AMERICAN SOCIETY FOR MICROBIOLOGY | Infection and Immunity®

MOLECULAR GENOMICS

Characterization of the Pathogenicity of *Streptococcus intermedius* TYG1620 Isolated from a Human Brain Abscess Based on the Complete Genome Sequence with Transcriptome Analysis and Transposon Mutagenesis in a Murine Subcutaneous Abscess Model

Noriko Hasegawa,[a,b] Tsuyoshi Sekizuka,[a] Yutaka Sugi,[a] Nobuhiro Kawakami,[c] Yumiko Ogasawara,[a] Kengo Kato,[a] Akifumi Yamashita,[a] Fumihiko Takeuchi,[a] Makoto Kuroda[a]

Laboratory of Bacterial Genomics, Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo, Japan[a]; Department of Biomolecular Science, Toho University, Chiba, Japan[b]; Department of Pediatrics, Toyonaka Municipal Hospital, Toyonaka, Osaka, Japan[c]

ABSTRACT *Streptococcus intermedius* is known to cause periodontitis and pyogenic infections in the brain and liver. Here we report the complete genome sequence of strain TYG1620 (genome size, 2,006,877 bp; GC content, 37.6%; 2,020 predicted open reading frames [ORFs]) isolated from a brain abscess in an infant. Comparative analysis of *S. intermedius* genome sequences suggested that TYG1620 carries a notable type VII secretion system (T7SS), two long repeat regions, and 19 ORFs for cell wall-anchored proteins (CWAPs). To elucidate the genes responsible for the pathogenicity of TYG1620, transcriptome analysis was performed in a murine subcutaneous abscess model. The results suggest that the levels of expression of small hypothetical proteins similar to phenol-soluble modulin β1 (PSMβ1), a staphylococcal virulence factor, significantly increased in the abscess model. In addition, an experiment in a murine subcutaneous abscess model with random transposon (Tn) mutant attenuation suggested that Tn mutants with mutations in 212 ORFs in the Tn mutant library were attenuated in the murine abscess model (629 ORFs were disrupted in total); the 212 ORFs are putatively essential for abscess formation. Transcriptome analysis identified 37 ORFs, including paralogs of the T7SS and a putative glucan-binding CWAP in long repeat regions, to be upregulated and attenuated *in vivo*. This study provides a comprehensive characterization of *S. intermedius* pathogenicity based on the complete genome sequence and a murine subcutaneous abscess model with transcriptome and Tn mutagenesis, leading to the identification of pivotal targets for vaccines or antimicrobial agents for the control of *S. intermedius* infections.

KEYWORDS brain abscess, genomics, murine model, *Streptococcus*, transposon mutagenesis, whole-genome sequence

Streptococcus intermedius, an intraoral commensal bacterium belonging to the *Streptococcus anginosus* group, is known to cause periodontitis and pyogenic infections in the brain and liver (1–3). Patients with invasive *S. intermedius* infections have significantly longer hospital stays and higher mortality rates than patients with other *S. anginosus* group infections (4), suggesting that species identification might be of importance for prognostication.

Received 19 October 2016 **Returned for modification** 11 November 2016 **Accepted** 19 November 2016

Accepted manuscript posted online 28 November 2016

Citation Hasegawa N, Sekizuka T, Sugi Y, Kawakami N, Ogasawara Y, Kato K, Yamashita A, Takeuchi F, Kuroda M. 2017. Characterization of the pathogenicity of *Streptococcus intermedius* TYG1620 isolated from a human brain abscess based on the complete genome sequence with transcriptome analysis and transposon mutagenesis in a murine subcutaneous abscess model. Infect Immun 85:e00886-16. https://doi.org/10.1128/IAI.00886-16.

Editor Nancy E. Freitag, University of Illinois at Chicago

Address correspondence to Makoto Kuroda, makokuro@niid.go.jp.

N.H. and T.S. contributed equally to this work.

S. intermedius causes brain and liver abscesses in the host; the brain abscess is a focal suppurative inflammation. Significant elevations of Th1 and Th17 cytokine levels have been detected in patients with brain abscesses caused by monomicrobial Gram-positive bacterial infections (e.g., those caused by *S. intermedius*), whereas a Th2 cytokine (interleukin-10) has been shown to be present in Gram-negative bacterial infections (e.g., those caused by *Bacteroides fragilis* and *Escherichia coli*) (5). Because bacterial brain abscesses are often caused by polymicrobial infections, testing by culture might not be appropriate for the detection of all associated pathogens. Massively parallel sequencing is a powerful tool for the simultaneous identification of multiple bacteria. Kommedal et al. suggested that *Aggregatibacter aphrophilus*, *Fusobacterium nucleatum*, and *S. intermedius* are key pathogens in the establishment of spontaneous polymicrobial brain abscesses; in particular, *S. intermedius* was frequently detected, with 24 *S. intermedius*-positive cases occurring among 52 patients (6).

S. intermedius bacteremia and liver abscesses are often seen in patients with a recent history of dental manipulation. Because *S. intermedius* is part of the commensal oral flora in humans, dental cleaning can cause bacteremia and seeding of the liver via the hematogenous route even in the absence of active oral infection (7). Indeed, a human epidural abscess caused by *S. intermedius* following a dental extraction has been reported (8). In terms of penetration through the blood-brain barrier, *Streptococcus pneumoniae* is the most well characterized of the streptococci (9), and the sialidase NanA, a cell wall-anchored protein (CWAP), contributes to the invasion by *S. pneumoniae* into human brain microvascular endothelial cells (hBMECs) (10). *S. intermedius* isolates carry an ortholog of the *nanA* gene (11), suggesting that *S. intermedius* might reach the brain via a mechanism similar to that employed by *S. pneumoniae*.

S. intermedius produces a unique hemolytic toxin, intermedilysin (ILY) (12), a cholesterol-dependent cytolysin (CDC) that binds specifically to human complement regulator CD59 (hCD59) (13) and that is thought to be the major virulence factor. However, specific factors involved in abscess formation have not been identified in *S. intermedius*, although it is significant that *S. intermedius* has been isolated from abscesses in children with life-threatening diseases (14).

Several complete or draft genome sequences of *S. intermedius* strains have been reported, including BA1, isolated from a human epidural abscess (15); B196, isolated from a human bronchopulmonary abscess in a patient with septic arthritis, osteomyelitis, and pyomyositis (16); and C270, isolated from a human bronchopulmonary abscess (16). In this study, we sequenced the entire genome of strain TYG1620, isolated from a brain abscess in an infant. The characterization of symptom-specific isolates, such as TYG1620, is crucial for understanding the pathogenicity of *S. intermedius*. We also performed a comprehensive genomics study with transcriptome analysis and random transposon (Tn) mutagenesis and identified a gene that was upregulated *in vivo* in a murine abscess model. Our findings suggest that unique phenol-soluble modulin (PSM)-like peptides and CWAPs are involved in the *in vivo* survival of *S. intermedius* TYG1620.

RESULTS AND DISCUSSION

Sequencing of *S. intermedius* TYG1620 complete genome. We obtained *S. intermedius* isolate TYG1620 from a brain abscess in a 16-month-old infant (Table 1 and Fig. 1A). TYG1620 showed strong aggregation in brain heart infusion (BHI) broth under anaerobic conditions (Fig. 1B). Whole-genome sequencing of *S. intermedius* TYG1620 was performed using an Illumina GAIIx platform (San Diego, CA, USA) (paired-end 81-mer; 6,314,564 total paired-end reads; N_{50}, 198,530 bp; estimated genome coverage, 944 times). The reads were assembled using CLC Genomics Workbench (v7.5) software, resulting in 77 contigs with a cumulative size of 1.95 Mb. The remaining gaps were closed using PCR, and the fragments obtained by pulsed-field gel electrophoresis (PFGE) were sequenced using an Illumina MiSeq platform (paired-end 300-mer). The reads were assembled using CLC Genomics Workbench (v7.5) or Platanus software (17) and the PRICE assembler (18). The complete genome sequence was verified by PFGE

TABLE 1 Background information for the *S. intermedius* strains[a]

Strain	Sequence status	Genome size (bp)	% GC content	No. of CDSs	No. of tRNAs	Avg length of CDS (nt)	% genome consisting of coding region	No. of pseudogenes	GenBank accession no.	Publication (PubMed identifer)	Extent of disease
TYG1620	Complete	2,006,877	37.6	2,007	60	874	87.9	13	AP014880	This study	Human brain abscess
B196	Complete	1,996,214	37.6	1,815	60	951	86.5	18	CP003857	24341328	Bronchopulmonary infection, sepsis, arthritis, osteomyelitis, pyomyositis
BA1	Draft	1,965,880	37.6	1,853	63	844	87.0	74	ANFT01000001-A	23405291	Human epidural abscess (complicated mastoiditis and osteomyelitis of the skull)
C270	Complete	1,960,728	37.6	1,778	60	949	86.0	24	CP003858	24341328	Bronchopulmonary infection
ATCC 27335	Draft	1,951,449	37.6	1,825	61	910	87.5	29	ATFK01000001-A	NA	NA
JTH08	Complete	1,933,610	37.6	1,793	67	978	86.2	28	NC_018073	NA	NA
SK54	Draft	1,919,718	37.6	1,879	28	899	88.4	0	AJKN01000001-A	NA	NA
F0413	Draft	1,921,346	37.6	1,812	62	885	88.3	0	AFXO01000001-A	NA	NA

[a]CDS, coding sequences; NA, not available.

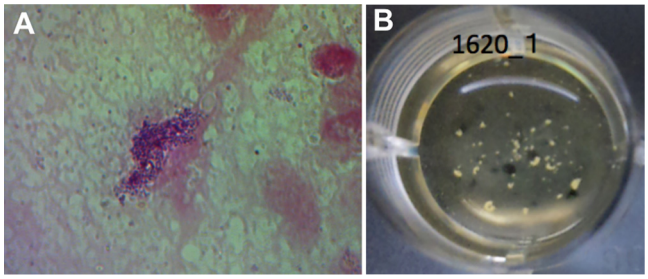

FIG 1 Laboratory testing of TYG1620. (A) Gram stain of the human brain abscess. Dark purple staining suggested a Gram-positive bacterial infection. (B) *Streptococcus intermedius* was isolated from the human brain abscess, and the overnight culture of *S. intermedius* TYG1620 (1620_1) in BHI broth under anaerobic condition showed strong bacterial cell aggregation.

using the restriction enzyme AscI (data not shown) and optical mapping with an Argus system (Fig. 2A).

Genome annotation was performed in Rapid Annotation using the Subsystems Technology (RAST; v2.0) (19), InterPro (v49.0) (20), and NCBI BLASTp/BLASTx programs. The final TYG1620 genome sequence was 2,006,877 bp and had a GC content of 37.6%, 60 tRNA genes, 4 rRNA operons, 2,007 predicted coding sequences, and 13 pseudogenes (Table 1 and Fig. 2B). The TYG1620 genome sequence had 19 CWAPs containing an LPXTG motif recognized by sortase A (Fig. 3). Of the CWAPs, SITYG_16380 had 6 tandem repeats of its 321-bp core unit, showing a collagen-binding B motif. Such long tandem repeats were highly problematic in closing the gap between scaffolds (see Fig. S1 in the supplemental material).

Comparative genomics among *S. intermedius* genome sequences. To elucidate specific genetic features of TYG1620, a comparative analysis of the TYG1620 genome sequence against other publicly available genome sequences was performed (Table 1). TYG1620 carried a potential type VII secretion system (T7SS) between SITYG_01160 and SITYG_01590 (Fig. 2B and C) that was integrated between *adhE* (bifunctional aldehyde/alcohol dehydrogenase) and *nanE* (*N*-acetylmannosamine-6-phosphate 2-epimerase). SITYG_1160 (96 amino acids [aa]) is similar to the *Mycobacterium* ESAT-6, which has been reported to be of fundamental importance in the virulence and protective immunity of *Mycobacterium tuberculosis* (21). The T7SS has been well characterized to be involved in the secretion of ESX proteins in *M. tuberculosis* and has a role in host-pathogen interactions (22). T7SSs have been found in members of the *Firmicutes*, such as *Bacillus* and *Clostridium* spp.; *Staphylococcus aureus*; *Streptococcus agalactiae*; and *Listeria monocytogenes* (23, 24). The results of the comparative genome analysis performed in the present study suggested that TYG1620 acquired T7SS in a horizontal manner and that it partially shares its T7SS with that from strains BA1, isolated from a human epidural abscess, and F0413, isolated from an unknown source (Fig. 2B), indicating that the T7SS could be associated with strain-specific pathogenicity, including abscess formation and septic arthritis.

Two homologous long repeats (44.0 kb; from nucleotides [nt] 70774 to 114782 and nt 1084062 to 1128070) were identified in the TYG1620 genome, and their localizations were confirmed by PFGE and optical mapping (Fig. 2A). The repeat sequence included possible plasmid-mediated DNA methyltransferases (SITYG_00620 and SITYG_00630) and a conjugal transfer protein and was present in strains SK54, ATCC 27335, and JTH08 but not strains BA1, B196, and F0413, suggesting that the long repeats could be strain-specific elements acquired from a certain plasmid. Intriguingly, two CWAPs were located in the repeat region; specifically, the putative glucan-binding protein SITYG_00870 (SITYG_10910 was on the second repeat) is a noteworthy CWAP that may be a potential virulence factor in the *in vivo* survival experiment described below (see Table 4, Fig. 5, and "Comparative analysis of the transcriptome *in vitro* and in the abscess *in vivo*" below).

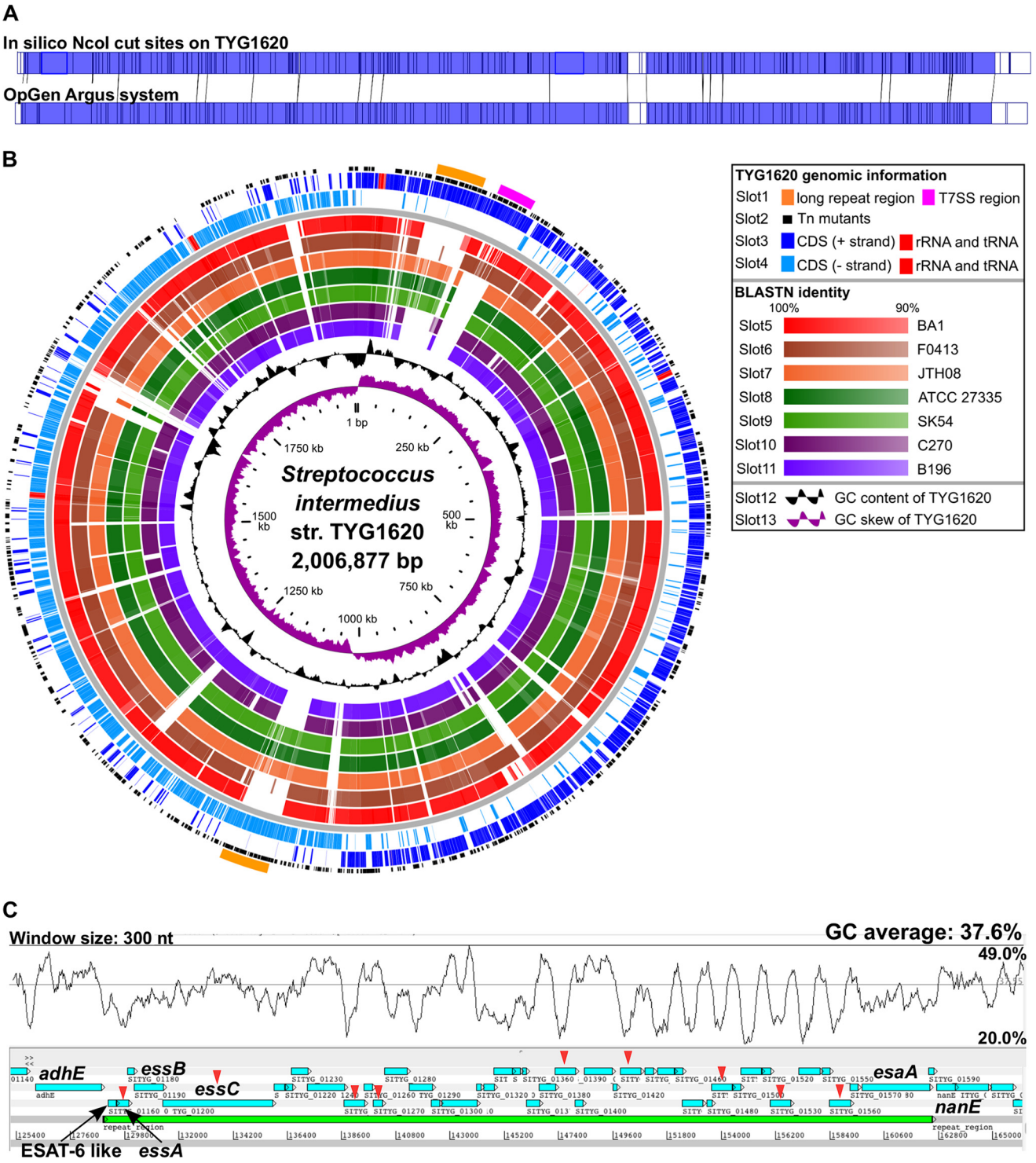

FIG 2 Basic genome information for *S. intermedius* strain TYG1620. (A) Optical mapping of TYG1620 genome DNA was performed using an Argus system. The upper and lower barcodes show the *in silico* NcoI restriction site on the TYG1620 complete chromosome DNA sequence and the actual NcoI-digested sites on the TYG1620 chromosome DNA molecule detected by the OpGen Argus system, respectively. The dark blue regions indicate the identical scaffoldings between the *in silico* and actual digestions. (B) Circular representation of the strain (str.) TYG1620 genome compared with the genomes of other *S. intermedius* strains. From the outside, slots 1 to 4 show TYG1620 genomic information (slot 1, TYG1620 genomic islands [T7SS]; slot 2, the IS*S1* insertion site in the Tn mutant library constructed by random transposon mutagenesis; slots 3 and 4, coding sequences and RNAs for the positive and minus strands, respectively); slots 5 to 11 show the results of comparative genome analysis of the TYG1620 genome with the genomes of *S. intermedius* strains BA1, F0413, JTH08, ATCC 27335, SK54, C270, and B196, respectively (the homology results are displayed as colored circles, as indicated in the box, with increasing color intensity signifying increased similarity); slot 12 shows the GC content; and slot 13 shows the GC content skew. (C) The gene organization of a possible T7SS region is shown as a light green bar. ORFs are shown as light blue boxes. The GC content with a 300-nt window is shown above the T7SS region. The vertical red arrowheads indicate the ORFs that were identified by a Tn mutant survival assay in the murine subcutaneous abscess model (see "Identification of *in vivo*-attenuated Tn mutants in a murine subcutaneous abscess model" in Materials and Methods and Fig. 5).

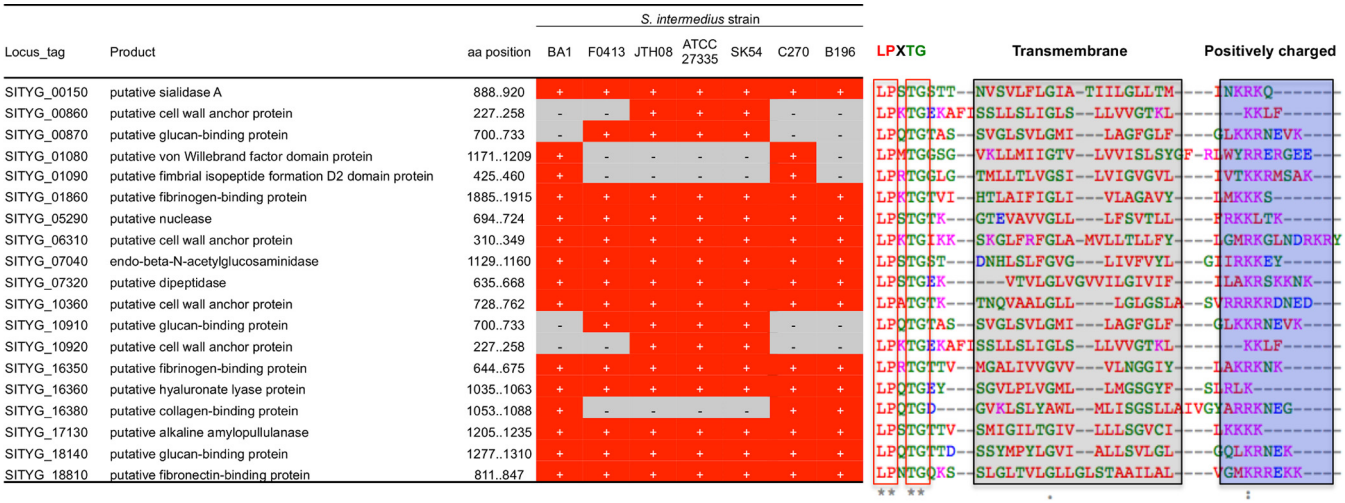

FIG 3 LPXTG motif-positive CWAPs in TYG1620. CWAPs are known to act as microbial surface proteins and contain an LPxTG cell-wall anchored motif recognized by sortase A, followed by a single transmembrane region and highly positive charged amino acids. TYG1620 has 19 CWAPs, and comparative genome analysis suggested that the differential possession of these CWAPs might play a role in strain-specific pathogenicity.

CWAPs. CWAPs are known to act as microbial surface components that recognize adhesive matrix molecules (MSCRAMMs), but in most cases their actual functions remain to be characterized. The TYG1620 genome sequence has 19 CWAPs containing an LPXTG motif recognized by sortase A; our comparative genome analysis suggested that these CWAPs might play a role in strain-specific pathogenicity (Fig. 3). Possible MSCRAMMs were predicted to be a fibronectin-binding protein (SITYG_18810), fibrinogen-binding proteins (SITYG_01860 and SITYG_16350), a collagen-binding protein (SITYG_16380), and putative glucan-binding proteins (SITYG_10910 and SITYG_18140). In addition, potential tissue-degrading factors, such as hyaluronate lyase, sialidase *nanA*, pullulanase, nuclease, and end-*β*-*N*-acetylglucosaminidase, were also identified.

Streptococcal genome sequences have indicated that streptococci generally carry multiple CWAPs; the redundancy depends on the species, as other members of the oral/pharyngeal bacterial flora, such as *Streptococcus mutans*, *Streptococcus mitis*, *Streptococcus pneumoniae*, and *Streptococcus pyogenes*, each carry multiple CWAPs (6, 14, 15, and 15, respectively) (25), suggesting that *S. intermedius* TYG1620 has a potential pathogenicity in the etiology of human brain abscesses.

Comparative analysis of the transcriptome *in vitro* and in the abscess *in vivo*. Although the comparative genome analysis revealed several noteworthy genetic features of TYG1620, the virulence factors crucial for the pathogenicity of *S. intermedius* remained to be clarified. Therefore, we established a murine subcutaneous abscess model and performed comparative transcriptome analysis to identify *in vivo*-specific gene expression in TYG1620 (Table S1). Indeed, subcutaneous inoculation of TYG1620 ($\sim 10^7$ CFU) into C57BL/6J mice resulted in reproducible abscess formation within 5 days in every trial (Fig. 4A). As a control for RNA expression, gene expression in two distinct growth phases (the early log and stationary phases) in BHI broth was selected for comparison with gene expression in the abscess (Table S2). At first we speculated that the gene expression in the stationary phase might be a better control for normalization because TYG1620 in an abscess could be subjected to diverse stresses, including conditions with low levels of nutrients and the host immune response. However, contrary to our speculation, a very good coefficient of determination (R^2 = 0.919) was obtained when gene expression in the early log phase was compared with that in the abscess, suggesting that early-log-phase cells in BHI broth might serve as a better control than stationary-phase cells (Fig. 4B and C).

Compared with the levels of expression in the early log phase, 279 open reading frames (ORFs) had significantly (≥3-fold) increased levels of expression in the abscess

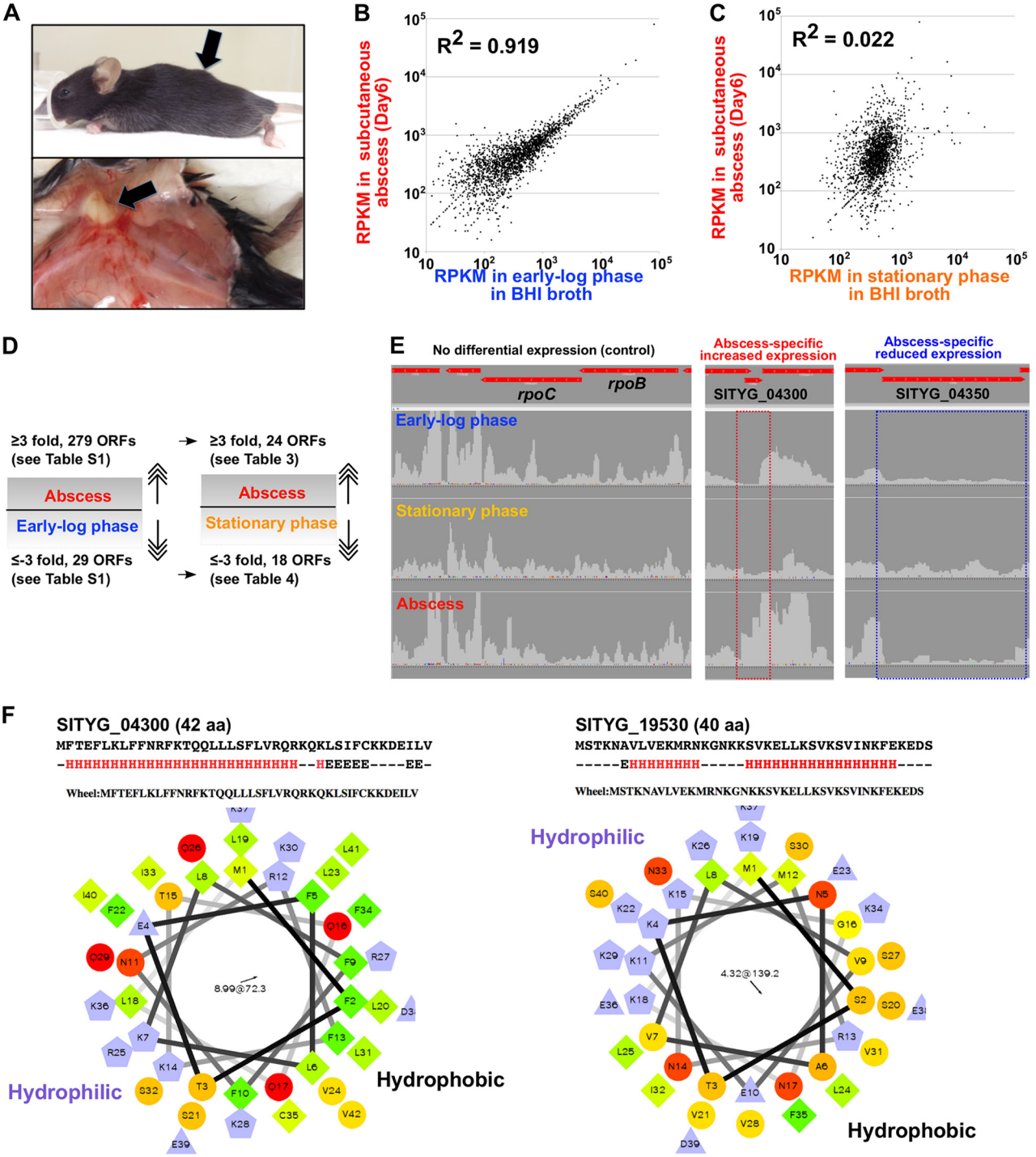

FIG 4 Mouse model of TYG1620-infected subcutaneous abscess. (A) A subcutaneous abscess on the back of a C57BL/6J mouse (arrow). (B) Comparative transcriptome analysis of 2,020 genes (including 13 pseudogenes) in TYG1620 from a subcutaneous abscess on day 6 and from *in vitro* early-log-phase growth in BHI broth. Each plot corresponds to the number of reads per kilobase per million (RPKM) of each gene. (C) Comparative transcriptome analysis of 2,020 genes (including 13 pseudogenes) in TYG1620 from a subcutaneous abscess on day 6 and from *in vitro* stationary-phase growth in BHI broth. (D) Scheme for the identification of *in vivo*-related gene expression from the transcriptome analyses. The coefficient of determination (R^2) suggested that gene expression in the early log phase in BHI broth could be a better control than that in the stationary phase; thus, significant (≥3-fold) up- or downregulated expression was first selected, followed by the use of expression at the stationary phase as a secondary control. This allowed the exclusion of general stress response genes that are not involved in *in vivo*-specific expression. (E) Validation using a read-mapping analysis for abscess-specific increased or decreased expression. Results for RNA polymerase subunit genes (*rpoB* and *rpoC*), used as internal expression controls, are shown. (F) SITYG_04300 is one of the genes with increased expression in the abscess (see Table 3 and panel E). Secondary structure prediction (JPred, v4) and a helical wheel projection suggested that the amino acid sequence (SITYG_04300 and SITYG_19530) shows an α-helical structure as amphipathic peptides, similar to PSMs. The H and E below the amino acid sequence represent α-helix and β-sheet, respectively. The number and arrow in the middle of the wheels correspond to the magnitude and the angle of the hydrophobic moment of the helical wheel, respectively.

TABLE 2 Abscess-specific increased gene expression

S. intermedius TYG1620 genome annotation			RNA-seq results				
			RPKM[a]			Expression ratio (fold change)	
Locus tag identifier	Product	Length (no. of aa)	Early log phase (BHI, 6 h)	Stationary phase (BHI, 24 h)	Subcutaneous abscess	Abscess/BHI, 6 h	Abscess/BHI, 24 h
SITYG_00160	Competence-specific global transcription modulator	157	166.3	139.1	614.4	4.0	3.7
SITYG_01550[b]	Hypothetical protein	96	57.0	0.0	276.1	5.3	>276.1
SITYG_02340	Hypothetical protein	210	249.0	173.6	1,031.2	4.5	5.0
SITYG_02390	Hypothetical protein	86	79.5	0.0	500.2	6.9	>500.2
SITYG_04050	Hypothetical protein	55	74.1	130.8	836.8	12.3	5.3
SITYG_04300[c]	Hypothetical protein	42	32.2	170.3	1,556.9	52.7	7.6
SITYG_06280	Hypothetical protein	83	0.0	87.2	318.8	>318.8	3.0
SITYG_07460	Hypothetical protein	42	32.2	0.0	389.2	13.2	>389.2
SITYG_07920	Hypothetical protein	244	28.2	59.8	368.9	14.2	5.1
SITYG_08020	Hypothetical protein	87	0.0	0.0	152.2	>152.2	>152.2
SITYG_08690	Hypothetical protein	57	262.2	0.0	865.7	3.6	>865.7
SITYG_08800	Hypothetical protein	258	165.5	141.4	542.8	3.6	3.2
SITYG_10090	Hypothetical protein	110	99.7	132.0	542.8	5.9	3.4
SITYG_10160	Hypothetical protein	54	0.0	0.0	121.7	>121.7	>121.7
SITYG_11120	Hypothetical protein	143	28.8	50.9	186.0	7.0	3.0
SITYG_12600	Putative membrane protein	187	58.8	272.7	997.1	18.5	3.0
SITYG_15030	Gcn5-related *N*-acetyltransferase	165	141.6	132.4	665.4	5.1	4.2
SITYG_16520	Hypothetical protein	62	21.9	116.3	531.3	26.4	3.8
SITYG_17490	Hypothetical protein	109	25.1	133.2	486.9	21.1	3.0
SITYG_19180	Hypothetical protein	172	55.9	84.7	483.7	9.4	4.8
SITYG_19530[c]	Hypothetical protein	40	33.7	178.7	979.7	31.6	4.6
SITYG_19560	Hypothetical protein	40	0.0	0.0	979.7	>979.7	>979.7
SITYG_19770	Hypothetical protein	54	100.6	133.2	1,338.9	14.5	8.4
SITYG_20170	Hypothetical protein	48	56.4	0.0	204.9	4.0	>204.9

[a]RPKM, number of reads per kilobase per million.
[b]ORF located in T7SS (Fig. 2C).
[c]Potential candidates for PSMs; see prediction of the secondary structure in Fig. 4F.

(Fig. 4D). Using the level of expression at the stationary phase as a secondary control, 24 ORFs were further identified to have significantly (≥3-fold) increased expression in the abscess (Table 2 and Fig. 4D). This highly stringent selection process, which excluded general stress response genes, enabled the identification of the specific ORFs with the most potential to be involved in the pathogenicity of *S. intermedius* expressed *in vivo*. Of 24 ORFs, 21 ORFs coded for hypothetical proteins, some of which were remarkably small ORFs, coding for less than 100 aa. Read mapping by next-generation sequencing indicated increased expression of some small ORFs in the abscess (see, e.g., SITYG_04300 in Fig. 4E). Some of these, including SITYG_04300 (Fig. 4F), appeared to have an α-helical structure similar to that of phenol-soluble modulins (PSMs), such as cytolysin, which is involved in staphylococcal pathogenesis in various eukaryotic cell types (26, 27). Secondary structure prediction suggested that two ORFs (SITYG_04300 and SITYG_19530) showed amphipathic characteristics and could be potential PSMs similar to *S. aureus* PSMβ1 (Fig. 4F). Recently, staphylococcal PSMs have been implicated in skin and soft tissue infection (SSTI) (28–30) and have shown significantly higher levels of expression in methicillin-resistant *S. aureus* (MRSA) strains isolated from SSTIs. These observations imply that the small ORFs detected in the present study might be candidates that contribute to *in vivo* abscess formation.

Although 18 ORFs with significantly (≥3-fold) decreased expression in the abscess were identified (Table 3 and Fig. 4D), there was no notable finding, such as immune escape by downregulation; thus, expression could be affected by a different culture broth *in vitro*.

Identification of *in vivo*-specific gene expression by Tn mutagenesis in a murine abscess model. A random transposon (Tn) mutant library of TYG1620 was constructed by the insertion of IS*S1* generated from the temperature-sensitive plasmid

TABLE 3 Abscess-specific decreased gene expression

S. intermedius TYG1620 genome annotation			RNA-seq RPKM[a]			Expression ratio (fold change)	
Locus tag identifier	Gene	Product	Early log phase (BHI, 6 h)	Stationary phase (BHI, 24 h)	Subcutaneous abscess	Abscess/BHI, 6 h	Abscess/BHI, 24 h
SITYG_00400	*purE*	Phosphoribosylaminoimidazole carboxylase catalytic subunit PurE	161.0	501.7	45.9	−3.23	−13.12
SITYG_00790		Hypothetical protein	111.3	168.4	0	<−111.3	<−168.4
SITYG_01360		Hypothetical protein	112.9	299.0	0	<−112.9	<−299.0
SITYG_02040		Hypothetical protein	212.7	1,502.5	57.2	−3.42	−31.53
SITYG_04350		Cysteine desulfurase, SufS subfamily	249.0	427.7	57.0	−4.01	−9.01
SITYG_05640		Hypothetical protein	181.9	241.0	22.0	−7.60	−13.15
SITYG_08670		Hypothetical protein	153.6	406.9	46.5	−3.03	−10.50
SITYG_08810		Gcn5-related *N*-acetyltransferase	216.6	353.0	60.5	−3.29	−7.01
SITYG_08820		Hypothetical protein	172.8	238.9	36.4	−4.36	−7.88
SITYG_09140	*coaA*	Pantothenate kinase	85.6	71.6	21.8	−3.61	−3.94
SITYG_09500		Hypothetical protein	109.2	192.8	0	<−109.2	<−192.8
SITYG_11090		Hypothetical protein	36.8	260.1	9.9	−3.43	−31.41
SITYG_14150		Hypothetical protein	188.6	222.0	0	<−188.6	<−222.0
SITYG_14680		Hypothetical protein	75.3	471.4	16.6	−4.18	−34.09
SITYG_16420		Hypothetical protein	108.4	143.6	0	<−108.4	<−143.6
SITYG_17820		Putative phosphotransferase system sugar-specific EII component	117.7	311.7	35.6	−3.04	−10.50
SITYG_18350		Predicted transcriptional regulators	287.0	414.6	31.6	−8.35	−15.74
SITYG_19570		Hypothetical protein	145.5	192.8	0	<−145.5	<−192.8

[a]RPKM, number of reads per kilobase per million.

pGh9:IS*S1* (see Fig. S2 for details). In total, 1,152 Tn mutant clones were constructed (Fig. 5A and Table S2). Comprehensive detection of the Tn insertion sites was performed by targeted DNA sequencing (DNA-seq) for the insertion of a specific PCR amplicon (Table S1). Read mapping to the TYG1620 genome suggested that 629 ORFs were disrupted by the insertion of IS*S1 tnpA in vitro* as an original Tn mutant library (Fig. 5C); the genes disrupted by Tn insertion are also shown in the circular representation of the TYG1620 genome (slot 2 in Fig. 2B). The insertion appeared to be random, and no insertions were found in the ribosomal protein operon (SITYG_18370 to SITYG_18710, from Mb 1.84 to 1.85) or the region around *ori* (SITYG_19850 to SITYG_20130, from Mb 1.97 to 1.99) because these genes are essential for cell growth. Long repeats and the T7SS region showed significantly more Tn insertions than other regions.

All cultures of the 1,152 Tn mutants were pooled and injected subcutaneously into C57BL/6J mice. On day 7, apparent subcutaneous abscesses were observed in all tested mice ($n = 4$) (Fig. 5B). To identify the *in vivo*-attenuated Tn mutant clones, targeted DNA-seq was performed using purified DNA from the subcutaneous abscess, with the result being that insertions in 212 ORFs were not detected in any of the tested mice ($n = 4$). This result indicates that these 212 ORFs might play crucial roles in bacterial survival and abscess formation rather than in *in vitro* growth (Fig. 5D and Table S3). Furthermore, these ORFs generally contribute to mostly fundamental cell growth functions, suggesting that such Tn mutants might be complemented by alternative gene functions and metabolic pathways in *in vitro* growth, whereas the impaired Tn mutants might no longer be alive in a subcutaneous abscess *in vivo*.

To further clarify the virulence factors pivotal for *in vivo* survival, the comparative transcriptome data described above were analyzed with the Tn mutant results. As a result, 37 ORFs showed significantly increased or decreased levels of gene expression in abscesses and were identified in all 212 *in vivo*-attenuated mutants (Table 4). These 37 ORFs might contribute to *in vivo* survival, including nucleotide biosynthesis (*purH*, *purE*, SITYG_15350), cell division (SITYG_19460), sugar metabolism (SITYG_07100, SITYG_07110, *glgA*, SITYG_13990), and competency (SITYG_10670, *comYB*, SITYG_10880). In particular, increased competence might lead to considerable genome evo-

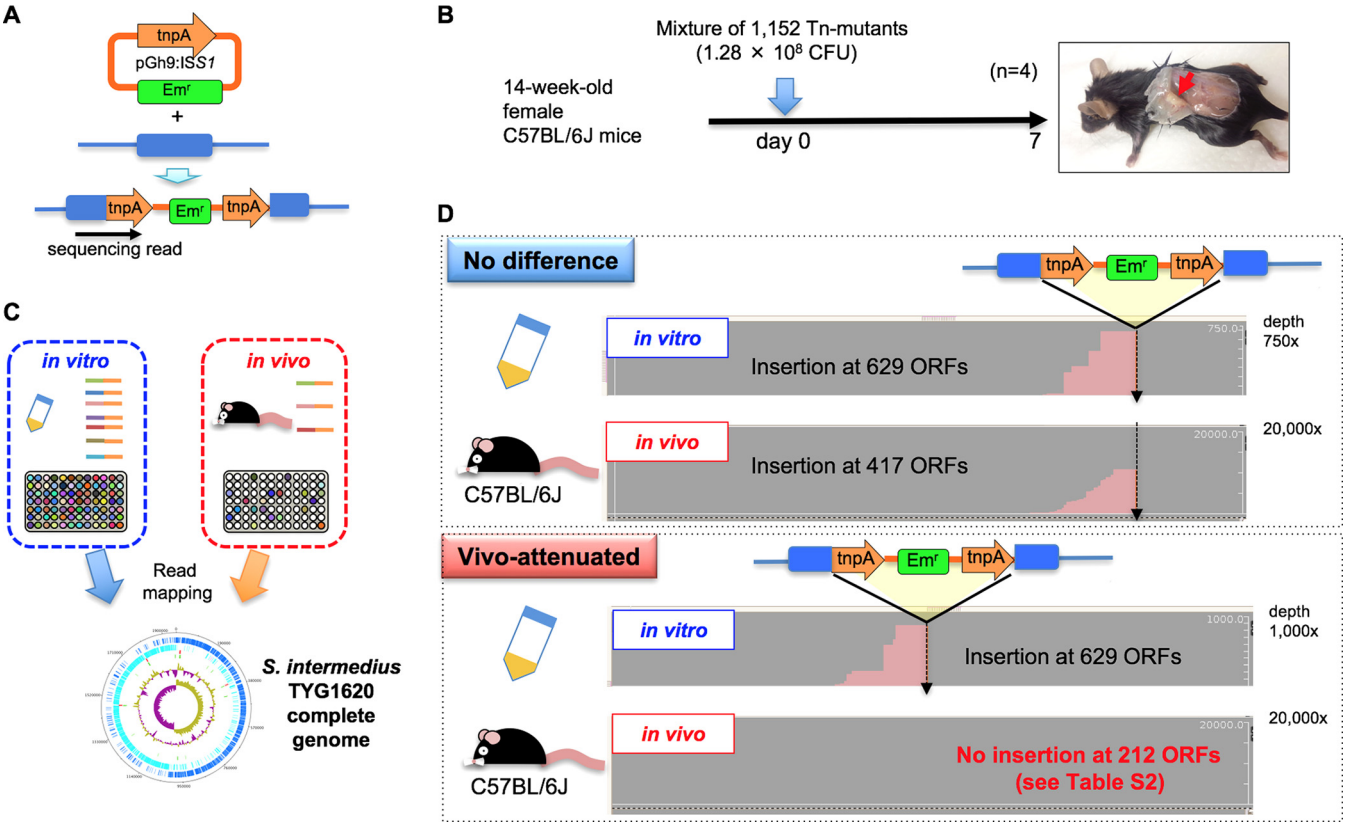

FIG 5 Identification of *in vivo*-attenuated Tn mutants. See Fig. S2 in the supplemental material for the construction of the random Tn mutant library. (A) Schematic representation of Tn insertion in TYG1620 chromosome DNA and detection of the insertion by targeted DNA-seq. (B) A mixture of 1,152 Tn mutants was inoculated subcutaneously into C57BL/6J mice ($n = 4$). At day 7, apparent subcutaneous abscesses were observed in all tested mice ($n = 4$). (C) DNA was prepared *in vitro* from a mixture of 1,152 Tn mutants and *in vivo* from subcutaneous abscesses, followed by DNA-seq to detect the Tn insertion sites. (D) Identification of *in vivo*-attenuated Tn mutants. Sequencing reads corresponding to the *tnpA* sequence were mapped to the TYG1620 genome sequence. In total, 629 ORFs were inserted by *tnpA in vitro* in the original Tn mutant library, while the insertion of 417 ORFs was detected in any tested mice ($n = 4$), indicating that nondetectable Tn mutants with mutations in the remaining 212 ORFs could play crucial roles in growth *in vivo* rather than *in vitro*.

lution via horizontal gene transfer. Indeed, a recent report demonstrated that horizontal transfer of plasmid and bacteriophage genes is greatly facilitated *in vivo* in the colonization of a gnotobiotic piglet by *S. aureus* (31).

The newly identified T7SS could be associated with *in vivo* survival, because Tn mutants with mutations in nine ORFs (SITYG_01170, SITYG_01200, SITYG_01250, SITYG_01270, SITYG_01390, SITYG_01430, SITYG_01490, SITYG_01530, and SITYG_01560) were attenuated in the abscess model (Fig. 2C and Table S2). In addition, SITYG_01560 (a hypothetical protein) was upregulated in the abscess formation compared with its level of regulation during *in vitro* growth (Table 4). Intriguingly, SITYG_01560 was found to be one of the paralogs in the T7SS locus. In total, nine ORFs (SITYG_01390, SITYG_01430, SITYG_01450, SITYG_01470, SITYG_01490, SITYG_01510, SITYG_01530, SITYG_01540, and SITYG_01560) shared multiple α-helix structures and at least 55% amino acid sequence similarity, showing variability in the sequence at the N terminus and a conserved amino acid sequence at the C terminus (Fig. S3).

Among 19 CWAP ORFs (Fig. 3), 8 ORFs (SITYG_00870, SITYG_01080, SITYG_01860, SITYG_10910, SITYG_10920, SITYG_16380, SITYG_17130, and SITYG_18140) were detected as *in vivo*-attenuated candidates (Table S3), implying that they are nonessential for *in vitro* growth but are necessary for adaptation to the severe environmental conditions in the abscess. As described above, in the long repeat region in TYG1620, two CWAPs (the putative glucan-binding proteins [32] SITYG_00870 and SITYG_10910) were selected to be crucial for *in vivo*-specific gene expression, and their significantly increased levels of expression in the abscess (Table 4) implied that these two CWAPs

TABLE 4 Crucial ORF candidates and differential expression for *in vivo* survival

S. intermedius TYG1620 genome annotation					Fold expression by RNA-seq		
Locus tag identifier	Gene	Product	COG[b]	LPXTG motif[c]	Abscess/BHI, 6 h	Abscess/BHI, 24 h	No. of Tn insertion sites *in vitro*[a]
SITYG_00320	*purH*	Bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	F		11.6		1
SITYG_00400	*purE*	Phosphoribosylaminoimidazole carboxylase catalytic subunit PurE	F		−3.2	−13.1	1
SITYG_00620		Putative D12 class N6 adenine-specific DNA methyltransferase	L		6.0		5
SITYG_00740		Putative endonuclease			3.8		4
SITYG_00760		Hypothetical protein			4.0		8
SITYG_00870		Putative cell wall anchor protein		LPXTG	4.2		7
SITYG_00920		Putative membrane protein	U		7.2		3
SITYG_01560[d]		Hypothetical protein			4.9		3
SITYG_02210		Hypothetical protein	M		3.4		1
SITYG_02610		Hypothetical protein	GEPR		3.5		2
SITYG_04200		ABC transporter, ATP-binding protein	V		3.5		4
SITYG_04440		Hypothetical protein			5.8		1
SITYG_06470	*celR*	Putative transcriptional regulator	K		4.4		2
SITYG_07100		ABC sugar transporter permease protein	G		6.6		1
SITYG_07110		ABC sugar transporter membrane spanning permease	G		5.3		3
SITYG_08430		Hypothetical protein	S		3.8		3
SITYG_08540		Hypothetical protein	R		3.1		2
SITYG_08620		Hypothetical protein			3.8		3
SITYG_08760		Putative ABC transporter ATPase	R		7.0		2
SITYG_09320	*pmrB*	Putative transport protein	GEPR		3.3		3
SITYG_10560		Putative RNase	S		4.2		4
SITYG_10670		Competence protein	R		5.5		2
SITYG_10740		Hypothetical protein			3.3		2
SITYG_10880		Putative conjugal transfer protein TraG			3.5		7
SITYG_10910		Putative cell wall anchor protein		LPXTG	3.8		7
SITYG_10960		DNA topoisomerase	L		3.3		5
SITYG_10980		Hypothetical protein			7.3		8
SITYG_11020		Hypothetical protein			3.7		8
SITYG_12930		Putative ABC transporter, ATP-binding protein	V		5.7		2
SITYG_13570	*glgA*	Glycogen synthase	G		3.5		2
SITYG_13990		Beta-galactosidase	G		3.4	−3.1	7
SITYG_14150		Hypothetical protein			<−188.6	<−222.0	1
SITYG_15350		Putative uridine phosphorylase			3.2		2
SITYG_16240		Hypothetical protein			5.1		2
SITYG_16560		Transposase	L		4.9		1
SITYG_18050	*comYB*	Competence protein ComYB	NU		3.7	−3.1	1
SITYG_19460		Cell division protein FtsK	D		6.6		2

[a]Number of Tn insertion sites *in vitro* in *in vivo*-attenuated TYG1620 Tn mutants in murine model of subcutaneous infection.
[b]COG, cluster of orthologous groups.
[c]See Fig. 3.
[d]ORF located in T7SS (Fig. 2C).

might contribute to *S. intermedius* pathogenesis by binding glucan molecules either in the abscess or in the oral environment.

S. intermedius possesses a species-unique toxin, ILY (*ily*, SITYG_01880), that specifically lyses human erythrocytes (12). The level of expression of ILY was increased by 3.5-fold in the abscess compared with that in an early-log-phase BHI culture, but the Tn mutant was not attenuated. ILY might not be involved in abscess formation; such secreted toxins may be available from the extracellular environments of other *ily*-positive Tn mutants. Thus, the contribution of ILY should be investigated using a specific *ily*-disrupted (or *ilv*-negative) Tn mutant.

Conclusions. We determined the complete genome sequence of *S. intermedius* TYG1620 isolated from a human brain abscess. Comparative genome analysis revealed that TYG1620 possesses a noteworthy pathogenicity island, including a T7SS and a possible repertoire of CWAP virulence factors. Transcriptome analysis and a random Tn mutant attenuation experiment in a murine subcutaneous abscess model identified

substantial virulence factors, in addition to ILY, that are important for *S. intermedius* pathogenicity. Specifically, *in vivo*-regulated genes similar to PSMβ1, paralogs of the T7SS and CWAPs, were identified. This study focused only on subcutaneous abscess formation; thus, the mechanisms of *S. intermedius* bacteremia and the subsequent internalization in the brain through the blood-brain barrier must still be elucidated to describe the complete pathogenicity of *S. intermedius*.

MATERIALS AND METHODS

Ethics statement. The study protocol was approved by the institutional Medical Ethics Committee of the National Institute of Infectious Diseases in Japan (approval no. 642), and it was conducted according to the Declaration of Helsinki principles. Prior to molecular diagnosis for etiological pathogens, written informed consent was obtained from the parents of the patient with a brain abscess to isolate potential etiological agents. The protocols for all experiments involving mice were approved by the guidelines of the Institutional Animal Care and Use Committee of the National Institute of Infectious Diseases, Japan (approval no. 115041), and the study was conducted according to Fundamental Guidelines for Proper Conduct of Animal Experiment and Related Activities in Research Institutions under the jurisdiction of the Ministry of Health, Labor and Welfare of Japan.

Bacterial strain. *S. intermedius* TYG1620 was isolated from a brain abscess in an infant. The strain was cultivated in BHI broth (Becton Dickinson, NJ, USA) or chocolate agar (Becton Dickinson, NJ, USA) under anaerobic conditions at 37°C.

PFGE. A Pulsed-field gel electrophoresis (PFGE) plug was prepared using a contour-clamped homogeneous electric field bacterial genomic DNA plug kit (Bio-Rad, CA, USA), replacing lysozyme with achromopeptidase (Wako, Osaka, Japan) for bacterial lysis. The plug was treated with ~40 units of the restriction enzyme AscI (New England Biolabs, USA), followed by PFGE (1% agarose gel, 0.5× TBE [Tris-borate-EDTA], 6 V/cm, 2.2- to 65.0-s pulse time, 120° angle, 20-h run time).

Whole-genome sequence analysis. Genomic DNA of *S. intermedius* TYG1620 was purified as follows: the bacterial cells were lysed with achromopeptidase (Wako), followed by phenol-chloroform extraction and further purification with a Qiagen DNA purification kit (Qiagen, Germany). Whole-genome sequencing of TYG1620 was performed using an Illumina GAIIx platform (paired-end 81-mer, 6,314,564 total paired-end reads). The reads were assembled using CLC Genomics Workbench (v7.5) software (Qiagen), followed by gap closing of specific PCR products and the DNA fragments obtained by PFGE using an Illumina MiSeq platform (paired-end 300-mer). The respective gap sequences were determined by the *de novo* assembly of a partial gap region using Platanus software (parameter, c-35) (17) and the PRICE assembler (parameter identity of 100%, minimum percent identity [mpi] of 99%, and target of 97%) (18).

Optical mapping of TYG1620 genome DNA was performed using the Argus system (OpGen, MD, USA) according to the manufacturer's protocol. Briefly, a large, genome-length DNA molecule of TYG1620 was prepared with an Argus HMW (high molecular weight) DNA isolation kit (Argus), and the prepared genome DNA was analyzed using a high-density MapCard kit (Argus) with the restriction enzyme NcoI to observe NcoI-specific digested sites on the card. The NcoI-digested DNA was detected as a barcode, followed by *de novo* assembly and visualization of the entire genome map with the MapSolver program (Argus).

Annotation was performed in Rapid Annotation using the Subsystems Technology (RAST; v2.0) (19), InterPro (v49.0) (20), and NCBI BLASTp/BLASTx software. SITYG_16380 has 6 tandem repeats of a 321-bp core unit showing a collagen-binding B motif. Such long tandem repeats were highly problematic in closing the gap between scaffolds (see Fig. S1 in the supplemental material). A 2.1-kb PCR product for the gap was obtained, followed by cloning of the PCR product into a pUC19 vector via in-fusion PCR cloning (Fig. S1).

Murine subcutaneous abscess model. Female C57BL/6J mice were obtained from Charles River Laboratories International, Inc., and were maintained under specific-pathogen-free conditions. *S. intermedius* TYG1620 was grown in BHI broth (BD Biosciences) under anaerobic conditions at 37°C for 18 h, followed by washing of the cells with phosphate-buffered saline (PBS). Two hundred microliters of the cell suspension (2×10^7 CFU) was inoculated subcutaneously into 8-week-old C57BL/6J mice anesthetized with isoflurane (Mylan, Tokyo, Japan).

RNA extraction. Subcutaneous abscesses were excised from the mice on day 6 after the *S. intermedius* TYG1620 inoculation and were kept at −80°C until use. Bacterial cells were recovered from the abscesses by the following procedure. A whole abscess (~50 mg) was homogenized in a 10× volume of the DNA/RNA Shield reagent (Zymo Research, Irvine, CA), followed by centrifugation at 15,000 rpm for 2 min to recover the *S. intermedius* cells as a pellet. The homogenization and cell recovery by centrifugation were repeated 3 additional times. The residual tissue, including *S. intermedius* cells, was washed twice with 10 mM Tris–10 mM EDTA (TE_{10} buffer), followed by resuspension in 100 μl of TE_{10} buffer including purified achromopeptidase (Wako) at 100 μg/ml, and the suspension was incubated at 37°C for 30 min. The cell lysate was subjected to RNA purification using a RecoverAll total nucleic acid isolation kit (Life Technologies) according to the manufacturer's instructions, with the modification that beating with 0.1-mm-diameter glass bead for 5 min was used to complete cell lysis. For extraction of RNA from *in vitro*-cultured *S. intermedius* TYG1620 in BHI broth, bacterial cells were harvested in the early log phase (i.e., after 6 h of cultivation) or stationary phase (i.e., after 24 h of cultivation). Each culture was collected and treated with the DNA/RNA Shield reagent for 5 min, after which the procedure described above was implemented.

RNA-seq analysis. Transcriptome sequencing (RNA-seq) libraries were prepared from approximately 30 ng of total RNA using a ScriptSeq (v2) RNA-seq library preparation kit (Epicentre Biotechnologies) according to the manufacturer's instructions. The RNA-seq libraries were sequenced as single-end 151-mers on a MiSeq sequencer using a MiSeq reagent kit (v3; Illumina, San Diego, CA) (Table S1). Transcriptome analysis was performed using CLC Genomics Workbench (v7.5) software (Qiagen K.K.). The significant ORFs were considered to be those with a false discovery rate (FDR)-normalized *P* value of less than 0.05. All RNA-seq raw data are available in Table S2.

Structure analysis. Secondary structure prediction was performed using the JPred (v4) server (http://www.compbio.dundee.ac.uk/jpred4/index_up.html) (33). Helical wheel projection was performed at http://rzlab.ucr.edu/scripts/wheel/wheel.cgi.

Construction of random Tn mutants. Random Tn insertion was performed using the plasmid pGh9:IS*S1* (GenBank accession number EU223008.1) carrying the insertion sequence IS*S1*, which facilitates random insertion. The plasmid can be depleted from the cells by cultivation at 38°C because of the presence of a temperature-sensitive replicon (34) (Fig. S2). Briefly, TYG1620 was grown in BHI broth under anaerobic conditions to mid-log phase, and the cells were washed with 0.5 M sucrose twice and 0.5 M sucrose–15% glycerol once to prepare electrocompetent cells. One hundred microliters of the prepared electrocompetent cells (1×10^7 CFU/ml) was mixed with 100 ng of pGh9:IS*S1* plasmid DNA, followed by electroporation with a MicroPulser apparatus (Bio-Rad) at 1.8 kV, 5.2 ms, and 200 Ω. The pulsed cells were immediately rescued with 1 ml of BHI broth containing 0.3 M sucrose and then incubated at 28°C for 2 h. The rescued cells were spread on BHI agar with 5 μg/ml erythromycin at 28°C for 24 h to obtain a stable clone harboring the pGh9:IS*S1* plasmid. Forty-one transformants were obtained and cultivated individually at 28°C for 8 h. To obtain the random Tn mutants, equal volumes of all transformant cultures were pooled, followed by incubation at a relatively high temperature (38°C) for 2 h to facilitate IS*S1* insertion into the chromosomal DNA and spread of the culture on BHI agar with 5 μg/ml erythromycin for incubation at 38°C for 62 h. In total, 1,152 colonies were picked and individually cultivated on 12 96-deep-well plates.

Identification of *in vivo*-attenuated Tn mutants in a murine subcutaneous abscess model. A Tn mutant library was prepared from a mixture of the 1,152 individually cultivated Tn mutants in BHI broth with 5 μg/ml erythromycin at 38°C under anaerobic conditions. The Tn mutant library mixture was washed twice with PBS and prepared at a cell density of 6.4×10^8 CFU/ml. Two hundred microliters of the cell suspension (1.3×10^8 CFU) was inoculated subcutaneously into 14-week-old C57BL/6J mice anesthetized with isoflurane (Mylan, Tokyo, Japan).

Subcutaneous abscesses were recovered at day 7. DNA was purified from the sample by the same procedure outlined in "RNA extraction" above without RNase treatment. A DNA-seq library for a targeted-only insertion site was prepared using a Nextera XT kit (Illumina). Briefly, the DNA was subjected to Tn*5* tagmentation with a Nextera XT kit, followed by neutralization according to the manufacturer's instruction. Generally, Nextera XT primer pairs (index 1 and index 2 primers) were used for the subsequent PCR enrichment and sample indexing, but here, the nucleotides of the Nextera XT index 1 primers (primers N701 to N712) were modified to anneal the IS*S1*-specific sequence as follows: P7 sequence (5′-CAAGCAGAAGACGGCATACGAGAT–3′)–index sequence (variable 8-mer)–the *tnpA* gene of the IS*S1*-specific sequence (5′-TCCTCGCTGTCATTTTTATTCAT-3′), corresponding to the sequence from nt 2272 to 2294 in pGh:IS*S1* (GenBank accession number EU223008.1). PCR enrichment was performed for 12 cycles according to the manufacturer's instruction. The targeted DNA-seq was performed by the single-end 330-mer MiSeq platform with indexing using the index 2 primers (primers N501 to N508) (Table S1).

The Tn insertion site was detected as follows: all sequencing reads were mapped to the sequence of *tnpA* (the transposase gene in pGh9:IS*S1*) by BWA-SW mapping (35) to collect the *tnpA* sequence-positive reads, the *tnpA* sequence was trimmed and subtracted according to an adapter-trimming procedure (36), and the resulting trimmed short reads were mapped to the TYG1620 chromosome DNA sequence (GenBank accession number AP014880) via BWA-SW mapping (35) to detect the single nucleotide at the IS*S1* insertion site. The insertion site was clearly shown as the high coverage peaks on the TYG1620 chromosome DNA sequence (Fig. 5D). All raw data for the Tn mutants are available in Table S2.

Accession number(s). The whole-genome sequence described here and its annotation are available in GenBank under accession number AP014880. The short-read sequences for RNA-seq and Tn mutant identification have been deposited in the DNA Data Bank of Japan (DDBJ accession number DRA005016).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/IAI.00886-16.

TEXT S1, PDF file, 0.3 MB.
TEXT S2, PDF file, 0.1 MB.
TEXT S3, PDF file, 0.2 MB.
TEXT S4, PDF file, 0.1 MB.
TEXT S5, PDF file, 0.5 MB.
TEXT S6, PDF file, 0.04 MB.

ACKNOWLEDGMENTS

The plasmid pGh9:IS*S1* was kindly provided by Toshifumi Tomoyasu (Tokushima University) under agreement with Alexandra Gruss (Unité Bactéries Lactiques et Pathogènes Opportunistes, France).

This study was supported by a Grant-in-Aid for Research on Emerging and Re-emerging Infectious Diseases, Labor and Welfare Programs from the Ministry of Health, Labor and Welfare of Japan (grant number H25-Shinko-Ippan-015) and the Research Program on Emerging and Re-emerging Infectious Diseases from the Japan Agency for Medical Research and Development (grant number 16fk0108119j0001).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Whiley RA, Fraser H, Hardie JM, Beighton D. 1990. Phenotypic differentiation of Streptococcus intermedius, Streptococcus constellatus, and Streptococcus anginosus strains within the "Streptococcus milleri group." J Clin Microbiol 28:1497–1501.
2. Whiley RA, Beighton D, Winstanley TG, Fraser HY, Hardie JM. 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.
3. Jacobs JA, Pietersen HG, Stobberingh EE, Soeters PB. 1995. Streptococcus anginosus, Streptococcus constellatus and Streptococcus intermedius. Clinical relevance, hemolytic and serologic characteristics. Am J Clin Pathol 104:547–553.
4. Junckerstorff RK, Robinson JO, Murray RJ. 2014. Invasive Streptococcus anginosus group infection—does the species predict the outcome? Int J Infect Dis 18:38–40. https://doi.org/10.1016/j.ijid.2013.09.003.
5. Bajpai A, Prasad KN, Mishra P, Singh AK, Gupta RK, Ojha BK. 2014. Distinct cytokine pattern in response to different bacterial pathogens in human brain abscess. J Neuroimmunol 273:96–102. https://doi.org/10.1016/j.jneuroim.2014.05.009.
6. Kommedal O, Wilhelmsen MT, Skrede S, Meisal R, Jakovljev A, Gaustad P, Hermansen NO, Vik-Mo E, Solheim O, Ambur OH, Saebo O, Hostmaelingen CT, Helland C. 2014. Massive parallel sequencing provides new perspectives on bacterial brain abscesses. J Clin Microbiol 52:1990–1997. https://doi.org/10.1128/JCM.00346-14.
7. Livingston LV, Perez-Colon E. 2014. Streptococcus intermedius bacteremia and liver abscess following a routine dental cleaning. Case Rep Infect Dis 2014:954046. https://doi.org/10.1155/2014/954046.
8. Heckmann JG, Pauli SU. 2015. Epidural abscess after dental extraction. Age Ageing 44:901. https://doi.org/10.1093/ageing/afv094.
9. Iovino F, Seinen J, Henriques-Normark B, van Dijl JM. 2016. How does Streptococcus pneumoniae invade the brain? Trends Microbiol 24:307–315. https://doi.org/10.1016/j.tim.2015.12.012.
10. Uchiyama S, Carlin AF, Khosravi A, Weiman S, Banerjee A, Quach D, Hightower G, Mitchell TJ, Doran KS, Nizet V. 2009. The surface-anchored NanA protein promotes pneumococcal brain endothelial cell invasion. J Exp Med 206:1845–1852. https://doi.org/10.1084/jem.20090386.
11. Yamaguchi M, Hirose Y, Nakata M, Uchiyama S, Yamaguchi Y, Goto K, Sumitomo T, Lewis AL, Kawabata S, Nizet V. 2016. Evolutionary inactivation of a sialidase in group B Streptococcus. Sci Rep 6:28852. https://doi.org/10.1038/srep28852.
12. Nagamune H, Ohnishi C, Katsuura A, Fushitani K, Whiley RA, Tsuji A, Matsuda Y. 1996. Intermedilysin, a novel cytotoxin specific for human cells secreted by Streptococcus intermedius UNS46 isolated from a human liver abscess. Infect Immun 64:3093–3100.
13. Giddings KS, Zhao J, Sims PJ, Tweten RK. 2004. Human CD59 is a receptor for the cholesterol-dependent cytolysin intermedilysin. Nat Struct Mol Biol 11:1173–1178. https://doi.org/10.1038/nsmb862.
14. Faden HS. 2016. Infections associated with Streptococcus intermedius in children. Pediatr Infect Dis J 35:1047–1048. https://doi.org/10.1097/INF.0000000000001227.
15. Planet PJ, Rampersaud R, Hymes SR, Whittier S, Della-Latta PA, Narechania A, Daugherty SC, Santana-Cruz I, Desalle R, Ravel J, Ratner AJ. 2013. Genome sequence of the human abscess isolate Streptococcus intermedius BA1. Genome Announc 1(1):e00117-12. https://doi.org/10.1128/genomeA.00117-12.
16. Olson AB, Kent H, Sibley CD, Grinwis ME, Mabon P, Ouellette C, Tyson S, Graham M, Tyler SD, Van Domselaar G, Surette MG, Corbett CR. 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. https://doi.org/10.1186/1471-2164-14-895.
17. Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama H, Kohara Y, Fujiyama A, Hayashi T, Itoh T. 2014. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Res 24:1384–1395. https://doi.org/10.1101/gr.170720.113.
18. Ruby JG, Bellare P, Derisi JL. 2013. PRICE: software for the targeted assembly of components of (meta) genomic sequence data. G3 (Bethesda) 3:865–880. https://doi.org/10.1534/g3.113.005967.
19. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206–D214. https://doi.org/10.1093/nar/gkt1226.
20. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador-Vegas A, Scheremetjew M, Rato C, Yong SY, Bateman A, Punta M, Attwood TK, Sigrist CJ, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P, Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, Wu CH, Orengo C, Sillitoe I, Mi H, Thomas PD, Finn RD. 2015. The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res 43:D213–D221. https://doi.org/10.1093/nar/gku1243.
21. Berthet FX, Rasmussen PB, Rosenkrands I, Andersen P, Gicquel B. 1998. A Mycobacterium tuberculosis operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). Microbiology 144(Pt 11):3195–3203.
22. Simeone R, Bottai D, Brosch R. 2009. ESX/type VII secretion systems and their role in host-pathogen interaction. Curr Opin Microbiol 12:4–10. https://doi.org/10.1016/j.mib.2008.11.003.
23. Pallen MJ. 2002. The ESAT-6/WXG100 superfamily—and a new Gram-positive secretion system? Trends Microbiol 10:209–212. https://doi.org/10.1016/S0966-842X(02)02345-4.
24. Abdallah AM, Gey van Pittius NC, Champion PA, Cox J, Luirink J, Vandenbroucke-Grauls CM, Appelmelk BJ, Bitter W. 2007. Type VII secretion—mycobacteria show the way. Nat Rev Microbiol 5:883–891. https://doi.org/10.1038/nrmicro1773.
25. Comfort D, Clubb RT. 2004. A comparative genome analysis identifies distinct sorting pathways in gram-positive bacteria. Infect Immun 72:2710–2722. https://doi.org/10.1128/IAI.72.5.2710-2722.2004.
26. Peschel A, Otto M. 2013. Phenol-soluble modulins and staphylococcal infection. Nat Rev Microbiol 11:667–673. https://doi.org/10.1038/nrmicro3110.
27. Cheung GY, Joo HS, Chatterjee SS, Otto M. 2014. Phenol-soluble modulins—critical determinants of staphylococcal virulence. FEMS Microbiol Rev 38:698–719. https://doi.org/10.1111/1574-6976.12057.
28. Qi R, Joo HS, Sharma-Kuinkel B, Berlon NR, Park L, Fu CL, Messina JA, Thaden JT, Yan Q, Ruffin F, Maskarinec S, Warren B, Chu VH, Fortes CQ, Giannitsioti E, Durante-Mangoni E, Kanafani ZA, Otto M, Fowler VG, Jr.

2016. Increased in vitro phenol-soluble modulin production is associated with soft tissue infection source in clinical isolates of methicillin-susceptible Staphylococcus aureus. J Infect 72:302–308. https://doi.org/10.1016/j.jinf.2015.11.002.
29. Syed AK, Reed TJ, Clark KL, Boles BR, Kahlenberg JM. 2015. Staphylococcus aureus phenol-soluble modulins stimulate the release of proinflammatory cytokines from keratinocytes and are required for induction of skin inflammation. Infect Immun 83:3428–3437. https://doi.org/10.1128/IAI.00401-15.
30. Berlon NR, Qi R, Sharma-Kuinkel BK, Joo HS, Park LP, George D, Thaden JT, Messina JA, Maskarinec SA, Mueller-Premru M, Athan E, Tattevin P, Pericas JM, Woods CW, Otto M, Fowler VG, Jr. 2015. Clinical MRSA isolates from skin and soft tissue infections show increased in vitro production of phenol soluble modulins. J Infect 71:447–457. https://doi.org/10.1016/j.jinf.2015.06.005.
31. McCarthy AJ, Loeffler A, Witney AA, Gould KA, Lloyd DH, Lindsay JA. 2014. Extensive horizontal gene transfer during Staphylococcus aureus co-colonization in vivo. Genome Biol Evol 6:2697–2708. https://doi.org/10.1093/gbe/evu214.
32. Sato Y, Yamamoto Y, Kizaki H. 1997. Cloning and sequence analysis of the gbpC gene encoding a novel glucan-binding protein of Streptococcus mutans. Infect Immun 65:668–675.
33. Drozdetskiy A, Cole C, Procter J, Barton GJ. 2015. JPred4: a protein secondary structure prediction server. Nucleic Acids Res 43:W389–W394. https://doi.org/10.1093/nar/gkv332.
34. Maguin E, Prevost H, Ehrlich SD, Gruss A. 1996. Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. J Bacteriol 178:931–935. https://doi.org/10.1128/jb.178.3.931-935.1996.
35. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595. https://doi.org/10.1093/bioinformatics/btp698.
36. Jiang H, Lei R, Ding SW, Zhu S. 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15:182. https://doi.org/10.1186/1471-2105-15-182.