



OPEN ACCESS

EDITED BY

Axel Cloeckaert,
Institut National de recherche pour
l'agriculture, l'alimentation et l'environnement
(INRAE), France

REVIEWED BY

Po-Xing Zheng,
Academia Sinica, Taiwan
Laura Maria Andrade De Oliveira,
Federal University of Rio de Janeiro, Brazil
Coralie Lemaire,
Centre Hospitalier Universitaire de Tours,
France

*CORRESPONDENCE

Barbara Spellerberg
✉ barbara.spellerberg@uniklinik-ulm.de

RECEIVED 17 March 2023

ACCEPTED 25 May 2023

PUBLISHED 15 June 2023

CITATION

Bauer R, Haider D, Grepels A, Roscher R,
Mauerer S and Spellerberg B (2023) Diversity of
CRISPR-Cas type II-A systems in *Streptococcus*
anginosus.

Front. Microbiol. 14:1188671.

doi: 10.3389/fmicb.2023.1188671

COPYRIGHT

© 2023 Bauer, Haider, Grepels, Roscher,
Mauerer and Spellerberg. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

Diversity of CRISPR-Cas type II-A systems in *Streptococcus anginosus*

Richard Bauer, Dorina Haider, Aline Grepels, Rebecca Roscher, Stefanie Mauerer and Barbara Spellerberg*

Institute of Medical Microbiology and Hygiene, University Medical Center, Ulm, Germany.

Streptococcus anginosus is a commensal Streptococcal species that is often associated with invasive bacterial infections. However, little is known about its molecular genetic background. Many Streptococcal species, including *S. anginosus*, harbor clustered regularly interspaced short palindromic repeats (CRISPR)-Cas systems. A CRISPR-Cas type II-A system as well as a type II-C system have been reported for this species. To characterize the CRISPR-Cas type II systems of *S. anginosus* in more detail, we conducted a phylogenetic analysis of Cas9 sequences from CRISPR-Cas type II systems with a special focus on streptococci and *S. anginosus*. In addition, a phylogenetic analysis of *S. anginosus* strains based on housekeeping genes included in MLST analysis, was performed. All analyzed Cas9 sequences of *S. anginosus* clustered with the Cas9 sequences of CRISPR type II-A systems, including the Cas9 sequences of *S. anginosus* strains reported to harbor a type II-C system. The Cas9 genes of the CRISPR-Cas type II-C systems of other bacterial species separated into a different cluster. Moreover, analyzing the CRISPR loci found in *S. anginosus*, two distinct *csn2* genes could be detected, a short form showing high similarity to the canonical form of the *csn2* gene present in *S. pyogenes*. The second CRISPR type II locus of *S. anginosus* contained a longer variant of *csn2* with close similarities to a *csn2* gene that has previously been described in *Streptococcus thermophilus*. Since CRISPR-Cas type II-C systems do not contain a *csn2* gene, the *S. anginosus* strains reported to have a CRISPR-Cas type II-C system appear to carry a variation of CRISPR-Cas type II-A harboring a long variant of *csn2*.

KEYWORDS

Streptococcus anginosus, CRISPR-Cas, classification, phylogenetic analysis, CSN2 gene

Introduction

Bacteria defend themselves against phage associated infection or other invading genetic elements (Barrangou and Marraffini, 2014). In this context the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) system has been described as a bacterial immunity system (Bernheim and Sorek, 2020). It contains distinct CRISPR-associated (*cas*) genes and a CRISPR array composed of unique spacer sequences interspersed with short repeats (Koonin and Makarova, 2009; Deveau et al., 2010; Le Rhun et al., 2019). The principle of the CRISPR-Cas system is based on the integration of new spacers derived from foreign genetic elements into the CRISPR array, that is further transcribed and processed into CRISPR RNAs (crRNAs), consisting of a part of the repeat and the spacer (Brouns et al., 2008; Wiedenheft et al., 2012; Makarova et al., 2015; Wei et al., 2015). Mature crRNAs guide Cas proteins to target sequences on invading

nucleic acid mediating specific cleavage (Marraffini and Sontheimer, 2008; Garneau et al., 2010; Hille et al., 2018).

Based upon *cas* gene content, repeat sequence and the organization of the CRISPR loci, CRISPR-Cas systems are currently divided into two classes including three main types each, ranging from type I to VI (class 1 comprising type I, III and IV; class 2 comprising type II, V, VI) and numerous subtypes (Shmakov et al., 2015; Le Rhun et al., 2019; Makarova et al., 2020). As opposed to class 1 systems employing multi-subunit Cas protein complexes, effector modules of class 2 systems only rely on a single Cas protein as corresponding type II systems with its prominent signature gene *cas9* (Sapranaukas et al., 2011; Koonin et al., 2017; Le Rhun et al., 2019). In type II systems the size of the *cas9* gene and the presence or absence of subtype-specific proteins besides Cas1, Cas2 and Cas9 are criterions for further subdivision into subtypes (II-A, II-C1, II-C2; Makarova et al., 2020). These type II systems are highly represented among pathogens (Fonfara et al., 2014; Louwen et al., 2014). While type II-A systems are characterized by the presence of a Csn2 protein, type II-C does not harbor any accessory protein (Mir et al., 2018; Makarova et al., 2020). Like Cas1 and Cas2, Csn2 is assumed to participate in spacer acquisition during the adaptation stage of CRISPR-mediated immunity (Heler et al., 2015; Wei et al., 2015).

Nearly half of all bacteria are equipped with at least one CRISPR-Cas system (Grissa et al., 2007). In streptococci all three major CRISPR types are represented with type I-C, type II-A and type III-A being the most frequently found subtypes (Louwen et al., 2014), whereas *S. pneumoniae* naturally lacks CRISPR-Cas loci (Bikard et al., 2012). Genome analysis of the *Streptococcus anginosus* group (SAG) revealed 7 of the 18 analyzed strains possessing a CRISPR-Cas system with most strains containing one CRISPR locus, typically a type II-A system (Olson et al., 2013). Type II CRISPR-Cas systems are widely distributed among a variety of different species, with type II-A representing the most commonly found subtype (Chylinski et al., 2014; Louwen et al., 2014). The type II-A system of *S. pyogenes* is very well studied and its Cas9 nuclease is extensively repurposed for genome engineering (Deltcheva et al., 2011; Jinek et al., 2012; Cebrian-Serrano and Davies, 2017). Type II-C CRISPR-Cas systems are the simplest type II systems regarding their structure and while nearly half of the multitude of Cas9 proteins discovered so far are part of the type II-C subtype, they have been sparsely investigated (Mir et al., 2018). Nevertheless, several type II-C systems in Streptococci were published and identified in public CRISPR databases as well (Gong et al., 2019; Lemaire et al., 2022). Considering phylogeny of Cas9, type II-A systems are assumed to be a derivative of type II-C, with the *csn2* gene acquired by type II-A ancestors, since the type II-A branch is embedded within type II-C (Chylinski et al., 2014).

Streptococcus anginosus is primarily a commensal of mucosal membranes colonizing many areas of the human body including the oral cavity, the gastrointestinal and the urogenital tract (Whiley et al., 1992; Paster et al., 2001; Facklam, 2002; Pilarczyk-Zurek et al., 2022). Together with the closely related species *Streptococcus constellatus* and *Streptococcus intermedius* it belongs to the SAG (Whiley and Beighton, 1991; Jensen et al., 2013). Bacteria of the SAG can frequently be isolated from blood cultures, abscesses, the respiratory tract of cystic fibrosis patients and have recently been associated with gastric cancer (Parkins et al., 2008; Grinwis et al., 2010; Mukae et al., 2016; Kobo et al., 2017; Zhou et al., 2022). Often overlooked in the past, *S. anginosus* has been increasingly identified in invasive infections

during the last years emphasizing its clinical importance as an emerging bacterial pathogen (Laupland et al., 2006; Reissmann et al., 2010; Siegman-Igra et al., 2012). Indeed, the incidence rate of SAG infections (8.65/100,000) even exceeds the combined incidence rates of group A and B streptococci in population-based surveillance studies (Laupland et al., 2006; Jiang et al., 2020).

We previously reported on the relationship between CRISPR-Cas systems and the presence of the β -hemolysin gene cluster in *Streptococcus anginosus* (Bauer et al., 2020). Similar to other human bacterial pathogens some *S. anginosus* strains harbor CRISPR-Cas type II systems (Olson et al., 2013; Louwen et al., 2014). To get a deeper insight into the CRISPR-Cas type II systems present in the species *S. anginosus* we performed a phylogenetic analysis of *S. anginosus* strains in our strain collection and publicly available whole genome sequences with a special focus on strains that were labeled as harboring a CRISPR-Cas type II-C locus.

Methods

Strain collection

Clinical isolates of *S. anginosus* strains originated from the University hospital Aachen and Ulm as described previously (Bauer et al., 2020). Streptococci were routinely grown on sheep blood agar plates (TSA + SB, Oxoid, Basingstoke, UK) at 37°C in a 5% CO₂ atmosphere. Liquid culture was performed in THY broth (Todd-Hewitt Broth [Oxoid] supplemented with 0.5% yeast extract [BD, Miami]). Strains used in this study are listed in [Supplementary Table S1](#).

Molecular methods

Bacterial genomic DNA was obtained according to manufacturer's instructions of standard commercial kits (GenElute™ Bacterial Genomic DNA Kits, Sigma-Aldrich, St. Louis, United States). For Polymerase Chain Reaction (PCR) Taq polymerase (Roche, Mannheim, Germany) was used with an initial denaturation step of 3 min at 94°C, 30 amplification cycles of 1 min at 94°C, 30 s at 50°C, 1 min at 72°C for MLST and 1.5 min at 72°C for CRISPR-Cas type II loci, followed by a final elongation step of 7 min at 72°C. Primers used for MLST and detection of CRISPR-Cas type II loci are listed in [Supplementary Table S2](#).

PCR-based identification of CRISPR-Cas type II systems in *Streptococcus anginosus*

Streptococcus anginosus whole genome sequences were analyzed for CRISPR-Cas type II systems by the CRIPRS finder program located at the following website: <https://crisprcas.i2bc.paris-saclay.fr/>. In the publicly available genomes of *S. anginosus*, two CRISPR-Cas type II systems can be detected that are encoded in two different genomic locations of GenBank entry NC_022239.1 (CRISPR_A: SANR_RS04955, SANR_RS04950; CRISPR_B: SANR_RS07450, SANR_RS07440). The genes encoding Cas9 proteins in these two different CRISPR loci display nucleotide differences allowing a

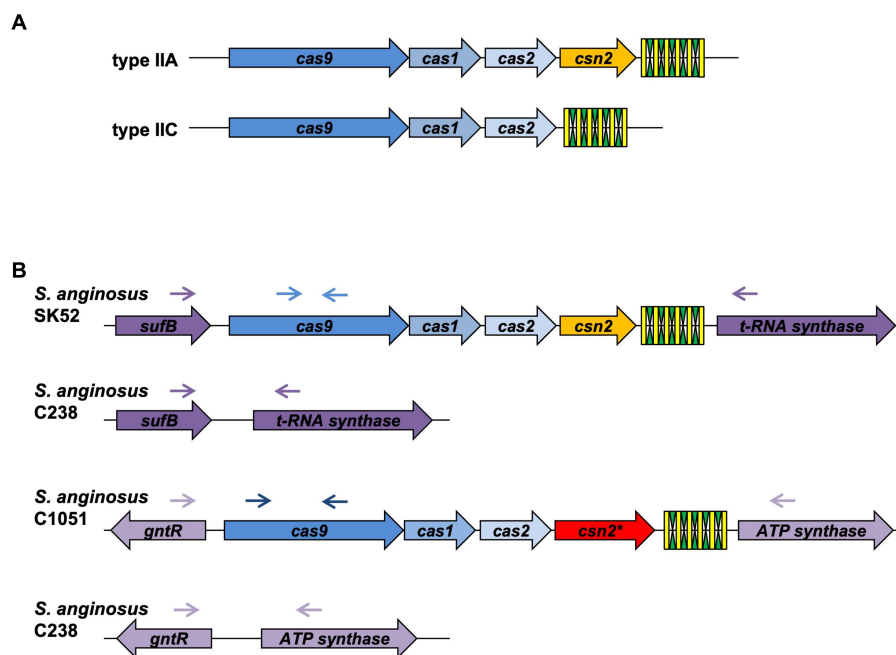


FIGURE 1

Genetic organization of CRISPR-Cas type II systems. (A) Conventional composition of CRISPR-Cas type II-A and II-C systems. (B) Genetic locus of CRISPR-Cas type II systems in *Streptococcus anginosus*. Green and yellow symbols represent repeat and spacer sequences of the CRISPR-Cas array. Genes present in all type II CRISPR-Cas systems are depicted in shades of blue, *Csn2* genes are shown in yellow for the canonical form and red for the long version. Lilac depicts genes, which are not part of the CRISPR-Cas locus. Arrows illustrate the location of primers used to screen for the presence and absence of CRISPR-Cas type II systems in *S. anginosus* isolates.

discrimination of the alleles by specific primers. The absence of a CRISPR locus in *S. anginosus* strains lacking a PCR product with primers targeting the two distinct CRISPR loci was confirmed through a set of primers adjacent to these typical CRISPR regions (SANR_RS04955, SANR_RS04950 and SANR_RS07450, SANR_RS07440). Primer sequences are listed in [Supplementary Table S2](#), primer binding sites are depicted in [Figure 1](#).

Phylogenetic analysis and statistics

The GenBank Database¹ served as source for nucleotide and protein sequences and the Basic Local Alignment Search Tool (BLAST) was used to identify homologous sequences. To correctly assign *S. anginosus* subspecies and genomosubgroups as described by [Jensen et al. \(2013\)](#) *S. anginosus* strains were subjected to phylogenetic analysis based on MLST data ([Figure 2](#)). For MLST analysis the sequence of seven housekeeping genes (*map*, *pfl*, *ppaC*, *pyk*, *rpoB*, *sodA*, and *tuf*) was determined and aligned as previously described ([Bishop et al., 2009](#)). MEGA version 7 was then used for phylogenetic analysis of the obtained sequences ([Kumar et al., 2018](#)) by applying the Minimal Evolution method ([Rzhetsky and Nei, 1993](#)). Evolutionary distances were computed using the Maximum Composite Likelihood method ([Tamura et al., 2004](#)). The rate variation among sites was modeled with a

gamma distribution (shape parameter = 20) and the minimum evolution tree was searched using the Close-Neighbor-Interchange (CNI) algorithm at a search level of 1. The initial tree was generated by using the Neighbor-joining algorithm ([Saitou and Nei, 1987](#)). The sequences of the seven housekeeping genes of *Streptococcus intermedius* (strain SV 101), *S. constellatus* (strain SV 019 and SV 100), *S. anginosus* (strain 557, 62CV, C1051, C238, F0211, J4206, J4211, OUP12, OUP25, SA1 and SK52) were retrieved from the GenBank database and served as reference in the analysis.

To examine the phylogeny of *S. anginosus* Cas9 variants, sequences of selected type II-A and II-C Cas9 nucleases ([Figure 3](#); [Supplementary Table S3](#)) as described previously in ([Chylinski et al., 2014](#); [Fonfara et al., 2014](#)) were aligned using MUSCLE algorithm with default parameters. MEGA version X was used for bootstrap analysis (100 replicates) of obtained sequences ([Felsenstein, 1985](#); [Kumar et al., 2018](#); [Stecher et al., 2020](#)) and conducted applying Minimal Evolution method ([Rzhetsky and Nei, 1993](#)). The evolutionary distances were computed using the JTT matrix-based method ([Jones et al., 1992](#)) and are depicted in the units of the number of amino acid substitutions per site. Rate variation among sites was modeled with a gamma distribution (shape parameter = 20). The minimal evolution tree was searched using CNI algorithm at a search level of 1 ([Nei and Kumar, 2000](#)). The Neighbor-joining algorithm was used to generate the initial tree ([Saitou and Nei, 1987](#)). Analysis involved 81 amino acid sequences and all ambiguous positions were removed for each sequence pair (pairwise deletion option).

The multiple sequence alignment of the *S. anginosus* Csn2 variants ([Figure 4](#)) was constructed using MUSCLE algorithm for each Csn2

¹ <http://www.ncbi.nlm.nih.gov/>

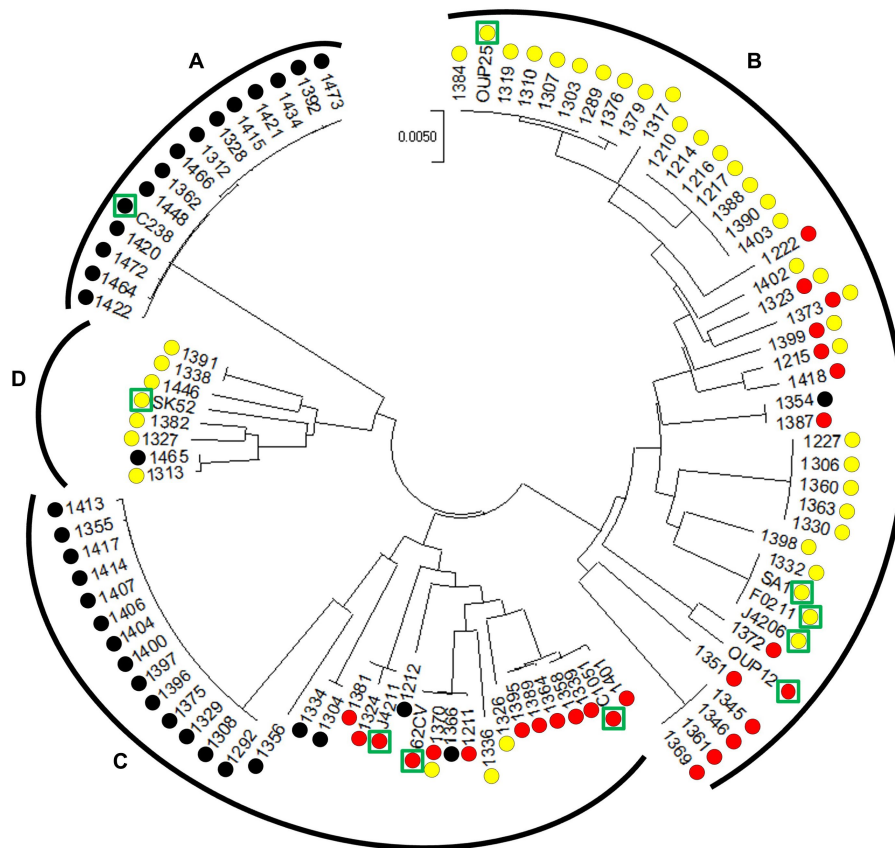


FIGURE 2

Circular minimal evolution tree constructed on the sequence of seven housekeeping genes illustrating the distribution of CRISPR-Cas type II systems in *S. anginosus* isolates. The units of the number of base substitutions per site are depicted. Subspecies [*S. anginosus* subsp. *whileyi* (A), *S. anginosus* subsp. *anginosus* (B)] and genomosubspecies [*S. anginosus* genomosubsp. *AJ1* (C), *S. anginosus* genomosubsp. *Vellorensis* (D)] are highlighted. Circles represent presence and absence of CRISPR-Cas type II systems. Yellow: type II-A. Red: type II-C. Black: no type II system. Green rectangles highlight sequenced strains derived from GenBank.

subfamily separately, with default settings. For each Csn2 subfamily, homologs were identified using HHPRED (Gabler et al., 2020) and obtained sequences *S. thermophilus* LMG 18311 (3ZTH) and *S. agalactiae* ATCC 13813 (3QHQ) were included.

GenBank entries of *S. anginosus cas9* and *csn2* newly obtained sequences can be accessed under the following numbers (OQ622068, OQ622069, OQ622070, OQ622071, OQ622072, OQ622073).

Results

Genomic organization of CRISPR-Cas in *Streptococcus anginosus*

Analyzing the CRISPR-Cas systems of whole genome sequences of *S. anginosus* strains available in GenBank and strains from our collection three different genotypes could be detected: Strains without any CRISPR arrays, isolates harboring one of two distinct CRISPR loci and strains harboring two different CRISPR arrays (Table 1). The most frequently found CRISPR locus displays the classical genomic organization of a CRISPR-Cas type II-A system with *cas9* encoding the nuclease followed by *cas1*, *cas2* and *csn2* genes and the CRISPR array (Figure 1). In all of the analyzed *S. anginosus* strains this typical type II-A locus is incorporated between the *sufB* and a tRNA synthase

gene. The second detected CRISPR locus also displays the canonical architecture of a CRISPR-Cas type II locus and is found between the *gntR* and an ATP synthase gene. At first glance it seems to be missing an easily identifiable *csn2* gene downstream of the *cas2* gene. However, closer analysis reveals that it contains a longer variant of a *csn2* gene in this genetic location. BLAST analysis of this *csn2* gene revealed closest homologies to the 350 amino acid (aa) long variant of the *csn2* gene of *Streptococcus thermophilus* LMG 18311 (42% identity and 65% similarity of the aa sequence; Figure 4). In each of the 27 strains containing this long variant of *csn2*, the corresponding CRISPR-Cas type II system was inserted downstream of *gntR*. CRISPR-Cas type I and CRISPR-Cas type III systems were not identified by the CRISPR finder program² for the *S. anginosus* strains listed in Table 1.

Phylogenetic analysis of *Streptococcus anginosus* strains

To analyze the population of *S. anginosus* in more detail and to investigate if strains carrying a specific CRISPR-Cas type II system cluster within specific subgroups of *S. anginosus*, a phylogenetic

² <https://crisprcas.i2bc.paris-saclay.fr/>

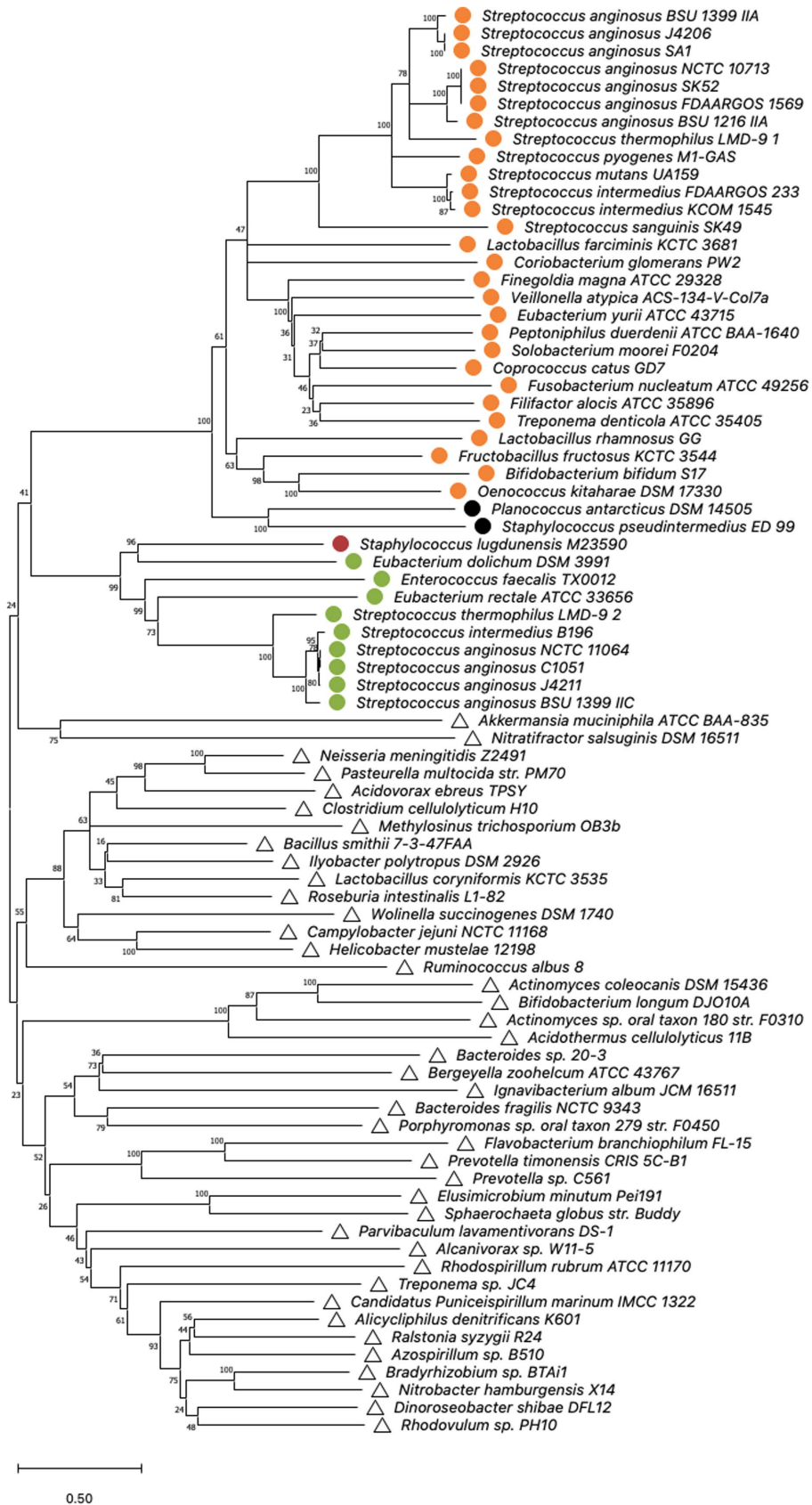


FIGURE 3
 Cas9 phylogeny. Minimal Evolution tree constructed on the sequences of selected type IIA and IIC cas9 nucleases published by Chylinski et al. (2014); Fonfara et al. (2014) by bootstrap analysis using MEGA X. Multiple alignment of Cas9 amino acid sequences was constructed using the MUSCLE

(Continued)

FIGURE 3 Continued

algorithm with default settings. The units of the number of amino acid substitutions per site are depicted. Type II-A systems are represented by circles, whereas type II-C systems are indicated by triangles. Corresponding *csn2* types stated by Chylinski et al. (2014) are highlighted. Orange: *csn2* type I, green: *csn2* type II, black: *csn2* type III and red: *csn2* type V.

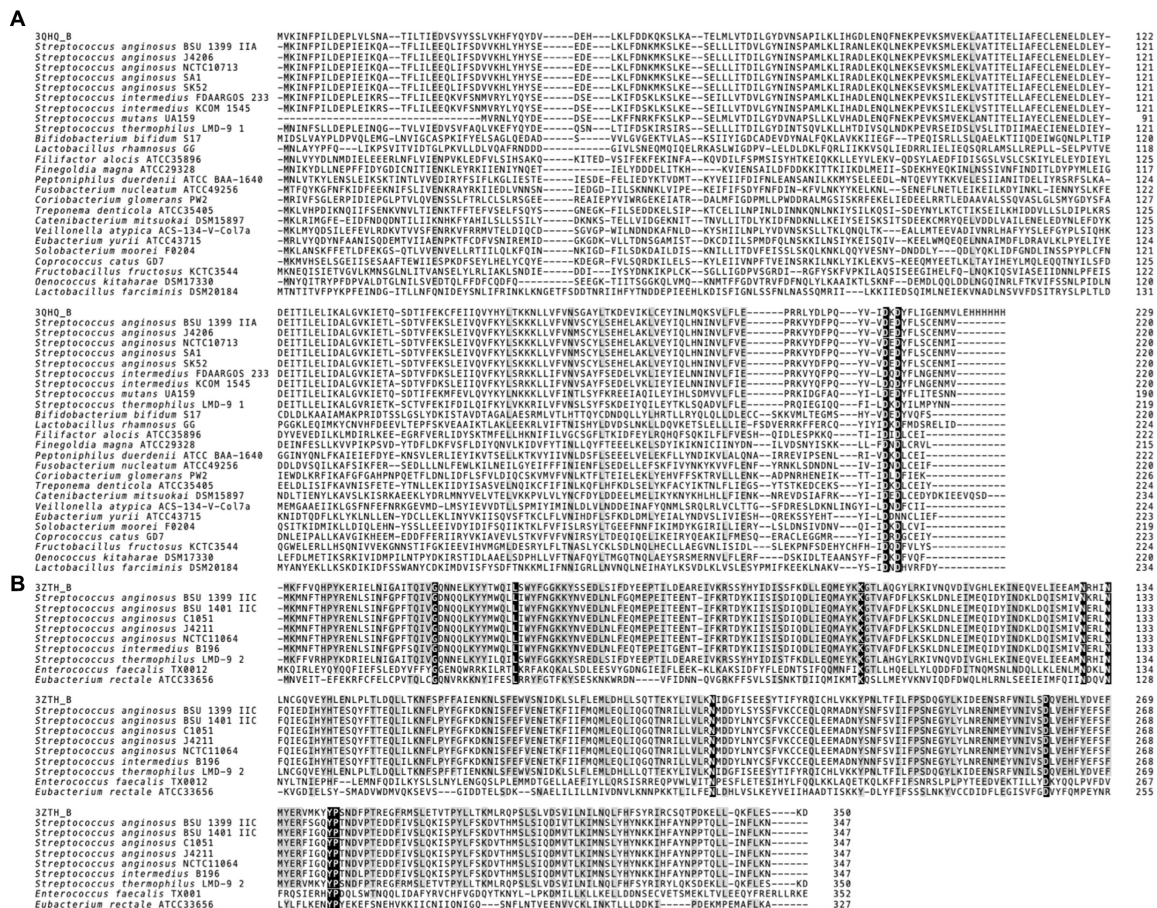


FIGURE 4 Multiple sequence alignment of Csn2 subfamilies found in *S. agalactiae*. (A) Canonical 220 aa variant (Csn2 type I). (B) 350 aa long Csn2 variant (Csn2 type II). For each Csn2 subfamily the multiple sequence alignment of selected protein sequences was constructed separately using the MUSCLE algorithm with default settings. Homologs of each Csn2 subfamily were identified using HHPRED. Retrieved Csn2 sequences of *S. thermophilus* LMG 18311 (PDB: 3ZTH) and *S. agalactiae* ATCC 13813 (PDB: 3QHQ) were included. The positions strongly conserved (>95% consensus) are shown by reverse shading, whereas above 70% consensus is highlighted in grey.

analysis based on housekeeping genes used for MLST analysis of viridans streptococci was conducted (Bishop et al., 2009). The analysis was carried out with the Mega7 program and included *S. anginosus* strains of clinical origin (Bauer et al., 2020) as well as whole genome sequences deposited in the GenBank database. The results showed a population structure representing the recognized subspecies and subgroups of *S. anginosus* as previously described (Jensen et al., 2013; Babbar et al., 2017; Figure 2). A correlation with the presence and absence of CRISPR-Cas type II systems demonstrated the absence of CRISPR-Cas in all 15 strains of *S. anginosus* subspecies *whyleyi*, while all of the 43 *S. anginosus* subspecies *anginosus* strains carried either one or both of the previously described CRISPR-Cas type II systems. The 34 strains of the *S. anginosus* genomosubspecies AJ displayed the most diverse picture, while 19 strains had no CRISPR-Cas type II

system, the remaining 15 strains harbored one of the two distinct CRISPR-Cas type II systems or both of them. In *S. anginosus* genomosubspecies *vellorensis* only the classical type II-A system could be detected in 7 of the 8 analyzed strains. In summary the strains harboring a CRISPR-Cas type II system, previously labeled as type II-C, did not cluster together within a certain subgroup of the *S. anginosus* strains.

Phylogeny of Cas9 nucleases from CRISPR type II-A and CRISPR type II-C systems

The structures and sequences of Cas9 nucleases associated with CRISPR-Cas type II systems have previously been shown to be distinct

TABLE 1 Clustered regularly interspaced short palindromic repeats (CRISPR) Cas system annotation of *Streptococcus anginosus* whole genome sequences.

Strain designation	CRISPR-Cas type II-A	CRISPR-Cas type II-C	none	# of spacers	Isolation source
<i>Streptococcus anginosus</i> BSU 1399	×	×		35	Soft tissue swab
<i>Streptococcus anginosus</i> C1051		×		29	Blood culture
<i>Streptococcus anginosus</i> C238			×	–	Respiratory tract
<i>Streptococcus anginosus</i> FDAARGOS_1155			×	–	Not available
<i>Streptococcus anginosus</i> FDAARGOS_1569	×			23	Not available
<i>Streptococcus anginosus</i> J4206	×			26	Blood culture
<i>Streptococcus anginosus</i> J4211		×		18	Hospital isolate
<i>Streptococcus anginosus</i> NCTC10713	×			23	Not available
<i>Streptococcus anginosus</i> NCTC11064		×		28	Materia alba tooth surface
<i>Streptococcus anginosus</i> SA1	×			35	Urine isolate
<i>Streptococcus anginosus</i> subsp. <i>whileyi</i> MAS624			×	–	Not available

for types II-A and II-C (Chylinski et al., 2014). To investigate if the Cas9 sequences of the two type II systems that are present in *S. anginosus* cluster either with type II-A or with type II-C a bootstrap analysis of available Cas9 sequences was performed using the MEGA X program (Figure 3). For comparison reasons, the analysis included Cas9 sequences from 39 different bacterial species including 18 streptococci carrying type II-A systems and 41 isolates from a variety of different bacterial species harboring CRISPR-Cas type II-C. All of the *S. anginosus* strains clustered together with the CRISPR type II-A Cas9 sequences, including the Cas9 sequences of strains reported to harbor a type II-C system. Also included in the analysis were Cas9 sequences from the GenBank database of *S. intermedius*, which belongs to the *S. anginosus* group. The bootstrap analysis located these sequences right next to the Cas9 sequences of *S. anginosus* strains within the CRISPR-Cas type II-A cluster, reflecting the close relationship between *S. intermedius* and *S. anginosus*. Interestingly none of the streptococcal sequences clustered in the vicinity of the Cas9 proteins of established CRISPR-Cas type II-C systems. In summary, the amino acid sequence analysis of available Cas9 proteins did not support the hypothesis that *S. anginosus* carries a CRISPR-Cas type II-C system.

Discussion

CRISPR-Cas type II systems are frequently found in pathogenic bacteria (Grissa et al., 2007; Jensen et al., 2013; Louwen et al., 2014), comprising Gram negative as well as Gram positive species. In comparison to type I and III, the CRISPR-Cas type II systems are simple, harboring only two to three CRISPR associated (Cas) genes. The canonical type II-A system of *Streptococcus pyogenes* consists of the gene encoding the Cas9 nuclease followed by *cas1*, *cas2* and *csn2* genes (Jansen et al., 2002; Le Rhun et al., 2019; Figure 1). Cas1 and Cas2 are conserved within CRISPR-Cas systems of different species. They function as endonucleases and mediate the spacer acquisition step of the CRISPR-Cas system (Nunez et al., 2014). Csn2 also interacts with double stranded DNA and participates in spacer acquisition (Ellinger et al., 2012; Lee et al., 2012; Chylinski et al., 2014). But, while Csn2 is always present in type II-A systems, it is

classically absent in CRISPR-Cas type II-C systems (Mir et al., 2018). Among pathogenic bacteria CRISPR-Cas type II-C systems are typically found in *Campylobacter*, *Corynebacteria*, *Neisseria* and *Pasteurella* species (Mir et al., 2018).

First genome projects showed the presence of typical CRISPR type II-A systems in *S. anginosus* strains (Olson et al., 2013), with several of the analyzed strains harboring more than one CRISPR array. Annotations in the CRISPR database (Grissa et al., 2007) show the typical CRISPR type II-A systems that have previously been described but also characterize several *S. anginosus* strains as harboring a CRISPR-Cas type II-C locus, since a typical *csn2* gene cannot readily be identified in the corresponding genomes (Gong et al., 2019; Lemaire et al., 2022). However, analyzing the CRISPR-Cas loci found in *S. anginosus* genomes that do not resemble the canonical type II-A system, we detected an additional gene downstream of *cas2*. Sequence comparison revealed significant homology to a Csn2 variant of *Streptococcus thermophilus* (Figure 4). This larger variant of Csn2, which has first been described in 2012 (Lee et al., 2012) is about 350 aa long and does not show any significant sequence homology to the canonical 220 aa long *csn2* gene of classical type II-A systems. The protein structure of this large Csn2 variant of *Streptococcus thermophilus* has been solved (Lee et al., 2012), showing a homotetramer that forms a central channel, binding linear double stranded DNA presumably through the interaction with basic amino acids. Its precise function remains however unclear, since neither a nuclease nor integrase activity has been demonstrated in functional assays. The Csn2 protein appears to be crucial for CRISPR-Cas type II-A systems, while it has not been found in any other CRISPR systems (Makarova et al., 2020). It has previously been proposed that CRISPR-Cas type II-A systems developed from CRISPR-Cas type II-C systems through the acquisition of a *csn2* gene (Chylinski et al., 2014).

To evaluate if phylogenetic analysis supports the hypothesis that *S. anginosus* carries two different CRISPR-Cas type II-A systems we performed Cas9 phylogeny. Cas9 phylogeny has previously been shown to allow a differentiation between different CRISPR-Cas systems (Chylinski et al., 2014). Comparison of the Cas9 sequences of *S. anginosus* present in the two distinct CRISPR-Cas type II systems with Cas9 sequences available in GenBank, showed a clustering of both *S. anginosus* Cas9 nucleases with Cas9 enzymes of other

CRISPR-Cas type II-A systems (Figure 3). The results further demonstrate a separate clustering of the Cas9 nucleases of CRISPR-Cas type II-C systems. None of the *S. anginosus* Cas9 sequences can be found in the type II-C clusters. Furthermore, an MLST analysis of multiple *S. anginosus* strains was performed to investigate, if the two distinct CRISPR-Cas type II systems associate with different *S. anginosus* subspecies or *S. anginosus* genomosubgroups (Figure 2). As observed in previous studies (Jensen et al., 2013; Babbar et al., 2017), known subspecies and genomosubgroups of the species are clearly identifiable. However, similar to the results of the Cas9 phylogeny, no indication of a clear separation of strains carrying one of the two type II systems can be detected.

With our enlarging knowledge about CRISPR-Cas systems it has become clear that besides their function in bacterial immunity CRISPR-Cas plays a role in virulence (Louwen et al., 2014) and may also have additional functions (Makarova et al., 2020). The association of CRISPR-Cas with virulence was first detected in *Francisella novicida*, where Cas9 inhibits transcription of a lipoprotein as an immune evasion mechanism (Weiss et al., 2007; Ratner et al., 2019). In streptococci several examples exist for the association of virulence and CRISPR-Cas. In *Streptococcus agalactiae* the highly virulent genetic lineage ST17 carries a different CRISPR-Cas profile (Lier et al., 2015). *Streptococcus mutans* strains carrying CRISPR-Cas display an increased biofilm formation and exopolysaccharide production, which is crucial for their virulence (Chen et al., 2017). In regard to hemolysin genes of *S. anginosus* we were able to show that the presence of *sag* genes is associated with the absence of CRISPR-Cas systems (Bauer et al., 2020). Increasing our knowledge about the diversity of *S. anginosus* CRISPR-Cas systems may thus also play a role in assessing the virulence potential of *S. anginosus* strains.

Taken together our data do not support the existence of a separate CRISPR Cas type II-C system in the *S. anginosus* strains we analyzed. In contrast to the currently used classifications a *csn2* variant could be detected in type II systems not fitting into the classical CRISPR-Cas type II-A structure. Phylogenetic analysis supports the hypothesis that these strains carry a variant CRISPR-Cas type II-A system.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession

number(s) can be found at: <https://www.ncbi.nlm.nih.gov/genbank/>, OQ622068, OQ622069, OQ622070, OQ622071, OQ622072, and OQ622073.

Funding

The work of RB and DH was supported through the Bausteine Program of the Medical Faculty, University of Ulm and the International Graduate School in Molecular Medicine Ulm.

Author contributions

BS and RB designed the study. BS, RB, and DH wrote the manuscript. RB, DH, AG, RR, and SM performed experiments and analyzed the data. AG, RR, and SM edited and modified the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1188671/full#supplementary-material>

References

- Babbar, A., Kumar, V. N., Bergmann, R., Barrantes, I., Pieper, D. H., Itzek, A., et al. (2017). Members of a new subgroup of *Streptococcus anginosus* harbor virulence related genes previously observed in *Streptococcus pyogenes*. *Int. J. Med. Microbiol.* 307, 174–181. doi: 10.1016/j.ijmm.2017.02.002
- Barrangou, R., and Marraffini, L. A. (2014). CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. *Mol. Cell* 54, 234–244. doi: 10.1016/j.molcel.2014.03.011
- Bauer, R., Neffgen, N., Grepfels, A., Furietsch, M., Mauerer, S., Barbaqadze, S., et al. (2020). Heterogeneity of *Streptococcus anginosus* ss-hemolysis in relation to CRISPR/Cas. *Mol. Oral. Microbiol.* 35, 56–65. doi: 10.1111/omi.12278
- Bernheim, A., and Sorek, R. (2020). The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.* 18, 113–119. doi: 10.1038/s41579-019-0278-2
- Bikard, D., Hatoum-Aslan, A., Mucida, D., and Marraffini, L. A. (2012). CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* 12, 177–186. doi: 10.1016/j.chom.2012.06.003
- Bishop, C. J., Aanensen, D. M., Jordan, G. E., Kilian, M., Hanage, W. P., and Spratt, B. G. (2009). Assigning strains to bacterial species via the internet. *BMC Biol.* 7:3. doi: 10.1186/1741-7007-7-3
- Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slikhuis, R. J., Snijders, A. P., et al. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960–964. doi: 10.1126/science.1159689
- Cebrian-Serrano, A., and Davies, B. (2017). CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm. Genome* 28, 247–261. doi: 10.1007/s00335-017-9697-4
- Chen, J., Li, T., Zhou, X., Cheng, L., Huo, Y., Zou, J., et al. (2017). Characterization of the clustered regularly interspaced short palindromic repeats sites in *Streptococcus mutans* isolated from early childhood caries patients. *Arch. Oral Biol.* 83, 174–180. doi: 10.1016/j.archoralbio.2017.07.023
- Chylinski, K., Makarova, K. S., Charpentier, E., and Koonin, E. V. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res.* 42, 6091–6105. doi: 10.1093/nar/gku241

- Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., et al. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607. doi: 10.1038/nature09886
- Deveau, H., Garneau, J. E., and Moineau, S. (2010). CRISPR/Cas system and its role in phage-bacteria interactions. *Annu. Rev. Microbiol.* 64, 475–493. doi: 10.1146/annurev.micro.112408.134123
- Ellinger, P., Arslan, Z., Wurm, R., Tschapek, B., MacKenzie, C., Pfeffer, K., et al. (2012). The crystal structure of the CRISPR-associated protein Csn2 from *Streptococcus agalactiae*. *J. Struct. Biol.* 178, 350–362. doi: 10.1016/j.jsb.2012.04.006
- Facklam, R. (2002). What happened to the streptococci: overview of taxonomic and nomenclature changes. *Clin. Microbiol. Rev.* 15, 613–630. doi: 10.1128/CMR.15.4.613-630.2002
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791. doi: 10.2307/2408678
- Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K. S., Lecrivain, A. L., Bzdrenga, J., et al. (2014). Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* 42, 2577–2590. doi: 10.1093/nar/gkt1074
- Gabler, F., Nam, S. Z., Till, S., Mirdita, M., Steinegger, M., Soding, J., et al. (2020). Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Curr. Protoc. Bioinformatics* 72: e108
- Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523
- Gong, T., Lu, M., Zhou, X., Zhang, A., Tang, B., Chen, J., et al. (2019). CRISPR-Cas Systems in Streptococci. *Curr. Issues Mol. Biol.* 32, 1–38. doi: 10.21775/cimb.032.001
- Grinwis, M. E., Sibley, C. D., Parkins, M. D., Eshaghrushan, C. S., Rabin, H. R., and Surette, M. G. (2010). Characterization of *Streptococcus milleri* group isolates from expectorated sputum of adult patients with cystic fibrosis. *J. Clin. Microbiol.* 48, 395–401. doi: 10.1128/JCM.01807-09
- Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinform.* 8:172. doi: 10.1186/1471-2105-8-172
- Heler, R., Samai, P., Modell, J. W., Weiner, C., Goldberg, G. W., Bikard, D., et al. (2015). Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 519, 199–202. doi: 10.1038/nature14245
- Hille, F., Richter, H., Wong, S. P., Bratovic, M., Ressel, S., and Charpentier, E. (2018). The biology of CRISPR-Cas: backward and forward. *Cells* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032
- Jansen, R., Embden, J. D., Gaastra, W., and Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43, 1565–1575. doi: 10.1046/j.1365-2958.2002.02839.x
- Jensen, A., Hoshino, T., and Kilian, M. (2013). Taxonomy of the Anginosus group of the genus *Streptococcus* and description of *Streptococcus anginosus* subsp. *whileyi* subsp. nov. and *Streptococcus constellatus* subsp. *viborgensis* subsp. nov. *Int. J. Syst. Evol. Microbiol.* 63, 2506–2519. doi: 10.1099/ijs.0.043232-0
- Jiang, S., Li, M., Fu, T., Shan, F., Jiang, L., and Shao, Z. (2020). Clinical characteristics of infections caused by *Streptococcus Anginosus* group. *Sci. Rep.* 10:9032. doi: 10.1038/s41598-020-65977-z
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829
- Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275
- Kobo, O., Nikola, S., Geffen, Y., and Paul, M. (2017). The pyogenic potential of the different *Streptococcus anginosus* group bacterial species: retrospective cohort study. *Epidemiol. Infect.* 145, 3065–3069. doi: 10.1017/S0950268817001807
- Koonin, E. V., and Makarova, K. S. (2009). CRISPR-Cas: an adaptive immunity system in prokaryotes. *F1000 Biol. Rep.* 1:95. doi: 10.3410/B1-95
- Koonin, E. V., Makarova, K. S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* 37, 67–78. doi: 10.1016/j.mib.2017.05.008
- Kumar, S., Stecher, G., Li, M., Nkay, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Laupland, K. B., Ross, T., Church, D. L., and Gregson, D. B. (2006). Population-based surveillance of invasive pyogenic streptococcal infection in a large Canadian region. *Clin. Microbiol. Infect.* 12, 224–230. doi: 10.1111/j.1469-0691.2005.01345.x
- Le Rhun, A., Escalera-Maurer, A., Bratovic, M., and Charpentier, E. (2019). CRISPR-Cas in *Streptococcus pyogenes*. *RNA Biol.* 16, 380–389. doi: 10.1080/15476286.2019.1582974
- Lee, K. H., Lee, S. G., Eun Lee, K., Jeon, H., Robinson, H., and Oh, B. H. (2012). Identification, structural, and biochemical characterization of a group of large Csn2 proteins involved in CRISPR-mediated bacterial immunity. *Proteins* 80, 2573–2582. doi: 10.1002/prot.24138
- Lemaire, C., Le Gallou, B., Lanotte, P., Mereghetti, L., and Pastuszka, A. (2022). Distribution, diversity and roles of CRISPR-Cas Systems in Human and Animal Pathogenic Streptococci. *Front. Microbiol.* 13:828031. doi: 10.3389/fmicb.2022.828031
- Lier, C., Baticle, E., Horvath, P., Haguenoer, E., Valentin, A. S., Glaser, P., et al. (2015). Analysis of the type II-A CRISPR-Cas system of *Streptococcus agalactiae* reveals distinctive features according to genetic lineages. *Front. Genet.* 6:214. doi: 10.3389/fgene.2015.00214
- Louwen, R., Staals, R. H., Endtz, H. P., van Baarlen, P., and van der Oost, J. (2014). The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol. Mol. Biol. Rev.* 78, 74–88. doi: 10.1128/MMBR.00039-13
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736. doi: 10.1038/nrmicro3569
- Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2020). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83. doi: 10.1038/s41579-019-0299-x
- Marraffini, L. A., and Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843–1845. doi: 10.1126/science.1165771
- Mir, A., Edraki, A., Lee, J., and Sontheimer, E. J. (2018). Type II-C CRISPR-Cas9 biology, mechanism, and application. *ACS Chem. Biol.* 13, 357–365. doi: 10.1021/acscchembio.7b00855
- Mukae, H., Noguchi, S., Naito, K., Kawanami, T., Yamasaki, K., Fukuda, K., et al. (2016). The importance of obligate anaerobes and the *Streptococcus anginosus* Group in Pulmonary Abscess: a clone library analysis using Bronchoalveolar lavage fluid. *Respiration Int. Rev. Thoracic Dis.* 92, 80–89. doi: 10.1159/000447976
- Nei, M., and Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. New York: Oxford University Press.
- Nunez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., and Doudna, J. A. (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* 21, 528–534. doi: 10.1038/nsmb.2820
- Olson, A. B., Kent, H., Sibley, C. D., Grinwis, M. E., Mabon, P., Ouellette, C., 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
- Parkins, M. D., Sibley, C. D., Surette, M. G., and Rabin, H. R. (2008). The Streptococcus milleri group—an unrecognized cause of disease in cystic fibrosis: a case series and literature review. *Pediatr. Pulmonol.* 43, 490–497. doi: 10.1002/ppul.20809
- Paster, B. J., Boches, S. K., Galvin, J. L., Ericson, R. E., Lau, C. N., Levanos, V. A., et al. (2001). Bacterial diversity in human subgingival plaque. *J. Bacteriol.* 183, 3770–3783. doi: 10.1128/JB.183.12.3770-3783.2001
- Pilarczyk-Zurek, M., Sitkiewicz, I., and Koziel, J. (2022). The clinical view on *Streptococcus anginosus* group—opportunistic pathogens coming out of hiding. *Front. Microbiol.* 13:956677. doi: 10.3389/fmicb.2022.956677
- Ratner, H. K., Escalera-Maurer, A., Le Rhun, A., Jaggavarapu, S., Wozniak, J. E., Crispell, E. K., et al. (2019). Catalytically active Cas9 mediates transcriptional interference to facilitate bacterial virulence. *Mol. Cell* 75:e495 doi: 10.1016/j.molcel.2019.05.029
- Reissmann, S., Friedrichs, C., Rajkumari, R., Itzek, A., Fulde, M., Rodloff, A. C., et al. (2010). Contribution of *Streptococcus anginosus* to infections caused by groups C and G streptococci, southern India. *Emerg. Infect. Dis.* 16, 656–663. doi: 10.3201/eid1604.090448
- Rzhetsky, A., and Nei, M. (1993). Theoretical foundation of the minimum-evolution method of phylogenetic inference. *Mol. Biol. Evol.* 10, 1073–1095.
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Sapranaukas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* 39, 9275–9282. doi: 10.1093/nar/gkr606
- Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., et al. (2015). Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol. Cell* 60, 385–397. doi: 10.1016/j.molcel.2015.10.008
- Siegmán-Igra, Y., Azmon, Y., and Schwartz, D. (2012). Milleri group streptococcus—a stepchild in the viridans family. *Eur. J. Clin. Microbiol. Infect. Dis.* 31, 2453–2459. doi: 10.1007/s10096-012-1589-7
- Stecher, G., Tamura, K., and Kumar, S. (2020). Molecular evolutionary genetics analysis (MEGA) for macOS. *Mol. Biol. Evol.* 37, 1237–1239. doi: 10.1093/molbev/msz312
- Tamura, K., Nei, M., and Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U. S. A.* 101, 11030–11035. doi: 10.1073/pnas.0404206101
- Wei, Y., Terns, R. M., and Terns, M. P. (2015). Cas9 function and host genome sampling in type II-A CRISPR-Cas adaptation. *Genes Dev.* 29, 356–361. doi: 10.1101/gad.257550.114

- Weiss, D. S., Brotcke, A., Henry, T., Margolis, J. J., Chan, K., and Monack, D. M. (2007). In vivo negative selection screen identifies genes required for *Francisella* virulence. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6037–6042. doi: 10.1073/pnas.0609675104
- Whiley, R. A., and Beighton, D. (1991). Emended descriptions and recognition of *Streptococcus constellatus*, *Streptococcus intermedius*, and *Streptococcus anginosus* as distinct species. *Int. J. Syst. Bacteriol.* 41, 1–5. doi: 10.1099/00207713-41-1-1
- Whiley, R. A., Beighton, D., Winstanley, T. G., Fraser, H. Y., and Hardie, J. M. (1992). *Streptococcus intermedius*, *Streptococcus constellatus*, and *Streptococcus anginosus* (the *Streptococcus milleri* group): association with different body sites and clinical infections. *J. Clin. Microbiol.* 30, 243–244. doi: 10.1128/jcm.30.1.243-244.1992
- Wiedenheft, B., Sternberg, S. H., and Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338. doi: 10.1038/nature10886
- Zhou, C. B., Pan, S. Y., Jin, P., Deng, J. W., Xue, J. H., Ma, X. Y., et al. (2022). Fecal signatures of *Streptococcus anginosus* and *Streptococcus constellatus* for noninvasive screening and early warning of gastric Cancer. *Gastroenterology* 162:e1918, 1933–1947.e18. doi: 10.1053/j.gastro.2022.02.015