Received: 17 July 2020 | Revised: 20 September 2020 | Accepted: 16 October 2020

DOI: 10.1111/mmi.14629

RESEARCH ARTICLE

WILEY

Genome-wide identification of *Streptococcus sanguinis* fitness genes in human serum and discovery of potential selective drug targets

Bin Zhu [1] | Shannon P. Green [1,2] | Xiuchun Ge[1] | Tanya Puccio [1] | Haider Nadhem[1] | Henry Ge[1] | Liang Bao[1] | Todd Kitten [1,2] | Ping Xu [1,2,3]

[1]Philips Institute for Oral Health Research, Virginia Commonwealth University, Richmond, VA, USA

[2]Department of Microbiology and Immunology, Virginia Commonwealth University, Richmond, VA, USA

[3]Center for Biological Data Science, Virginia Commonwealth University, Richmond, VA, USA

Correspondence
Ping Xu, Philips Institute for Oral Health Research, Virginia Commonwealth University, Richmond, VA, 23298, USA.
Email: pxu@vcu.edu

Funding information
CCTR Endowment Fund, Grant/Award Number: 2-94579; National Institute of Allergy and Infectious Diseases, Grant/Award Number: R01AI114926; National Institute of Dental and Craniofacial Research, Grant/Award Number: R01DE023078

Abstract

Streptococcus sanguinis is a primary colonizer of teeth and is associated with oral health. When it enters the bloodstream, however, this bacterium may cause the serious illness infective endocarditis. The genes required for survival and proliferation in blood have not been identified. The products of these genes could provide a rich source of targets for endocarditis-specific antibiotics possessing greater efficacy for endocarditis, and also little or no activity against those bacteria that remain in the mouth. We previously created a comprehensive library of *S. sanguinis* mutants lacking every nonessential gene. We have now screened each member of this library for growth in human serum and discovered 178 mutants with significant abundance changes. The main biological functions disrupted in these mutants, including purine metabolism, were highlighted via network analysis. The components of an ECF-family transporter were required for growth in serum and were shown for the first time in any bacterium to be essential for endocarditis virulence. We also identified two mutants whose growth was reduced in serum but not in saliva. This strategy promises to enable selective targeting of bacteria based on their location in the body, in this instance, treating or preventing endocarditis while leaving the oral microbiome intact.

KEYWORDS
infective endocarditis, purine metabolism, sequencing, *Streptococcus sanguinis*, virulence

1 | INTRODUCTION

Streptococcus sanguinis is a gram-positive facultative anaerobe and a primary colonizer of the oral cavity (Zhu et al., 2018). Epidemiological studies have shown that *S. sanguinis* has significantly increased abundance in healthy versus diseased subgingiva as well as higher abundance on sound enamel versus carious lesions (Griffen et al., 2012; Stingu et al., 2008; Zhu et al., 2018). Additionally, in vitro experiments have shown that *S. sanguinis* can compete against oral pathogens such as *Streptococcus mutans* and *Porphyromonas gingivalis* via hydrogen peroxide production (Chen et al., 2012; Kreth et al., 2005; Zhu et al., 2018, 2019). These studies suggest that *S. sanguinis* contributes to the homeostasis of the oral cavity. It is also known, however, that oral surgery, mastication, or routine hygiene practices

can produce lesions in the oral mucosa through which *S. sanguinis* and other oral streptococci can enter the bloodstream (Wilson et al., 2007; Wray et al., 2008). If this happens in a person who has certain preexisting cardiac conditions, these bacteria can colonize the damaged heart valves or other endocardial surfaces, causing a disease called infective endocarditis (IE) (Baddour et al., 2015).

Antibiotics are often used prophylactically to prevent IE (European Society of Cardiology, 2015; Wilson et al., 2007) and therapeutically to treat it (Baddour et al., 2015; Wilson et al., 2007). In the former case, the primary indication is patients with high-risk cardiac conditions (Thornhill et al., 2018) who are preparing to undergo an invasive dental procedure. In the latter case, treatment can last 4 to 6 weeks (Rasmussen et al., 2017), a duration that would be expected to contribute to both the development of antibiotic resistance and to the disruption of the oral microbiome (Shaw et al., 2019). Moreover, efficacy is suboptimal, with valve surgery required in addition to antibiotics in approximately 50% of cases. Even with all available treatments, 1-year readmission and mortality rates of 65% and 33%, respectively, have been reported (Rasmussen et al., 2017).

We previously proposed targeted antimicrobial therapy based on the different essential gene targets present in beneficial versus pathogenic bacterial species in the microbiome (Stone et al., 2015; Stone & Xu, 2017). Here, we propose a second method to selectively target pathogens—in this instance, based on the different genetic requirements of a single bacterial species when present at a disease site versus its usual anatomical location. As applied to IE caused by oral streptococci, we propose to take advantage of potential differences in available nutrients in the oral cavity compared to infected heart valves. If an essential nutrient is present at a lower concentration in the bloodstream than in the oral cavity, genes required for its uptake should be much more important for bacteria on blood-bathed endocardial surfaces than in the oral cavity. A drug that specifically inhibits *S. sanguinis* in the bloodstream may prove more efficacious than conventional antibiotics for IE prevention or treatment while leaving the oral microbiome intact. In addition, several studies have identified *S. sanguinis* strains from IE cases that are resistant to penicillin (Pericàs et al., 2019), linezolid (Mendes et al., 2013), and vancomycin (Safdar & Rolston, 2006). Thus, finding new drug targets for the selective inhibition of *S. sanguinis* in the bloodstream could lead to fewer treatment failures.

In the process of identifying the essential genes in *S. sanguinis* SK36, we constructed a comprehensive library of nonessential genes containing 2046 single-gene mutants (Xu et al., 2011). In this study, we tested the growth of these mutants in human serum (HS) and brain heart infusion (BHI) broth using a sequencing protocol similar to transposon sequencing (Tn-seq) that we refer to as ORF-seq. The network of genes whose corresponding mutants had significantly reduced growth in HS was constructed by predicting the protein–protein interactions in the STRING database (Szklarczyk et al., 2019). Biological functions and metabolic pathways related to the growth of *S. sanguinis* in HS were identified. Several genes found to be important for growth in HS were also shown to be important for IE virulence, either previously or in this study. Additionally, two mutants were discovered with defective growth in serum, but not in saliva. These mutants are potential drug targets for selective inhibition of *S. sanguinis* growth in the bloodstream.

2 | RESULTS

2.1 | Identification of fitness genes of *S. sanguinis* in HS using ORF-seq

Separate cultures of 2046 *S. sanguinis* mutants, each of which contains a deletion in a single nonessential gene, were pooled together, harvested, and inoculated into HS, which was incubated at 37°C under microaerobic conditions (6% O_2) to mimic the nutrient conditions in the bloodstream. The pooled bacterial cells were collected after 24 hr incubation (T_{24}) and compared with the bacterial cells in the original library (T_0). The abundance change of mutants was determined by a procedure with similarities to Tn-seq that we have named ORF-seq. (See Materials and Methods for details.) A total of 178 mutants were identified with significant abundance changes ($\log_2$ fold change ≥ 0.585 or ≤ −0.585 and p value ≤ .05) in HS including 138 mutants whose abundance was decreased (SI Spreadsheet 1).

2.2 | Network analysis of *S. sanguinis* fitness genes in HS

We used STRING, a comprehensive database for the prediction of protein–protein interactions (Szklarczyk et al., 2017), to investigate the linkage between the 178 proteins whose corresponding mutants exhibited a significant change in abundance in HS. The linkage information was put into Gephi for visualization and analysis of the network (Bastian et al., 2009). There are 149 proteins that were annotated in STRING and present in the network (Figure 1 and SI Spreadsheet 2). Others are uncharacterized proteins or proteins without information in the STRING database, whose functions cannot be predicted. The main network contains 135 proteins and contains several small, independent proteins groups.

Most of the mutants with altered abundance had reduced rather than increased abundance after 24 hr of incubation in HS (Figure 1). Four of these mutants correspond to genes that we identified previously as encoding virulence factors for IE in a rabbit model: *ssaB*, *purB*, *lgt*, and *ccpA* (Bai et al., 2019; Das et al., 2009; Paik et al., 2005). This is not surprising because a mutant that cannot grow in serum is unlikely to grow on an infected heart valve with blood as its primary source of nutrition.

The 10 mutants with the largest decrease in abundance were deleted for *purA*, *rpiA*, *clpP*, *priA*, *hom*, *purB*, *guaA*, *nadD*, *guaB*, or *rpoE* (Figure 1 and SI Spreadsheet 2). Four of the mutants (*purA*, *purB*, *guaA*, and *guaB*) belong to the purine metabolism category in the Clusters of Orthologous Groups (COG) database (Tatusov et al., 2000), indicating that purine metabolism is required for normal growth of *S. sanguinis* in

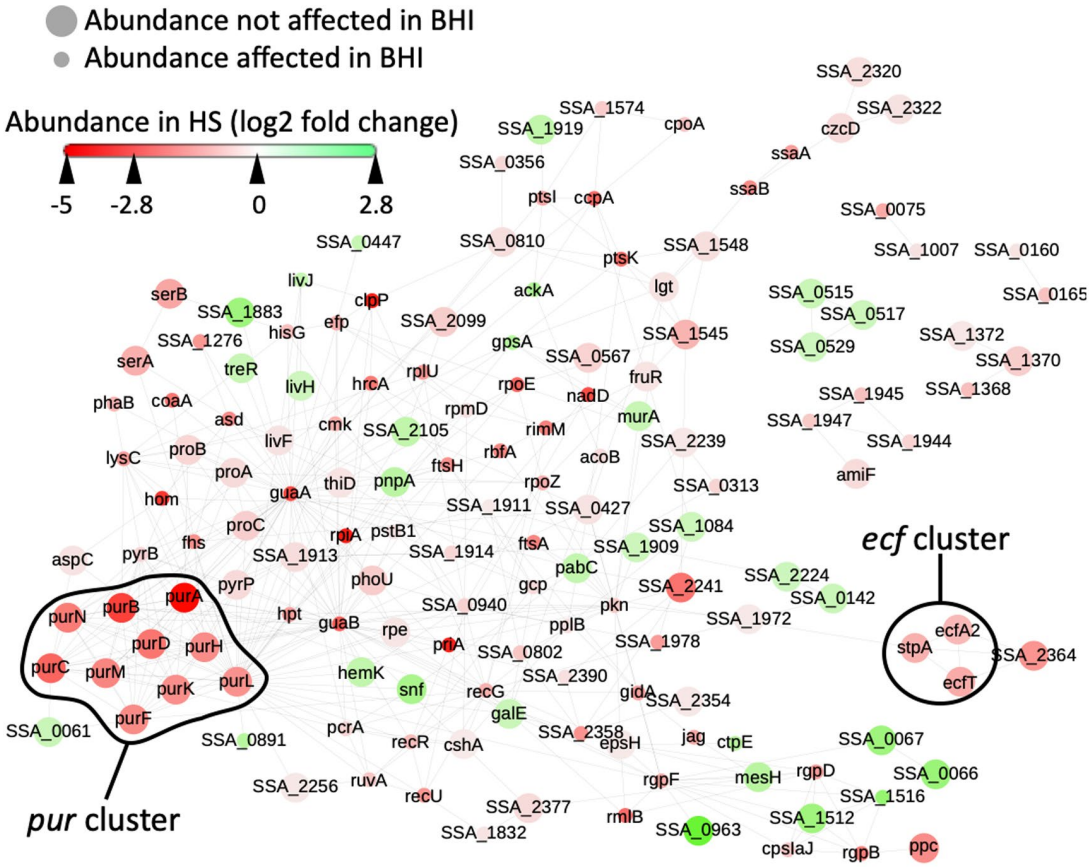

FIGURE 1 HS-related protein network and abundance change of mutants in HS. The mutant library was incubated in HS for 24 hr and the abundance of mutants was measured by ORF-seq. Mutants that had significant changes in abundance after 24 hr of incubation in HS that were also annotated in STRING are displayed. The abundance change is indicated by coloration as depicted in the color scale. The node sizes are larger for mutants whose abundance was not affected in BHI. Significance was analyzed by Rockhopper software and determined by a *p* value adjusted for a false discovery rate of 1% [Colour figure can be viewed at wileyonlinelibrary.com]

HS. Similar findings have been obtained previously with other bacterial species (Le Breton et al., 2013; Samant et al., 2008; Zhang et al., 2017). Details of these four genes will be discussed below.

Functional enrichment analysis was performed using the DAVID database (Huang et al., 2009) for the 138 genes whose mutants exhibited significant decreases in abundance in HS. Three KEGG pathways—ABC transporters, purine metabolism, and arginine and proline metabolism—were found to be significantly enriched (Figure 2a). The biological functions of the proteins in the network were annotated using the COG database and sorted by the number of proteins involved in these functions (Tatusov et al., 2000). The top eight functions are depicted in Figure 2b. The pathway enrichment results were similar to those obtained from the COG annotation.

The size of nodes in the network is positively related to the betweenness centrality of proteins (Figure 2b). There exists at least one shortest path between two nodes in a network. The betweenness centrality describes the number of different shortest paths passing through a node in a network (Freeman, 1977). A protein with a larger betweenness centrality is more central in a network. In our network, GuaA, GuaB, and PurL are the three largest nodes (Figure 2b). Their biological functions are related to the purine metabolism pathway.

Other hub proteins (betweenness centrality ≥ 400) are also apparent, such as Pkn, Gcp, Lgt, and RgpF (Figure 2b). The direct link from these four hub proteins to other proteins in the network is shown in Figure S1. In our ORF-seq results, all four of these hub gene's mutants had reduced growth in HS (Figure 1). However, the growth reduction exhibited by these four mutants was less pronounced than that of other mutants with less betweenness centrality. For example, Δ*ssaB* and Δ*stpA* both had larger decreases in abundance in HS than Δ*lgt*, (Figure 1). These data suggest that, although the four hub proteins interacted with other proteins or might modulate other proteins at the posttranslational level (Bugrysheva et al., 2011), proteins directly related to nutrient uptake and synthesis such as PurA, PurB, SsaB, and StpA seemed to be more important (Figure 1).

2.3 | The impact of purine metabolism on *S. sanguinis* growth in HS

We next repeated the ORF-seq experiment as before, except in BHI. Because the library was constructed in BHI-grown cells, we knew

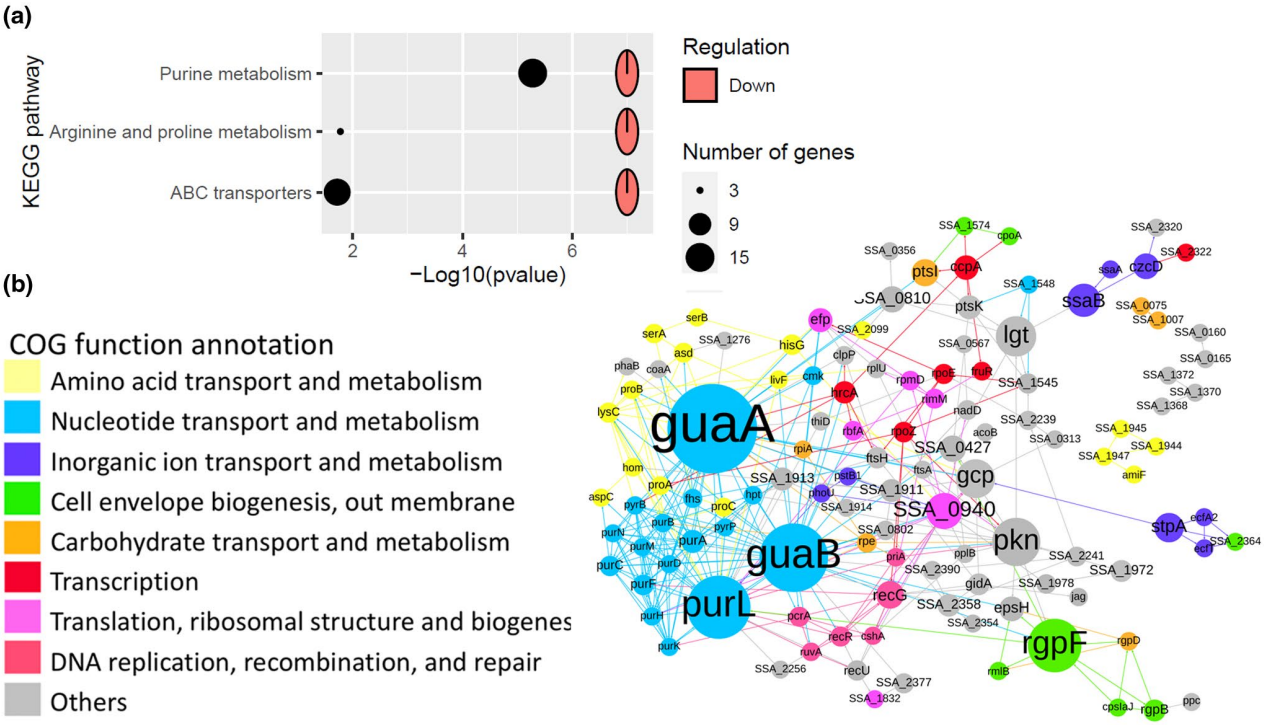

FIGURE 2 Functional analyses of fitness proteins in HS. (a) All 138 proteins whose corresponding mutants had significantly decreased abundance in HS were entered into the DAVID database for enrichment analysis. The significantly enriched KEGG pathways (p value ≤ .05) implicated in the growth of *S. sanguinis* in HS are shown. (b) The 138 fitness proteins were classified by COG function annotation and are colored accordingly. The size of nodes in the network is positively related to the betweenness centrality [Colour figure can be viewed at wileyonlinelibrary.com]

that all the mutants should be capable of growth, although not at the same rate. We performed the same enrichment analysis described above with genes whose mutants had decreased abundance after 24 hr of incubation in BHI. Interestingly, homologous recombination was the only pathway that was significantly enriched in BHI (SI Spreadsheet 3). By comparing the results of the two ORF-seq studies, we identified 88 mutants whose relative abundance was altered with a $\log_2$ value ≥|0.585| in HS but not in BHI. These 88 mutants are depicted by a bigger node size in Figure 1, in blue in Figure S2 and their locus tags are highlighted in yellow in SI Spreadsheet 1. Although any gene required for growth in HS is of interest, we reasoned that those that were not needed for growth in the rich medium BHI might be more likely to be HS-specific. We noted that within this group, 10 of the 12 mutants with the greatest defect in HS abundance belonged to the purine metabolic pathway. As depicted in the KEGG database, purine metabolism is one of two components of the nucleotide transport and metabolism pathway (Figure 3a) (Kanehisa et al., 2017). The genes on the left side of phosphoribosyl pyrophosphate (PRPP) are responsible for purine salvage and synthesis while the genes on the right side are responsible for pyrimidines. These pathways have been well characterized (Kilstrup et al., 2005; Leonard et al., 2018). Many of the genes in both pathways were found to be essential, as indicated in Figure 3a by the blue coloring. The corresponding mutants could not be constructed and their phenotypes could therefore not be assessed (Xu et al., 2011).

The deletion of *pur* genes whose products catalyze de novo purine synthesis from PRPP/L-glutamine to AMP/GMP reduced the growth of *S. sanguinis* in HS, suggesting that de novo synthesis of AMP and GMP was important in HS (Figure 3a). In contrast, most of the genes required solely for purine salvage had little impact on growth, including *punA*, *deoD*, *apt*, and *xpt* (Figure 3a). This could either be because salvage is only important for serum growth when de novo synthesis is blocked by mutation or because the gene products, which are nucleotide phosphorylases (PunA, DeoD) or phosphoribosyltransferases (Apt, Xpt), are at least partially redundant. The mutation of *hpt*, which also encodes a phosphoribosyltransferase (hypoxanthine → IMP), did reduce abundance in HS significantly ($\log_2$ fold change = −2.15), suggesting that purine salvage contributed to the growth of *S. sanguinis* in HS (Figure 3a). Nevertheless, the deletion of *hpt* had less impact than deletion of genes such as *purH*, *purC*, and *purE*, indicating that de novo synthesis was more important than purine salvage from inosine for growth in HS (Figure 3a and SI Spreadsheet 1). Finally, it should also be noted that the four mutants (*purA*, *purB*, *guaA*, and *guaB*) that displayed the largest reductions in abundance in HS out of all the purine mutants tested are each blocked in the conversion of inosine to either AMP or GMP, which are reactions shared by the de novo and salvage pathways. The combined results suggest that salvage plays a role in purine replenishment in HS, but that de novo synthesis makes a greater contribution.

Given that salvage requires less energy than synthesis, we predicted that its limited contribution to growth in HS likely results from purine nucleoside concentrations being too low to support normal growth of *S. sanguinis*. To test this hypothesis, we consulted the Human

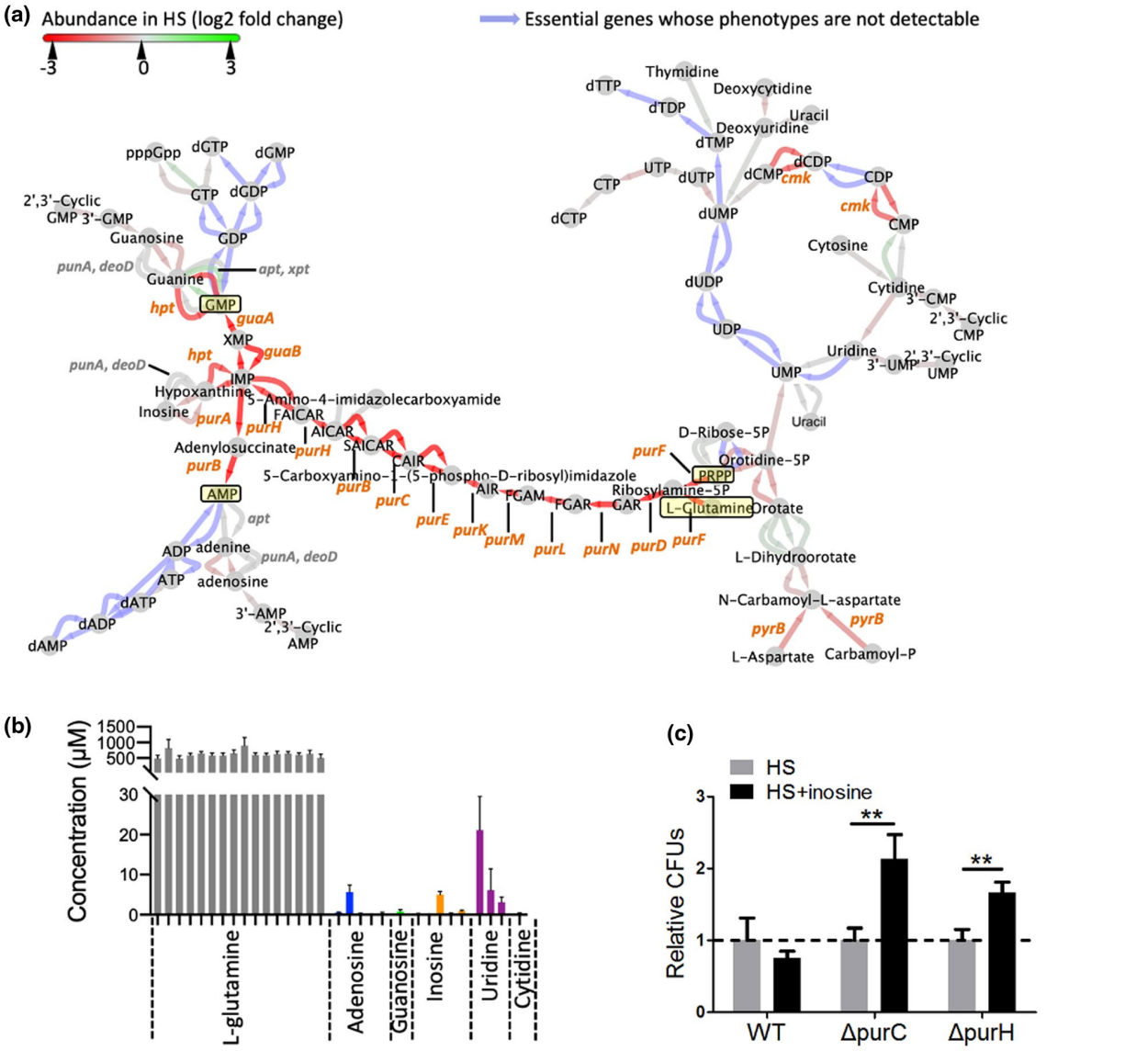

FIGURE 3 The effect of nucleotide metabolism on the growth of *S. sanguinis* in HS. (a) The KEGG nucleotide transport and metabolism pathway is shown. The abundance changes of the respective mutants in HS in ORF-seq are indicated by color according to the scale shown. (b) Metabolite concentrations in blood in the form of mean ± standard deviation were downloaded from the Human Metabolome Database (HMDB) and are depicted here. (C) The growth of Δ*purC* and Δ*purH* was measured after 24 hr of incubation with or without supplementation with 0.15 mg/ml inosine. The CFUs of each strain in HS+ inosine are expressed relative to that of the same strain in HS. Means and standard deviations from triplicate experiments are shown. ***p* value ≤ .01, Student's *t* test [Colour figure can be viewed at wileyonlinelibrary.com]

Metabolome Database (HMDB) (Wishart et al., 2018), which showed that the concentration of L-glutamine in blood was more than 500 μM, whereas the concentrations of adenosine, guanosine, and inosine were less than 1 μM for most of the results reported (Figure 3b), in agreement with our expectations. We then added 0.15 mg/ml (560 μM) inosine to HS and the growth of *S. sanguinis* wild type (WT), the Δ*purC* mutant, and the Δ*purH* mutant was measured after 24 hr of incubation in individually inoculated tubes. The addition of inosine did not significantly impact the colony-forming unit (CFU) numbers for WT (Figure 3c), but did increase the CFUs of Δ*purC* and Δ*purH*, which was also consistent with our hypothesis (Figure 3c). In other words, only when the pathway from L-glutamine/PRPP to AMP/GMP was blocked, would the production of AMP/GMP from inosine salvage noticeably improve growth. When cells were grown in BHI, the precursors for de novo synthesis and nucleosides for salvage were sufficient to support normal growth of the *pur* mutants (Δ*purA*, Δ*purB*, Δ*purC*, Δ*purD*, Δ*purE*, Δ*purF*, Δ*purH*, Δ*purK*, Δ*purL*, Δ*purM*, and Δ*purN*) (Figure S2 and SI Spreadsheet 1). Interestingly, however, the *guaA* and *guaB* mutants grew poorly in BHI (although better than in serum), suggesting that guanosine levels in BHI are not sufficient to support normal growth in the absence of conversion of inosine to GMP (SI Spreadsheet 1).

Compared to purine metabolism, genes involved in pyrimidine metabolism had less impact on *S. sanguinis* growth in HS (Figure 3a). The genes responsible for biosynthesis from PRPP/L-aspartate/L-glutamine to UMP had little to no effect on the abundance of *S. sanguinis* in HS (Figure 3a). One possibility was that there was enough

UMP from another pathway to compensate for loss of de novo biosynthesis. The HMDB indicated concentrations of uridine greater than 3 μM for all three reported results, whereas most of the reported concentrations of adenosine, guanosine, and inosine were lower than 1 μM, which may explain why de novo biosynthesis of UMP was less important than AMP/GMP (Figure 3b).

2.4 | Transporters important for the growth of *S. sanguinis* in HS

In gram-positive bacteria, transport proteins localize to the sole cell membrane. Therefore, many of the component proteins are surface-accessible making them appealing drug targets. Using the KEGG pathway annotation tool (Kanehisa et al., 2017), we identified four types of ABC transporters whose mutants had reduced abundance in HS. These included six putative metal transport genes (*ecfA2*, *ecfT*, *stpA*, *ssaA*, *ssaB*, and *czcD*), four peptide and nickel transport genes (*oppF*, *oppD*, *oppB*, and *amiF*), three phosphate and amino acid transport genes (*pstB1*, *livF*, and SSA_2099), and one sugar transporter (SSA_1007) (SI Spreadsheet 1). Deletion of the metal transport genes caused the greatest reduction in *S. sanguinis* growth in HS (SI Spreadsheet 1). We examined the most promising mutants by individual inoculation into tubes containing HS or saliva, followed by dilution plating and colony counts.

The *czcD* gene has been studied in *S. pneumoniae*, where it is responsible for transporting Zn out of cells (Martin & Giedroc, 2016). The ortholog in *S. sanguinis* (SSA_2321) is a closely related to the pneumococcal protein (SPD_1638), and the two are reciprocal best hits by BlastP analysis. The *czcD* deletion mutant is sensitive to Zn toxicity (Martin & Giedroc, 2016). Metabolomics data from HMDB showed that the concentration of Zn in blood is higher than that in saliva (Figure 4a), which might explain why Δ*czcD* had poor growth in HS. In individual inoculation assays, growth of Δ*czcD* was normal in saliva but growth was much lower in HS after 6 hr of incubation, which was consistent with this hypothesis (Figure 4b). However, the final CFUs of Δ*czcD* cultures were similar to WT after 24 hr, suggesting that the deletion of *czcD* reduced the growth rate rather than achievable density (Figure 4b).

The Mn transporter mutant Δ*ssaA* was similar to Δ*ssaB*, with both showing defective growth in HS as assessed by ORF-Seq (SI Spreadsheet 1). Although the Δ*ssaC* mutant displayed a decreased but insignificant abundance change by ORF-seq (SI Spreadsheet 1), the reduction in growth in HS was significant when the *ssaC* mutant was assessed in an individual inoculation assay (Figure 4d). Metabolomics data from HMDB showed a seemingly higher concentration of Mn in saliva than in blood or serum, although the difference was not significant (Figure 4c). Individual inoculation growth studies performed in HS and saliva revealed that the three *ssa* gene deletion mutants demonstrated reduced growth in both HS and saliva (Figure 4d).

ECF genes belong to the energy-coupling factor (ECF) family of ABC transporters (Rempel et al., 2019). Deletion of *ecfA2*, *ecfT*, or *stpA* attenuated the growth of *S. sanguinis* in both HS and saliva by individual inoculation (Figure 4e). However, the growth defects of these three mutants were more severe in HS than in saliva (Figure 4e); the average decrease of these three mutants in HS was ~2.4 and ~6 times larger than that in saliva at 6 and 24 hr, respectively. These three mutants will be examined further as described below.

2.5 | Mutants with defective growth in HS but normal growth in saliva

In addition to the transport mutants described above, we selected a small number of mutants whose growth was reduced in HS but not in BHI and compared their growth in HS versus saliva by individual inoculation. Two mutants, ΔSSA_0048 and ΔSSA_2241, were identified that were recovered from HS at significantly lower numbers than WT at 24 hr, but demonstrated normal growth in saliva (Figure 4f). Both of the corresponding genes are annotated as encoding hypothetical proteins (Xu et al., 2007). Further studies will be required to determine their functions.

2.6 | Characterization of ECF transporters and contribution to endocarditis virulence

Given the importance of the ECF transporters to serum growth and their lesser effect on growth in saliva, we decided to examine them further. We began by attempting to determine the function of these genes, which were annotated as cobalt (Co) transporters (Xu et al., 2007). We first deleted the *ecfA2*, *ecfT*, and *stpA* genes, which belong to the same operon, to create a triple mutant (ΔECF), then, complemented that mutant by placing the three genes at an ectopic expression site used previously for this purpose (Senty Turner et al., 2009). We then employed inductively coupled plasma-optical emission spectroscopy (ICP-OES) to measure the levels of Co and other divalent cations in SK36, the ΔECF mutant, and the complemented mutant. As shown in Figure 5a, there was no significant difference in Co levels in the mutant compared to SK36 or the complemented mutant. This was also true for Fe and Zn. Surprisingly, Mn levels were significantly lower in the ΔECF mutant and significantly higher in the complemented mutant (Figure 5a). We were interested in further investigating this finding, since we had previously linked the virulence reduction of an *ssaB* mutant to reduced Mn uptake (Crump et al., 2014). We therefore combined the Δ*ssaACB* mutation with the ΔECF mutation and tested this combined mutant alongside SK36, the two ECF strains, and the Δ*ssaACB* mutant by ICP-OES. As shown in Figure 5a, Mn levels in the ΔECF mutant, while lower than WT, were significantly higher than those in the Δ*ssaACB* mutant, and deletion of the ECF genes from the Δ*ssaACB* mutant did not result in a significant reduction in Mn levels compared to the Δ*ssaACB* mutation alone. We also tested the same strains for their ability to grow in rabbit serum in 12% of O_2, the same level found in arterial blood (Atkuri et al., 2007), since this would allow a direct comparison with data collected previously for the *ssaB* (Crump et al., 2014) and Δ*ssaACB* (Murgas et al., 2020) mutants. As shown in Figure 5b, all of the

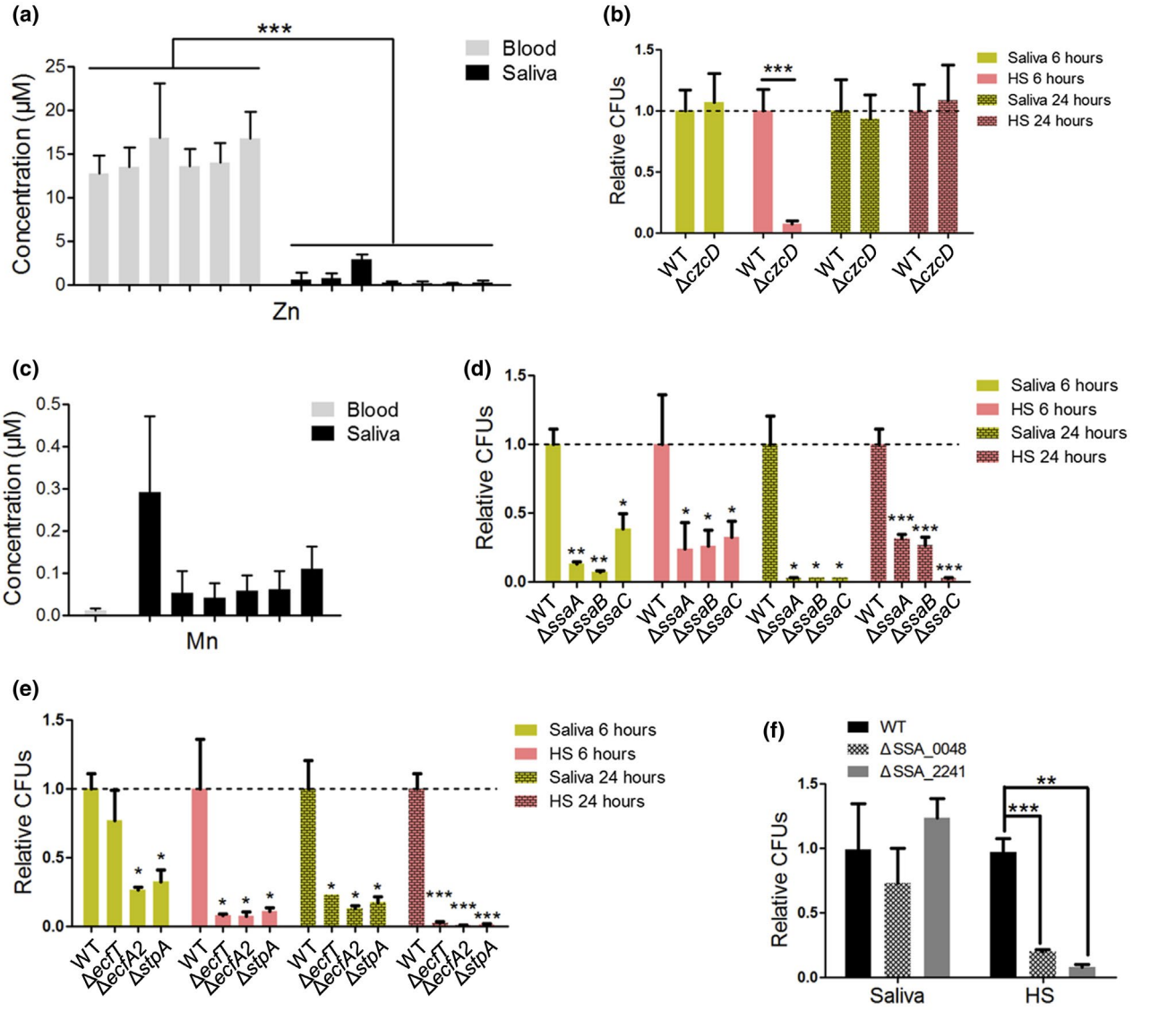

FIGURE 4 Comparison of HS and saliva for metal content and cultivation of *S. sanguinis* strains. (a and c) Metal concentrations in blood and saliva downloaded from HMDB in the form of mean ± standard deviation. Data were analyzed by unpaired *t* test, ****p* value ≤ .001. (b, d, and e) Growth of strains inoculated singly into saliva or HS and incubated for 6 or 24 hr. The CFUs of mutants were expressed relative to those of WT grown under the same condition. (f) Growth of strains inoculated singly into saliva or HS and incubated for 24 hr. CFUs are expressed relative to WT grown under the same condition. Means and standard deviations from triplicate experiments are shown for all growth studies. **p* value ≤ .05, ***p* value ≤ .01, ****p* value ≤ .001, Student's *t* test [Colour figure can be viewed at wileyonlinelibrary.com]

strains grew indistinguishably in the overnight BHI precultures used to inoculate the rabbit serum ("0 h" samples); however, the ΔECF and the Δ*ssaACB* mutants showed no apparent growth, whereas the complemented ΔECF mutant grew to WT levels. Interestingly, the combined ΔECF Δ*ssaACB* mutant appeared to experience net death during the course of the 24-hr incubation in rabbit serum, exhibiting CFU values that were significantly lower than at the time of inoculation and also significantly lower than every other strain at 24 hr.

In a previous study that included SK36 and an *ssaB* mutant, growth in rabbit serum in 12% of O_2 predicted IE virulence in a rabbit model (Crump et al., 2014), with SK36 recovered from infected hearts in numbers ~10^5-fold greater than the *ssaB* mutant. Given that the ΔECF mutant exhibited a defect in serum growth similar to that of the Δ*ssaACB* mutant, we were interested in testing the ΔECF mutant in our rabbit endocarditis model as well (Crump et al., 2014). The ΔECF mutant, the complemented ΔECF mutant, and JFP36, a derivative of SK36 that possesses a gene for erythromycin resistance (Em^r) inserted into the same ectopic locus as the ECF complemented mutant, were co-inoculated into each rabbit. We then used plating on selective antibiotics to determine CFU numbers in the inoculum and the infected vegetations, as we have done previously (Crump et al., 2014). Figure 6 shows that, as with the serum growth study shown in Figure 5b, the ΔECF mutant was recovered at levels ~10^5-fold lower than the Em^r WT strain. The complemented mutant was recovered at levels 1.6-fold higher than—though not significantly different from—the WT strain. Thus, the virulence defect of the ΔECF mutant is comparable to that of the *ssaB* mutant.

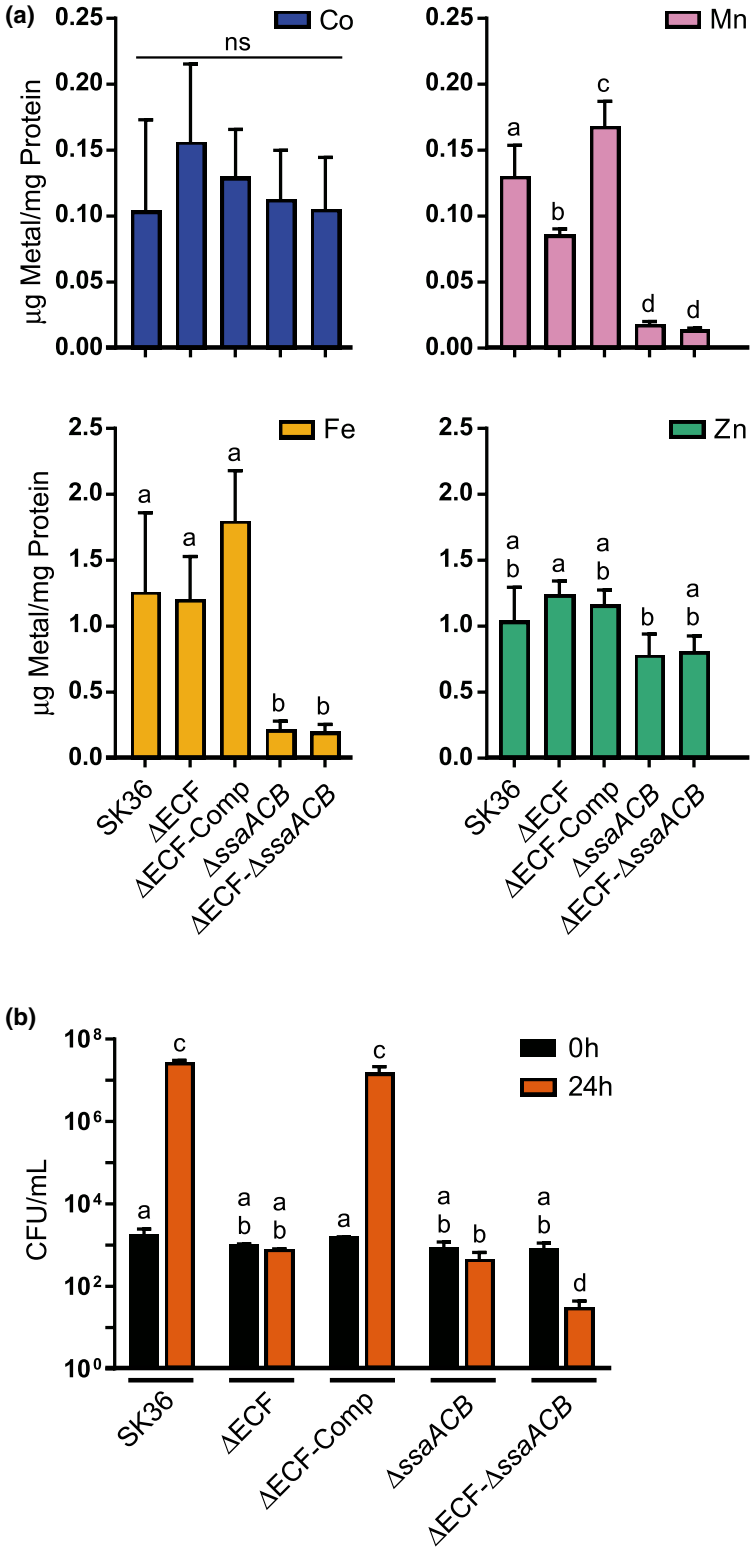

FIGURE 5 Contribution of the ECF and SsaACB transporters to metal uptake and serum growth. The strains listed were examined for (a) metal content by ICP-OES after growth in BHI or (b) growth in pooled rabbit serum for 24 hr. For both assays, all strains were inoculated singly into separate tubes, and three replicates were analyzed by ANOVA (InStat v.3.1, GraphPad software). If a significant difference was detected (p value ≤ .05), samples were further compared with a Tukey posttest. In (a), means and standard deviations are shown, and a separate statistical comparison was performed for each metal. In (b), geometric means and standard deviations are shown, and all samples were compared together. Bars that share a letter indicate values that are not significantly different (p value > .05) within the same comparison group as just defined [Colour figure can be viewed at wileyonlinelibrary.com]

3 | DISCUSSION

Antibiotic resistance is a risk of long-term antibiotic use. To address this issue, we have proposed the idea of developing new drugs to selectively inhibit the growth of bacteria in the bloodstream but not in other locations of the body. Such drugs could potentially be more effective for treatment, but their greatest benefit would likely be for IE prevention. Advisory groups differ in opinion regarding the merit of antibiotic prophylaxis prior to dental procedures, with those opposed suggesting that the cumulative risk from daily bacteremias likely exceeds the risk from infrequent dental procedures (Cahill et al., 2017). An antibiotic that targets bacteria in the bloodstream but not at other body sites might find acceptance by both sides for daily use in patients who have cardiac conditions that put them at

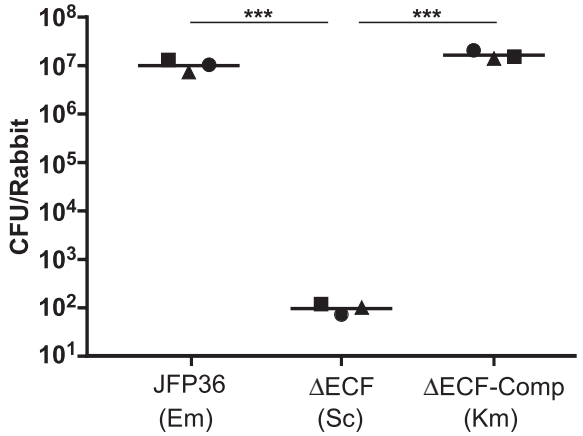

FIGURE 6 Assessment of virulence in an animal model of endocarditis. The strains listed were cultured separately, combined, and inoculated into three rabbits. Bacteria recovered ~20 hr later from infected heart valves were enumerated by plating on replicate BHI plates containing the indicated antibiotic. Horizontal lines indicate geometric mean values. Like symbols indicate the same animal. Em, erythromycin; Sc, spectinomycin; Km, kanamycin; ***p value ≤ .001 by repeated measures ANOVA, followed by Tukey posttest. Other values were not significantly different (p value > .05)

high risk for IE (Thornhill et al., 2018). This strategy may be also helpful for other diseases in which antibiotic prophylaxis is used to prevent illnesses caused by blood-borne bacteria, as in the case of children with sickle cell disease (Gaston et al., 1986).

Tn-seq using random transposon mutant pools has recently been applied to study virulence for a number of pathogens (Le Breton et al., 2013; Miller et al., 2017; Rowe et al., 2019). In this study, a Tn-seq-like approach, ORF-seq, was used to explore genes related to *S. sanguinis* growth in HS. Our results indicated that purine metabolism and several transporters were important for the growth of *S. sanguinis* in HS. Many of the observed results of this ORF-seq screen are consistent with previous findings (Le Breton et al., 2013; Samant et al., 2008; Zhang et al., 2017). For example, synthesis of nucleotides has been found to be required for growth in blood or serum in a number of bacterial species including *Streptococcus pyogenes* (Le Breton et al., 2013), *Escherichia coli*, *Salmonella enterica*, *Bacillus anthracis*, (Samant et al., 2008), and *Enterococcus faecium* (Zhang et al., 2017), and for virulence in others (Chiang & Mekalanos, 1998; Hava & Camilli, 2002, Jones et al., 2000; Mei et al., 1997; Polissi et al., 1998). The Mn transporter SsaB has been well studied in *S. sanguinis* as a virulence factor in IE (Crump et al., 2014; Das et al., 2009). These results suggest that our ORF-seq method worked as intended and may be used to predict fitness genes in other species.

Despite its obvious successes, our ORF-seq approach has some limitations. One is that, as with Tn-seq, all of the mutants are combined into pools. This may mask the effect that some mutations would have on the ability of the corresponding mutant to grow because a growth-limiting metabolite could be provided by the other mutants. Such fitness genes cannot be identified using this method. Another potential issue could be polar effects on the expression of downstream genes, which might lead to misleading results. Both of these problems could be partially addressed by a systems biology approach, in which conclusions come from multiple lines of evidence or from different mutants. For example, our conclusion concerning the importance of purine metabolism for bacterial growth in HS was verified through findings derived from multiple mutants.

We discovered two potential selective drug targets for the inhibition of *S. sanguinis* growth in HS that we predict would preserve growth in saliva. To our knowledge, this is the first report of selective gene targets suppressing bacterial growth in HS but not in saliva. In our previous study, we identified three essential functions of *S. sanguinis*: cell envelope development, energy production, and genetic information processing (Kong & Zhu, 2019; Xu et al., 2011). Based on differences in the essential genes of different bacterial species, we have developed a selective approach that targets key pathogens in a microbiome while sparing commensal bacteria (Stone et al., 2015). Here, we propose another selective approach to target pathogens, based on their relative locations and environments. While we are interested in the oral cavity and blood-borne diseases, the same principle can be applied to develop drugs that target or bypass bacteria at other body sites as well.

We also characterized a mutant lacking the common components of an ECF transport system. As far as we are aware, this is only the second such system to have been identified as important for virulence, with the first being an apparent heme transporter in *S. pyogenes* (Chatterjee et al., 2020). This is also the first ECF system shown to be important for IE virulence in any bacterium. Based on our ICP-OES analysis, the ECF proteins that we characterized play no role in Co transport in *S. sanguinis*. This is not surprising. The *ecfA2*, *ecfT*, and *stpA* genes appear to encode the common components of a Group II transport system. In such systems, the common components drive the transport of multiple substrates by interacting with distinct substrate-binding proteins encoded by noncontiguous genes (Rempel et al., 2019). Moreover, SK36 is annotated as containing a complete Co transport system of the Group I variety, where the substrate-binding protein is encoded adjacent to the other three components. This system is located within a larger locus devoted to synthesis of cobalamin, or vitamin B_{12} (Xu et al., 2007). Considering that cobalamin requires Co as a cofactor, this annotation is inherently plausible. In addition, misannotation of EcfA and EcfT proteins as Co transporters is a common error (Rempel et al., 2019) and the recent reannotation of these genes in GenBank no longer suggests Co transport.

These proteins may play a role in Mn transport. It seems unlikely that reduced Mn transport is responsible for the poor serum growth or virulence of the ΔECF mutant, however, given that this mutant transported far more Mn than the Δ*ssaACB* mutant (Figure 5a) and that serum growth was not restored to the ΔECF mutant by addition of Mn to 100 μM (data not shown). For comparison, we found previously that 5 μM Mn was sufficient to fully restore the serum growth of the Δ*ssaACB* mutant (Murgas et al., 2020). Finally, the ECF components that we characterized presumably interact with all other Group I substrate-binding protein components present in the cell. It is not clear

how many such components there are or the substrates they transport, although the RAST-based annotation of the SK36 genome present on the PATRIC website (https://patricbrc.org/) identifies eight, most of which are thought to transport vitamins or cofactors. It is thus not surprising that the ΔECF mutant displays a severe growth defect.

To date, very few *S. sanguinis* genes have been confirmed as virulence factors using an animal model (Bai et al., 2019; Crump et al., 2014; Ge et al., 2008; Martini et al., 2020). We have likely identified many additional promising candidates. By ORF-seq, the Δ*ssaB* and Δ*purB* mutants exhibited a $\log_2$ fold change in HS of −2.49 and −3.99, respectively. If these values are compared to the abundance of these mutants relative to WT in an IE animal model ($\log_2$ fold change = −11.75 and −7.31, respectively), the defect is clearly much greater in the animals. Similarly, the $\log_2$ fold changes in HS for the individual ECF mutants ranged from −1.58 to −1.85, whereas the $\log_2$ fold change for the triple mutant in the IE model was −16.67 (Figure 6). In contrast, in our previous study (Crump et al., 2014) growth of Δ*ssaB* in rabbit serum matched the results of the IE model ($\log_2$ fold change = −14.14 for IE vs. −19.19 for serum), just as it did for the ΔECF mutant in this study ($\log_2$ fold change = −15.95 for serum vs. −16.67 for IE; Figure 5b vs. Figure 6). The ECF result was not due to a difference between the triple ΔECF mutant versus the individual mutants or human versus rabbit serum, since we observed similar results when each of the three individual ECF mutants was grown in rabbit serum (not shown). The explanation is most likely the degree to which the inocula were diluted for the two experiments. For ORF-seq, the mutant pool was diluted 100-fold into HS, whereas each strain was diluted 10^6-fold into rabbit serum. Thus, a WT strain would be expected to attain an abundance ~100-fold that of a nongrowing strain by ORF-seq, whereas a ~10^6-fold difference would be expected in the rabbit serum studies. The latter experiment is a better model of in vivo disease, where we have observed colonization of cardiac tissues by ~10^2 CFU and recovery the following day of ~10^7 CFU per animal (Crump et al., 2014). We can therefore conclude that the ORF-seq study provides a highly conservative view of the virulence defects that we would likely see in the IE model or in human disease. Thus, many of the mutants displaying lesser defects by ORF-seq may be worth further examination with our standard rabbit serum growth assay, followed by testing in the in v*i*vo model.

4 | MATERIALS AND METHODS

4.1 | Bacterial strains

Strains used in the ORF-seq analysis were from our comprehensive mutant pool, which was generated from strain SK36 (Kilian et al., 1989) in our previous study (Ge & Xu, 2012). Additional strains used in this study are described below and listed in Table S1. The ΔECF mutant was constructed by replacing the three ECF genes (SSA_2367, SSA_2366, and SSA_2365, in that order) with the *aad9* cassette encoding Sc^r (Senty Turner et al., 2009) by overlap extension PCR (Ho et al., 1989). The *aad9* cassette, including its promoter, was inserted immediately downstream from the stop codon of the upstream gene, SSA_2368. We note that we relied on the 2020 GenBank annotation for this deletion, in which the start codon of SSA_2367 is separated by 1 bp from the stop codon of SSA_2368, rather than the two ORFs overlapping as in the earlier annotation. The *aad9* cassette extended exactly to the stop codon of SSA_2365, which was 210 bp from the downstream ORF, SSA_2364. Based on expression data from multiple RNA-seq studies, this ORF appears to be in a separate operon from the ECF genes (data not shown). For complementation of this mutant, the same three genes were placed behind the *aphA-3* gene with its own promoter used for construction of the nonessential gene library (Ge & Xu, 2012). These four genes were then inserted into the SSA_0169 gene. Although this disrupts the gene, we have shown previously that this has no effect on growth, IE virulence, or a number of other phenotypes (Senty Turner et al., 2009). To make it possible to use different antibiotics for selection of the ΔECF mutant and the complemented mutant, we performed an exact replacement of the *aad9* cassette in the complemented mutant with the *tetM* cassette including its promoter, encoding tetracycline resistance (Senty Turner et al., 2009). The Δ*ssaACB* mutant (JFP169) was constructed by an exact replacement of the *ssaACB* genes with the promoterless version of the *aphA-3* gene, also used for construction of the nonessential gene library (Ge & Xu, 2012). This strain relies on the strong, native *ssaACB* promoter for expression of the *aphA-3* gene. This strain has been characterized previously (Murgas et al., 2020) (also Puccio, et al, 2020). For construction of the ΔECF-ΔssaACB mutant, the *aad9* cassette with its flanking DNA in the ΔECF mutant was amplified by PCR and introduced into the Δ*ssaACB* mutant by transformation. All mutants were confirmed to possess the expected sequences by Sanger sequencing (Retrogen, Inc.).

4.2 | Growth conditions

Unless otherwise stated, strains were grown under microaerobic conditions (6% O_2, 7.2% CO_2, 7.2% H_2, and 79.6% N_2) at 37°C using an Anoxomat® system (Spiral Biotech, Norwood, MA). HS was prepared from frozen serum obtained commercially from anonymous donors (normal pool, MP Biomedicals™, cat No. MP092930149) that was thawed at room temperature and heat inactivated by incubation for 30 min at 56°C. Heat inactivation was performed to eliminate possible differences in complement levels among different serum lots. For the ORF-seq experiments, the mutant library was grown in BHI (Difco Inc., Detroit, MI) or HS for 24 hr. For in vitro growth studies involving individual mutants, strains were grown individually in BHI overnight, and then, diluted 100-fold into BHI, HS, or human saliva obtained commercially from anonymous donors (pooled human donors, LeeBio, Catalog Number: 991-05-P) for 24 hr unless otherwise indicated. Growth in rabbit serum was examined as described previously (Baker et al., 2018). Briefly, cells were grown overnight at 37°C in BHI with 6% of O_2, then, diluted 10^6-fold into pooled rabbit serum (Gibco) that had been preincubated in 12% of O_2 (12% O_2, 4.3% H_2, 4.3% CO_2, and 80% N_2). Incubation was continued in this atmosphere for 24 hr, then, 1 ml was removed, sonicated, diluted in

PBS, and plated onto BHI agar for enumeration. Apart from the minute amounts in the inocula, BHI was not added to HS, human saliva, or rabbit serum to support growth. When needed, antibiotics were added at the following concentrations: kanamycin (Km), 500 μg/ml; spectinomycin (Sc), 200 μg/ml; tetracycline (Tc), 5 μg/ml; erythromycin (Em), 10 μg/ml.

4.3 | ORF-seq

The tag that was linked to the kanamycin resistance gene (*aphA*-3) used to create each mutant (Xu et al., 2011) served as the basis for determination of the abundance of each mutant in the pooled library via ORF-seq. Our method is a modification of that described by Turner et al. (2015), which employed a poly(C) tail. In ORF-seq, the *aphA*-3 gene is present in every mutant and is thus analogous to the transposon present in every mutant in the Tn-seq procedure. A 5′ primer complementary to the 3′ end of *aphA*-3 was used in conjunction with a 3′ primer containing a poly-G tail to amplify DNA sequences located adjacent to the *aphA*-3 gene to identify each mutant. Each mutant in the library was separately cultured, combined, and washed with PBS. The library mixture was then diluted 100-fold into HS or BHI and incubated microaerobically for 24 hr at 37°C. Cells were harvested by centrifugation at 1,700 g for 15 min at 4°C and washed with PBS. Genomic DNA was isolated using a QIAamp DNA Mini Kit (Qiagen, USA). Subsequently, genomic DNA was sheared by sonication using a Covaris S2 Ultrasonicator (Qsonica, USA). PolyC was added to 3′ ends of the fragmented library by Terminal Deoxynucleotidyl Transferase (Promega, USA). The DNA fragments with polyC were purified by AMPure XP beads (Beckman, USA) for use as template for the next step of PCR. PCR was performed using primers olj 376 and K10_Truseq (Table S1). The PCR program was as follows: 98°C for 45 s, then, 20 cycles of 98°C for 10 s, 60°C for 30 s, and 72°C for 30 s. This was followed by 72°C for 5 min and a hold at 4°C. The PCR products were purified by AMPure XP beads, following by the second round of PCR using the primers Truseq_HT01 for the BHI set and Truseq_HT02 for the HS set (Table S1). PE1npKan was used in both sets as the 3′ primer (Table S1). The PCR program was as follows: 98°C for 45 s, 10–15 cycles of 98°C for 10 s, 60°C for 30 s, and 72°C for 30 s, followed by 72°C for 5 min and a hold at 4°C. The PCR products were purified by AMPure XP Beads and sent to the VCU DNA Core Facility for NGS sequencing. Each experiment was performed in triplicate.

4.4 | Metal analysis

Cellular metal content was determined by ICP-OES exactly as described previously (Baker et al., 2018). In brief, cells grown in BHI were harvested, washed in cold, Chelex-treated (Bio-Rad) PBS, and resuspended in the same buffer. An aliquot was removed for lysis and protein determination using a bicinchoninic acid (BCA) protein assay kit (Pierce). The remainder of the sample was centrifuged, resuspended in Chelex-treated deionized water, to which was then added concentrated (67% to 70%) trace metal-grade nitric acid. The mixture was then placed into a modified polytetrafluoroethylene (TFM) digestion vessel in a Multiwave Go microwave digestion system (Anton Paar) for digestion at 20×10^5 Pa. Following dilution in Chelex-treated water, samples were analyzed on an Agilent 5110 ICP-VDV OES with a Pb internal standard (MSPB-10PPM; Inorganic Ventures) against a standard curve made with a CMS-5 metal standard (Inorganic Ventures, Inc.). Each sample was analyzed for Co, Mn, Fe, and Zn. Results were expressed relative to protein amounts determined above.

4.5 | Virulence studies

Specific-pathogen-free, male New Zealand White rabbits (2–4 kg; Charles River Labs Inc., Wilmington MA) were used to assess virulence as previously described (Crump et al., 2014). Rabbits were sedated and anesthetized before insertion of a catheter into the right carotid artery. The catheter was trimmed and sutured in place to induce minor damage to the aortic valve. The incision site was closed with staples. The following day, all strains were inoculated into BHI for overnight growth in 6% of O_2. The next morning, the cultures were diluted 10-fold into fresh, prewarmed BHI and incubated for 3 hr before harvesting by centrifugation and washing twice in PBS. The washed cells were diluted and combined in PBS to yield an inoculum containing approximately 10^7 CFU/ml of each strain. Bacterial enumeration of the remaining inoculum was performed by removing 1 ml of the prepared inoculum, sonicating, and plating a dilution series onto replicate BHI agar plates, each of which contained one of the three antibiotics employed in the experiment, using an Eddy Jet 2 spiral plater (Advanced Instruments, Inc., Norwood, MA). The peripheral ear vein of each rabbit was injected with 0.5 ml of the combined inoculum. Rabbits were sacrificed via intravenous injection of Euthasol approximately 20 hr after inoculation. Catheter placement was assessed upon necropsy and cardiac vegetations were collected, and then, placed into PBS and homogenized. The homogenates were sonicated, serially diluted, and plated onto replicate BHI agar plates, each containing one of three antibiotics as per the inoculum, and incubated overnight anaerobically. Resulting colonies were counted to determine CFU/rabbit for each strain, with numbers normalized to account for small differences in inoculum numbers for each strain. All animal experiments were approved by the Institutional Animal Care and Use Committee of the Virginia Commonwealth University, and were performed in compliance with local, state, and federal regulations.

4.6 | Data analysis

Reads obtained from ORF-seq were aligned against the *S. sanguinis* SK36 genome using Rockhopper (Tjaden, 2015). Analyses were run using

default parameter settings. Significance was determined by a *p* value adjusted for a false discovery rate of 1%. Three replicates were examined for data analysis. The set of genes whose mutants had significant abundance changes was input into the STRING database for the construction of a protein–protein interaction network (Szklarczyk et al., 2017). The results of the STRING analysis were imported to the Gephi software program for the visualization and calculation of the betweenness centrality (Bastian et al., 2009). Genes in the STRING network were classified by COG annotation (Tatusov et al., 2000). Genes identified as affecting abundance in HS or BHI were input into the DAVID database for the KEGG pathways enrichment analyses. The significantly enriched KEGG pathways (p value ≤ .05) were identified. The purine and pyrimidine metabolism pathways were analyzed by mapping related *S. sanguinis* genes onto the KEGG nucleotide transport and metabolism pathway (Kanehisa et al., 2017). The function of transporters in *S. sanguinis* was also classified based on the information obtained from the KEGG database (Kanehisa et al., 2017). The HMDB was used to explore the concentration of compounds in blood and saliva (Wishart et al., 2018). The data of metabolite concentrations in the form of mean ± standard deviation was viewed. Note that samples are labeled as "blood" in this database even if they are actually plasma or serum (Wishart et al., 2018). The growth of individual *S. sanguinis* mutants in HS and saliva was investigated by dilution plating of bacterial cultures. Three replicates were performed to measure means and standard deviations. Significance was assessed using Student's *t* test. The growth of strains in rabbit serum and their recovery in the rabbit endocarditis model were assessed by ANOVA and Tukey posttest applied to log-transformed data.

ACKNOWLEDGMENTS

We thank Karina Kunka for assistance with figure preparation, proofreading, and statistical analysis, Dr. Seon-Sook An and Nicai Zollar for assistance with the animal surgeries, and Dr. Joseph Turner for assistance with the ICP-OES instrument. This work was supported by National Institutes of Health grants R01DE023078 (PX) and R01AI114926 (TK) and the CCTR Endowment Fund (PX). The funders had no role in the study design, data collection or interpretation, or the decision to submit the work for publication.

CONFLICT OF INTERESTS

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

XG and PX conceived and designed the ORF-seq study. XG carried out ORF-seq assays. BZ, LB, HN, and HG performed HS growth studies. SG designed and performed the metal analysis and rabbit serum growth studies. SG, TP, and TK designed and performed the virulence study. Data analysis was performed by BZ, HG, and SG. BZ, TK, HN, and PX wrote this manuscript with assistance from SG and TP. All authors reviewed and discussed the manuscript.

DATA AVAILABILITY STATEMENT

All data generated or analyzed during this study are included in this published article (and its supplementary files).

ORCID

Bin Zhu https://orcid.org/0000-0003-2829-2925
Shannon P. Green https://orcid.org/0000-0002-5925-246X
Tanya Puccio https://orcid.org/0000-0002-0223-4885
Todd Kitten https://orcid.org/0000-0001-6097-7583
Ping Xu https://orcid.org/0000-0002-8368-801X

REFERENCES

Atkuri, K.R., Herzenberg, L.A., Niemi, A.-K., Cowan, T. & Herzenberg, L.A. (2007) Importance of culturing primary lymphocytes at physiological oxygen levels. *Proceedings of the National Academy of Sciences*, *104*, 4547–4552.

Baddour, L.M., Wilson, W.R., Bayer, A.S., Fowler, V.G. Jr., Tleyjeh, I.M., Rybak, M.J., et al. (2015) Infective endocarditis in adults: diagnosis, antimicrobial therapy, and management of complications: A scientific statement for healthcare professionals from the American Heart Association. *Circulation*, *132*, 1435–1486.

Bai, Y., Shang, M., Xu, M., Wu, A., Sun, L. & Zheng, L. (2019) Transcriptome, phenotypic, and virulence analysis of *Streptococcus sanguinis* SK36 wild type and its CcpA-Null Derivative (ΔCcpA). *Frontiers in Cellular and Infection Microbiology*, *9*, 411.

Baker S. P., Nulton T. J., Kitten T. (2018) Genomic, Phenotypic, and Virulence Analysis of Streptococcus sanguinis Oral and Infective-Endocarditis Isolates. *Infection and Immunity*, *87* (1), e00703–18. http://dx.doi.org/10.1128/iai.00703-18

Bastian, M., Heymann, S. & Jacomy, M. (2009) Gephi: an open source software for exploring and manipulating networks. In: *International AAAI Conference on Web and Social Media*. San Jose, CA: 361–362. https://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154

Bugrysheva, J., Froehlich, B.J., Freiberg, J.A. & Scott, J.R. (2011) Serine/threonine protein kinase Stk is required for virulence, stress response, and penicillin tolerance in *Streptococcus pyogenes*. *Infection and Immunity*, *79*, 4201–4209.

Cahill, T.J., Harrison, J.L., Jewell, P., Onakpoya, I., Chambers, J.B., Dayer, M., et al. (2017) Antibiotic prophylaxis for infective endocarditis: a systematic review and meta-analysis. *Heart*, *103*, 937–944.

Chatterjee, N., Cook, L.C.C., Lyles, K.V., Nguyen, H.A.T., Devlin, D.J., Thomas, L.S., et al. (2020) A novel heme transporter from the ECF family is vital for the Group A *Streptococcus* colonization and infections. *Journal of Bacteriology*.*202*(14), e00205–20.

Chen, L., Ge, X., Wang, X., Patel, J.R. & Xu, P. (2012) SpxA1 involved in hydrogen peroxide production, stress tolerance and endocarditis virulence in *Streptococcus sanguinis*. *PloS One*, *7*, e40034.

Chiang, S.L. & Mekalanos, J.J. (1998) Use of signature-tagged transposon mutagenesis to identify *Vibrio cholerae* genes critical for colonization. *Molecular Microbiology*, *27*, 797–805.

Crump, K.E., Bainbridge, B., Brusko, S., Turner, L.S., Ge, X., Stone, V., et al. (2014) The relationship of the lipoprotein SsaB, manganese and superoxide dismutase in *Streptococcus sanguinis* virulence for endocarditis. *Molecular Microbiology*, *92*, 1243–1259.

Das, S., Kanamoto, T., Ge, X., Xu, P., Unoki, T., Munro, C.L., et al. (2009) Contribution of lipoproteins and lipoprotein processing to endocarditis virulence in *Streptococcus sanguinis*. *Journal of Bacteriology*, *191*, 4166–4179.

European Society of Cardiology. (2015) The 2015 ESC Guidelines for the management of infective endocarditis. *European Heart Journal*, *36*, 3036–3037.

Freeman, L.C. (1977) A set of measures of centrality based on betweenness. *Sociometry*, *40*, 35–41.

Gaston, M.H., Verter, J.I., Woods, G., Pegelow, C., Kelleher, J., Presbury, G., et al. (1986) Prophylaxis with oral penicillin in children with sickle cell anemia. A randomized trial. *The New England Journal of Medicine*, *314*, 1593–1599.

Ge, X., Kitten, T., Chen, Z., Lee, S.P., Munro, C.L. & Xu, P. (2008) Identification of *Streptococcus sanguinis* genes required for biofilm formation and examination of their role in endocarditis virulence. *Infection and Immunity*, *76*, 2551–2559.

Ge, X. & Xu, P. (2012) Genome-wide gene deletions in *Streptococcus sanguinis* by high throughput PCR. *Journal of Visualized Experiments: JoVE.23*(69), 4356

Griffen, A.L., Beall, C.J., Campbell, J.H., Firestone, N.D., Kumar, P.S., Yang, Z.K., et al. (2012) Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *The ISME Journal*, *6*, 1176–1185.

Hava, D.L. & Camilli, A. (2002) Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Molecular Microbiology*, *45*, 1389–1406.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. & Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, *77*, 51–59.

Huang, D.W., Sherman, B.T. & Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*, 44–57.

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

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, *45*, D353–D361.

Kilian, M., Mikkelsen, L. & Henrichsen, J. (1989) Taxonomic study of viridans streptococci: description of *Streptococcus gordonii* sp. nov. and emended descriptions of *Streptococcus sanguis* (White and Niven 1946), *Streptococcus oralis* (Bridge and Sneath 1982), and *Streptococcus* mitis (Andrewes and Horder 1906). *International Journal of Systematic Evolutionary Microbiology*, *39*, 471–484.

Kilstrup, M., Hammer, K., Ruhdal Jensen, P. & Martinussen, J. (2005) Nucleotide metabolism and its control in lactic acid bacteria. *FEMS Microbiology Reviews*, *29*, 555–590.

Kong, X. & Zhu, B. (2019) ePath: an online database towards comprehensive essential gene annotation for prokaryotes. *Scientific Reports*, *9*, 12949.

Kreth, J., Merritt, J., Shi, W. & Qi, F. (2005) Competition and coexistence between *Streptococcus mutans* and *Streptococcus sanguinis* in the dental biofilm. *Journal of bacteriology*, *187*, 7193–7203.

Le Breton, Y., Mistry, P., Valdes, K.M., Quigley, J., Kumar, N., Tettelin, H., et al. (2013) Genome-wide identification of genes required for fitness of group A *Streptococcus* in human blood. *Infection and Immunity*, *81*, 862–875.

Leonard, A., Gierok, P., Methling, K., Gómez-Mejia, A., Hammerschmidt, S. & Lalk, M. (2018) Metabolic inventory of *Streptococcus pneumoniae* growing in a chemical defined environment. *International Journal of Medical Microbiology*, *308*, 705–712.

Martin, J.E. & Giedroc, D.P. (2016) Functional determinants of metal ion transport and selectivity in paralogous cation diffusion facilitator transporters CzcD and MntE in *Streptococcus pneumoniae*. *Journal of Bacteriology*, *198*, 1066–1076.

Martini, A.M., Moricz, B.S., Ripperger, A.K., Tran, P.M., Sharp, M.E., Forsythe, A.N., et al. (2020) Association of novel *Streptococcus sanguinis* virulence factors with pathogenesis in a native valve infective endocarditis model. *Frontiers in Microbiology*, *11*, 10.

Mei, J.M., Nourbakhsh, F., Ford, C.W. & Holden, D.W. (1997) Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. *Molecular Microbiology*, *26*, 399–407.

Mendes, R.E., Deshpande, L.M., Kim, J., Myers, D.S., Ross, J.E. & Jones, R.N. (2013) *Streptococcus sanguinis* isolate displaying a phenotype with cross-resistance to several rRNA-targeting agents. *Journal of Clinical Microbiology*, *51*, 2728–2731.

Miller, D.P., Hutcherson, J.A., Wang, Y., Nowakowska, Z.M., Potempa, J., Yoder-Himes, D.R., et al. (2017) Genes contributing to *Porphyromonas gingivalis* fitness in abscess and epithelial cell colonization environments. *Frontiers in Cellular and Infection Microbiology*, *7*, 378.

Murgas, C.J., Green, S.P., Forney, A.K., Korba, R.M., An, S.S., Kitten, T., et al. (2020) Intracellular metal speciation in *Streptococcus sanguinis* establishes SsaACB as critical for redox maintenance. *ACS Infectious Diseases.6*(7), 1906–1921.

Paik, S., Senty, L., Das, S., Noe, J.C., Munro, C.L. & Kitten, T. (2005) Identification of virulence determinants for endocarditis in *Streptococcus sanguinis* by signature-tagged mutagenesis. *Infection and Immunity*, *73*, 6064–6074.

Pericàs, J.M., Nathavitharana, R., Garcia-de-la-Mària, C., Falces, C., Ambrosioni, J., Almela, M., et al. (2019) Endocarditis caused by highly penicillin-resistant viridans group Streptococci: still room for vancomycin-based regimens. *Antimicrobial Agents and Chemotherapy*, *63*(8).e00516–19.

Polissi, A., Pontiggia, A., Feger, G., Altieri, M., Mottl, H., Ferrari, L. & Simon, D. (1998) Large-scale identification of virulence genes from *Streptococcus pneumoniae*. *Infection and Immunity*, *66*, 5620–5629.

Puccio, T., Kunka, K.S., Zhu, B., Xu, P. & Kitten, T. (2020) Manganese depletion leads to multisystem changes in the transcriptome of the opportunistic pathogen Streptococcus sanguinis. *Frontiers in Microbiology*, *11* https://doi.org/10.3389/fmicb.2020.592615/full.

Rasmussen, T.B., Zwisler, A.D., Thygesen, L.C., Bundgaard, H., Moons, P. & Berg, S.K. (2017) High readmission rates and mental distress after infective endocarditis—results from the national population-based CopenHeart IE survey. *Infection and Immunity*, *235*, 133–140.

Rempel, S., Stanek, W.K. & Slotboom, D.J. (2019) ECF-Type ATP-binding cassette transporters. *Annual Review of Biochemistry*, *88*(1), 551–576. .

Rowe, H.M., Karlsson, E., Echlin, H., Chang, T.C., Wang, L., van Opijnen, T., et al. (2019) Bacterial factors required for transmission of *Streptococcus pneumoniae* in mammalian hosts. *Cell Host & Microbe*, *25*, 884–891.e886.

Safdar, A. & Rolston, K.V. (2006) Vancomycin tolerance, a potential mechanism for refractory gram-positive bacteremia observational study in patients with cancer. *Cancer*, *106*, 1815–1820.

Samant, S., Lee, H., Ghassemi, M., Chen, J., Cook, J.L., Mankin, A.S., et al. (2008) Nucleotide biosynthesis is critical for growth of bacteria in human blood. *PLoS Pathog*, *4*, e37.

Senty Turner, L., Das, S., Kanamoto, T., Munro, C.L. & Kitten, T. (2009) Development of genetic tools for in vivo virulence analysis of *Streptococcus sanguinis*. *Microbiology (Reading, England)*, *155*, 2573–2582.

Shaw, L.P., Bassam, H., Barnes, C.P., Walker, A.S., Klein, N. & Balloux, F. (2019) Modelling microbiome recovery after antibiotics using a stability landscape framework. *ISME Journal*, *13*, 1845–1856.

Stingu, C.S., Eschrich, K., Rodloff, A.C., Schaumann, R. & Jentsch, H. (2008) Periodontitis is associated with a loss of colonization by *Streptococcus sanguinis*. *Journal of Medical Microbiology*, *57*, 495–499.

Stone, V.N., Parikh, H.I., El-rami, F., Ge, X., Chen, W., Zhang, Y., et al. (2015) Identification of small-molecule inhibitors against meso-2, 6-diaminopimelate dehydrogenase from *Porphyromonas gingivalis*. *PloS One*, *10*, e0141126.

Stone, V.N. & Xu, P. (2017) Targeted antimicrobial therapy in the microbiome era. *Molecular Oral Microbiology*, *32*, 446–454.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*, D607–D613.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Research*, *45*, D362–D368.

Tatusov, R.L., Galperin, M.Y., Natale, D.A. & Koonin, E.V. (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research*, *28*, 33–36.

Thornhill, M.H., Gibson, T.B., Cutler, E., Dayer, M.J., Chu, V.H., Lockhart, P.B., et al. (2018) Antibiotic prophylaxis and incidence of endocarditis before and after the 2007 AHA recommendations. *Journal of the American College of Cardiology*, *72*, 2443–2454.

Tjaden, B. (2015) De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biology*, *16*, 1.

Turner, K.H., Wessel, A.K., Palmer, G.C., Murray, J.L. & Whiteley, M. (2015) Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 4110–4115.

Wilson, W., Taubert, K.A., Gewitz, M., Lockhart, P.B., Baddour, L.M., Levison, M., et al. (2007) Prevention of infective endocarditis: guidelines from the American heart association: a guideline from the American heart association rheumatic fever, endocarditis, and Kawasaki disease committee, council on cardiovascular disease in the young, and the council on clinical cardiology, council on cardiovascular surgery and anesthesia, and the quality of care and outcomes research interdisciplinary working group. *Circulation*, *116*(15), 1736–1754.

Wishart, D.S., Feunang, Y.D., Marcu, A., Guo, A.C., Liang, K., Vazquez-Fresno, R., et al. (2018) HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research*, *46*, D608–D617.

Wray, D., Ruiz, F., Richey, R. & Stokes, T. (2008) Prophylaxis against infective endocarditis for dental procedures–summary of the NICE guideline. *British Dental Journal*, *204*, 555–557.

Xu, P., Alves, J.M., Kitten, T., Brown, A., Chen, Z., Ozaki, L.S., et al. (2007) Genome of the opportunistic pathogen *Streptococcus sanguinis*. *Journal of Bacteriology*, *189*, 3166–3175.

Xu, P., Ge, X., Chen, L., Wang, X., Dou, Y., Xu, J.Z., et al. (2011) Genome-wide essential gene identification in *Streptococcus sanguinis*. *Scientific Reports*, *1*, 125.

Zhang, X., de Maat, V., Guzmán Prieto, A.M., Prajsnar, T.K., Bayjanov, J.R., de Been, M., et al. (2017) RNA-seq and Tn-seq reveal fitness determinants of vancomycin-resistant *Enterococcus faecium* during growth in human serum. *BMC Genomics*, *18*, 893.

Zhu, B., Macleod, L.C., Kitten, T. & Xu, P. (2018) *Streptococcus sanguinis* biofilm formation & interaction with oral pathogens. *Future Microbiology*, *13*, 915–932.

Zhu, B., Macleod, L.C., Newsome, E., Liu, J. & Xu, P. (2019) *Aggregatibacter actinomycetemcomitans* mediates protection of *Porphyromonas gingivalis* from *Streptococcus sanguinis* hydrogen peroxide production in multi-species biofilms. *Scientific Reports*, *9*, 4944.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

How to cite this article: Zhu B, Green SP, Ge X, et al. Genome-wide identification of *Streptococcus sanguinis* fitness genes in human serum and discovery of potential selective drug targets. *Mol Microbiol*. 2021;115:658–671. https://doi.org/10.1111/mmi.14629