han, J., Zeng, G., Pan, Y., Sun, X., *et al.* (2015). Proteomic analysis on the antibacterial activity of a Ru(II) complex against Streptococcus pneumoniae. *Journal of proteomics 115*, 107-116. doi:10.1016/j.jprot.2014.11.018

Yoshida, A., Niki, M., Yamamoto, Y., Yasunaga, A., and Ansai, T. (2015). Proteome analysis identifies the Dpr protein of Streptococcus mutans as an important factor in the presence of early streptococcal colonizers of tooth surfaces. *PLoS ONE 10*,

e0121176. doi:10.1371/journal.pone.0121176

Yun-Juan, B., Zhong, L., Jeffrey, A.M., Shaun, W.L., Victoria, A.P., and Francis, J.C. (2015). CovRS-Regulated Transcriptome Analysis of a Hypervirulent M23 Strain of Group A Streptococcus pyogenes Provides New Insights into Virulence Determinants. *J Bacteriol 197*, 3191-3205.

Yuzenkova, Y., Gamba, P., Herber, M., Attaiech, L., Shafeeq, S., Kuipers, O.P., Klumpp, S., Zenkin, N., and Veening, J.W. (2014). Control of transcription elongation by GreA determines rate of gene expression in Streptococcus pneumoniae. *Nucleic Acids Res 42*, 10987-10999. doi:10.1093/nar/gku790

Zeng, L., and Burne, R. (2015). Sucrose- and Fructose-Specific Effects on the Transcriptome of Streptococcus mutans, as Determined by RNA Sequencing. *Appl Environ Microbiol 82*, 146-156.

Zhou, Y., Wang, Y., Deng, L., Zheng, C., Yuan, F., Chen, H., Bei, W., and Li, J. (2015). Evaluation of the protective efficacy of four novel identified membrane associated proteins of Streptococcus suis serotype 2. *Vaccine 33*, 2254-2260. doi:10.1016/j.vaccine.2015.03.038