bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 # Novel genes required for the fitness of *Streptococcus pyogenes*
2 # in human saliva

3

4

5 **Luchang Zhu[1], Amelia R. L. Charbonneau[3,4], Andrew S. Waller[3],**
6 **Randall J. Olsen[1,2], Stephen B. Beres[1], and James M. Musser[1,2#]**

7

8 [1]Center for Molecular and Translational Human Infectious Diseases Research,
9 Houston Methodist Research Institute, and Department of Pathology and
10 Genomic Medicine, Houston Methodist Hospital, Houston, Texas, USA
11 [2]Department of Pathology and Laboratory Medicine, Weill Medical College of
12 Cornell University, New York, New York, USA
13 [3]Animal Health Trust, Lanwades Park, Newmarket, Suffolk, UK
14 [4]Department of Veterinary Medicine, University of Cambridge, Cambridge, UK

15

16 [#]Address correspondence to James M. Musser,
17 jmmusser@houstonmethodist.org

18

19 Running Head: *Streptococcus pyogenes* fitness genes in human saliva

20

21 Abstract Word Count: 247
22 Manuscript Word Count:

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

23 **ABSTRACT** *Streptococcus pyogenes* (group A *Streptococcus*, or GAS) causes
24 600 million cases of pharyngitis each year. Despite this considerable disease
25 burden, the molecular mechanisms used by GAS to infect, cause clinical
26 pharyngitis, and persist in the human oropharynx are poorly understood. Saliva is
27 ubiquitous in the human oropharynx and is the first material GAS encounters in
28 the upper respiratory tract. Thus, a fuller understanding of how GAS survives and
29 proliferates in saliva may provide valuable insights into the molecular
30 mechanisms at work in the human oropharynx. We generated a highly saturated
31 transposon insertion mutant library in serotype M1 strain MGAS2221, a strain
32 genetically representative of a pandemic clone that arose in the 1980s and
33 spread globally. The transposon mutant library was exposed to human saliva to
34 screen for GAS genes required for wild-type fitness in this clinically relevant fluid.
35 Using transposon-directed insertion site sequencing (TraDIS), we identified 92
36 genes required for GAS fitness in saliva. The more prevalent categories
37 represented are genes involved in carbohydrate transport/metabolism, amino
38 acid transport/metabolism, and inorganic ion transport/metabolism. Using six
39 isogenic mutant strains, we confirmed that each of the mutants are significantly
40 impaired for growth or persistence in human saliva *ex vivo*. Mutants with an
41 inactivated *spy0644* (*sptA*) or *spy0646* (*sptC*) gene have especially severe
42 persistence defects. This study is the first use of TraDIS to study bacterial fitness
43 in human saliva. The new information we obtained is valuable for future
44 translational maneuvers designed to prevent or treat human GAS infections.
45

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

46 **IMPORTANCE** The human bacterial pathogen *Streptococcus pyogenes*
47 (group A streptococcus, GAS) causes more than 600 million cases of pharyngitis
48 annually worldwide, 15 million of which occur in the United States. The human
49 oropharynx is the primary anatomic site for GAS colonization and infection, and
50 saliva is the first material encountered. Using a genome-wide transposon mutant
51 screen, we identified 92 GAS genes required for wild-type fitness in human
52 saliva. Many of the identified genes are involved in carbohydrate
53 transport/metabolism, amino acid transport/metabolism, and inorganic ion
54 transport/metabolism. The new information is potentially valuable for developing
55 novel GAS therapeutics and vaccine research.
56
57
58
59 KEYWORDS *Streptococcus pyogenes*, saliva, transposon screen, TraDIS,
60 fitness genes
61
62
63
64
65
66
67
68

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

69 **[INTRODUCTION]**

70 Bacterial pathogens have evolved highly specialized molecular strategies to
71 survive and persist in diverse host niches (1, 2). Understanding the molecular
72 mechanisms contributing to bacterial fitness in human environments is valuable
73 for developing therapeutic strategies to treat and potentially prevent infections.
74 *Streptococcus pyogenes* (group A streptococcus, or GAS) is a significant human
75 pathogen that causes extensive health and economic impact globally (3). The
76 human oropharynx is the primary anatomic site for GAS colonization and
77 infection (3-5). This pathogen causes 600 million cases of pharyngitis annually
78 worldwide, 15 million of which occur in the United States (3). The annual direct
79 health care costs associated with GAS pharyngitis are estimated to be 2 billion
80 dollars annually in the United States alone (3, 5). The organism is also
81 responsible for an additional 100 million cases of other human infections each
82 year, many of which occur after initial colonization of the oropharynx (3). These
83 additional infections include acute rheumatic fever and subsequent rheumatic
84 heart disease, and as a consequence is the most common cause of preventable
85 pediatric heart disease globally (3, 6). The majority of cases of rheumatic fever
86 occur following human upper respiratory tract infection. Despite the extensive toll
87 on human health, the molecular mechanisms used by GAS to successfully
88 colonize, cause acute pharyngitis and persist in the human oropharynx remain
89 largely unknown or poorly understood (7, 8). This lack of knowledge constitutes a
90 critical knowledge gap in our understanding of GAS pathogenesis, and thus
91 represents an opportunity for enhanced understanding of the molecular

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

92 mechanisms at work during the critical initial pathogen interaction with the human
93 host.
94 The oropharynx is the primary site of entry for GAS into the body and the
95 major portal of person-to-person transmission (9-11). Several observations have
96 stimulated our interest in studying the molecular genetic basis of interaction
97 between human saliva and GAS. Saliva is ubiquitous in the human oropharynx
98 and is the first host material contacted by GAS in its common transmission cycle.
99 Compared to individuals with clinical pharyngitis who lack GAS in their saliva,
100 patients with GAS present in their saliva are more likely to transmit the organism
101 to a new host by aerosolization (10-12). Thus, understanding how GAS survives
102 and proliferates in saliva and the oropharynx may provide valuable insights into
103 the molecular mechanisms underlying successful bacterial interaction in this
104 niche.
105 Previous studies addressing GAS-saliva interactions have identified
106 several factors that contribute to bacterial fitness (13-18), but knowledge is
107 limited. The studies conducted by Sitkiewicz et al. (13) and Virtaneva et al. (14,
108 15) investigated gene responses of a serotype M1 GAS strain grown in human
109 saliva *ex vivo* (16). Subsequently, Shelburne et al (17) identified a key two-
110 component transcriptional regulatory system (SptR/S) of previously unknown
111 function that plays central role in optimizing GAS persistence in saliva. The
112 SptR/S two-component system influences multiple GAS metabolic pathways and
113 production of many virulence factors (17). For example, the secreted virulence
114 factors streptococcal inhibitor of complement (Sic) and a potent extracellular

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

115 cysteine protease (SpeB) made by GAS during growth *ex vivo* in saliva
116 contribute significantly to GAS persistence in this fluid (16). Additional
117 information about GAS interaction in the oropharynx was provided by a 20-
118 monkey study that investigated global transcriptome changes occurring over 86
119 days of the infection cycle, including initial colonization, acute clinical pharyngitis,
120 and ultimately asymptomatic colonization (15, 18). Taken together, these studies
121 have provided a broad overview of some of the GAS processes at work in the
122 oropharynx; however, much remains to be learned.
123 Especially lacking is a detailed understanding of the genes required for
124 successful growth and persistence in saliva. To address this important
125 knowledge gap, we conducted a genome-wide screen to identify GAS genes
126 contributing to fitness in human saliva *ex vivo*. Using **tra**nsposon-**d**irected
127 **i**nsertion **s**ite sequencing (TraDIS) (19-28), we generated a highly saturated
128 transposon insertion library (140,249 unique transposon insertions) in serotype
129 M1 reference strain MGAS2221 (29-31). This serotype was used because it is
130 usually the most common cause of pharyngitis and other human infections in
131 western countries (3). Strain MGAS2221 is genetically representative of the
132 pandemic clone that arose in the 1980s and rapidly spread globally (29-32). The
133 transposon mutant library was exposed to human saliva *ex vivo* for 48 hrs to
134 identify GAS genes contributing to fitness over time in this clinically relevant fluid.
135 Saliva studies conducted with six isogenic mutant strains validated the findings.
136 The new information we obtained substantially increases our overall
137 understanding of the molecular genetic basis of pathogen-saliva interactions and

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

138 is valuable for future translational research designed to treat or prevent human
139 GAS infections.
140
141
142 **RESULTS**
143
144 **Construction of a highly saturated transposon insertion library in**
145 **serotype M1 GAS strain MGAS2221.** A transposon insertion mutant library was
146 generated using serotype M1 strain MGAS2221 as the parental organism. Strain
147 MGAS2221 was chosen for transposon mutagenesis because (i) it is genetically
148 representative of a pandemic clone that arose in the 1980s and disseminated
149 worldwide (30-32), (ii) MGAS2221 has wild-type alleles of major transcriptional
150 regulators that affect virulence, such as *covR*/*covS*, *ropB*, *mga*, and *rocA*, and
151 (iii) it has been used in many mouse and primate infection studies (31). Using
152 plasmid pGh9:IS*S1* (26), we successfully generated a dense transposon mutant
153 library in strain MGAS2221 containing 140,249 unique transposon insertions. On
154 average, the library contains one transposon insertion every 13 nucleotides.
155 93.4% of the genes in the MGAS2221 genome (1720 out of 1841) have at least
156 one transposon insertion. The nearly random distribution of transposon insertion
157 and high density of transposon insertions in the *S. pyogenes* genome is
158 illustrated in Fig. 1. By analyzing the mutant library using the tradis_essentiality
159 TraDIS toolkit script (25), we identified 432 genes (~23.5% of the genes in the
160 genome) that are essential for the serotype M1 GAS strain MGAS2221 in our

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

161 experimental conditions (40°C, in THY broth supplemented with 0.5 μg/ml
162 erythromycin, see Materials and Methods). The list of identified essential genes
163 is presented in Table S1.

164 **Genes contributing to fitness of GAS over time in human saliva.** We
165 exposed the transposon mutant library to pooled human saliva *ex vivo* to screen
166 for genes contributing to fitness in this fluid. TraDIS was used to identify genes
167 with significantly altered mutant frequency in the output mutant pools compared
168 to the input pool at 12, 24, and 48 hrs after saliva inoculation. Genes with
169 significantly decreased mutant frequency (fold-change >1.5, and q value < 0.1) in
170 the output mutant pools were regarded as important for saliva growth and
171 persistence, which can be referred to as fitness. To ensure that the statistical
172 power was adequate, genes with fewer than 10 transposon insertions in any of
173 the four input mutant pools were excluded from the analysis, as recommended
174 by van Opijnen et al (33).

175 We identified 30 (12 hrs), 42 (24 hrs) and 83 (48 hrs) genes with
176 significantly decreased mutant frequencies, providing evidence that these genes
177 contribute to GAS fitness in human saliva (Fig. 2 and Fig. 3, Table S2). In total,
178 92 genes were identified at the three time points (Fig. 3A, Table S2). Clusters of
179 orthologous groups (COG) classification of the 92 genes showed that
180 numerically, the three more prevalent categories include genes involved in
181 carbohydrate transport and metabolism (n = 10 genes), amino acid transport and
182 metabolism (n = 8 genes), and inorganic ion transport and metabolism (n = 7
183 genes) (Fig. 3B). Our previously published data from an experimental pharyngitis

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

184 infection study, involving 20 cynomolgus macaques (15), identified genes
185 expressed during GAS oropharyngeal infection. Of the 92 saliva-fitness genes
186 identified by TraDIS, ~74% were also expressed during GAS oropharyngeal
187 infection (Table S2). Moreover, many of these 92 genes (e.g., *nagA*, *pstS*, *oppA*,
188 and *malX*) are upregulated during GAS oropharyngeal infection in cynomolgus
189 macaques (15). Together, our results suggest that many of the GAS genes
190 contributing to fitness in saliva *ex vivo* may also contribute to pathogen fitness in
191 the oropharynx of non-human primates (NHPs). However, experiments will be
192 required to directly test this hypothesis.

193 Additionally, we determined that some genes at 12 hr (n = 8 genes), 24 hr
194 (n = 21 genes) and 48 hr (n = 41 genes) were significantly associated with
195 potentially increased GAS fitness in saliva. The magnitude of the fold-change in
196 these genes was relatively modest compared to the genes that were decreased
197 in fitness (Fig. 2A-C, Table S3).

198 The fitness score of all genes in the M1 GAS genome (including the genes
199 with less than 10 insertions) with significant change in sequence read counts at
200 the three time points (12 hr, 24 hr and 48 hr) are listed in Table S4, Table S5,
201 and Table S6, respectively.

202 **Validation of six genes required for wild-type GAS fitness in human**
203 **saliva *ex vivo*.** To validate the TraDIS screen findings, we analyzed the saliva
204 fitness phenotype of each isogenic mutant strain generated from six genes
205 (*spy0644*, s*py0646*, *lacR.1*, *carB*, *nifS1*, and *pstS*) identified in the screen.
206 These six genes were chosen for analysis because (i) they have not been

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

207 previously shown to participate in GAS fitness in human saliva, (ii) transposon
208 insertions into these genes represented a range of altered fitness fold-change
209 values, (iii) the genes are present in the core genome of all sequenced GAS
210 genomes, (iv) the genes are known to be expressed in the oropharynx of NHPs
211 during experimental infection (15), and (v) these genes participate in a variety of
212 biological pathways: *spy0644* and s*py0646* (putative ABC transporter), *lacR.1*
213 (carbohydrate metabolism), *carB* (pyrimidine and arginine synthesis), *nifS1*
214 (amino acid metabolism) and *pstS* (phosphate transport). To test the hypothesis
215 that inactivating each of these six genes impaired GAS fitness in human saliva *ex*
216 *vivo*, we used targeted insertional mutagenesis (31, 34, 35) to create isogenic
217 mutant strains from wild-type parental strain MGAS2221. The genome of each
218 isogenic strain was sequenced before use to ensure that no spurious mutations
219 had been introduced during mutant construction. Consistent with our hypothesis,
220 the results (Fig. 4) confirmed that these six isogenic mutant strains had
221 significantly decreased fitness in human saliva compared to parental strain
222 MGAS2221 (Fig. 4). Importantly, we discovered that mutant strains Δ*spy0644*
223 and Δ*spy0646* had severely impaired fitness in saliva (Fig. 4A). Greater than
224 99% of the Δ*spy0644* and Δ*spy0646* inocula were not present as viable cells at
225 the 24 hr time point post-saliva inoculation (Fig. 4A). Genes *spy0644*, *spy0645*,
226 and *spy0646* likely constitute an operon that encodes an ABC transporter system
227 (Fig. 4A and Table S2). On the basis of genome sequencing of thousands of
228 strains of 20 evolutionarily diverse M protein serotypes commonly causing
229 human infections, these three genes are part of the core genome of GAS (32,

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

230 36). That is, these genes are present in all GAS genomes sequenced and
231 moreover, they are highly conserved in genome location and context, and
232 primary amino acid sequence. In addition, homologs of this three-gene region
233 are present in related species of pathogenic streptococci, including
234 *Streptococcus agalactiae, Streptococcus dysgalactiae*, *Streptococcus equi*,
235 *Streptococcus gallolyticus*, *Streptococcus mutans* and others, and are conserved
236 in location downstream of *carB* (Fig. 5). This suggests a functional relationship
237 exists between this ABC transporter and the metabolic activities of CarB. In this
238 regard, we also identified *carB* to be important for wild-type GAS fitness in saliva
239 (Table S2), and confirmed that the Δ*carB* isogenic mutant strain is significantly
240 impaired in persistence in saliva (Fig. 4B). *carB* encodes the large subunit of
241 carbamoylphosphate synthetase (37, 38). Carbamoylphosphate is a precursor for
242 both pyrimidine and arginine synthesis (38). Interestingly, *carB* also was reported
243 to be required for GAS fitness in human blood, a body fluid with a very different
244 chemical composition to saliva (39). These results indicate that pyrimidine and
245 arginine synthesis mediated in part by *carB* contributes to GAS fitness in multiple
246 host niches. The Δ*pstS* mutant strain is also significantly attenuated for growth in
247 human saliva (Fig. 4E). *pstS* encodes a putative phosphate binding protein and is
248 part of a six-gene operon encoding a phosphate uptake system (Fig. 4E). A
249 genome-wide transposon mutagenesis screen found that *pstS* is required for the
250 fitness of *Streptococcus pneumoniae* in human saliva *ex vivo* (40). We also note
251 that *pst* operon genes have been reported to contribute to the virulence of
252 multiple gram-negative pathogens, including *Proteus mirabilis* and *Escherichia*

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

253 *coli* (41-45). These results suggest efficient phosphate uptake is important for
254 fitness of multiple human and animal pathogens that must successfully interact
255 with their specific host niches, including saliva in the oropharynx. Growth of the
256 six mutants in rich medium *in vitro* (THY) showed that none of the mutants has
257 severe growth defects (Fig. 4F). To summarize, the saliva growth phenotype of
258 these six isogenic mutant strains strongly reflected the fitness data obtained from
259 the high-throughput TraDIS screen.
260
261
262 **DISCUSSION**
263
264 Our results present for the first time a genome-wide view of the GAS genes
265 contributing to pathogen fitness in human saliva, only the second pathogen (40)
266 for which such a screen has been performed. The work also represents the first
267 application of TraDIS to GAS, a human-specific pathogen responsible for greater
268 than 700 million infections each year, including 600 million pharyngitis cases (3).
269 Some years ago we initiated study of GAS–human saliva molecular
270 interactions (16) with the goal of obtaining new insight into GAS gene activity
271 during the earliest stage of oropharyngeal infection. A common theme that has
272 emerged from many studies (16, 17, 46, 47) is that genes involved in complex
273 carbohydrate catabolism play important roles in growth and persistence in human
274 saliva. Saliva in the human oropharynx contains many nutrients and diverse
275 molecules critical to innate and acquired immunity (48-50). Knowledge gained

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

276 about how GAS responds to saliva contributes to a broader understanding of
277 host–pathogen interaction and microbial persistence on mucosal surfaces.
278 Expression microarray analysis, immunologic methods, and *in vivo* gene
279 quantification identified a genetic program used by GAS to survive in human
280 saliva *ex vivo* (14, 16, 17). A key discovery was that a two-component regulatory
281 system (TCS) of previously unknown function played a central role in pathogen
282 survival in saliva (17), and revealed an intimate link between metabolism,
283 virulence factor production, and GAS persistence in saliva. However, the
284 strategy used was unable to directly identify the specific genes contributing to
285 GAS persistence in saliva. Our TraDIS analysis discovered that 25 of 92 (27%)
286 genes contributing to fitness in human saliva *ex vivo* are involved in
287 carbohydrate, amino acid and inorganic ion transport and/or metabolism.
288 Inactivation of genes in these categories is likely to significantly impair core
289 metabolic processes such as nutrient acquisition and use. Our results add to the
290 important theme of an intimate linkage between metabolism and GAS
291 persistence in human saliva. We note that a similar spectrum of genes was
292 abundantly represented in an *ex vivo* human saliva transposon mutagenesis
293 screen conducted for *S. pneumoniae* (40). For example, the *opp* operon (oligo
294 peptide transport) and *pst* operon (phosphate uptake) are required for fitness in
295 both GAS and *S. pneumoniae* (40) (Table S2), suggesting certain mechanisms
296 contributing to bacterial fitness in saliva are shared by multiple pathogens in the
297 oropharyngeal niche.

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

298 By mining data available from our previously conducted NHP pharyngitis
299 study, we found that a large majority (74%, see Table S2) of the 92 genes
300 discovered herein to contribute to fitness in human saliva also were expressed *in*
301 *vivo* in the monkey oropharynx (15). Several explanations may account for the
302 lack of evidence for expression of 24% of the genes. First, the NHP study was
303 conducted relatively early in our understanding of the annotation of the genome
304 of the input serotype M1 strain MGAS5005. It is possible that not all of the
305 genome of MGAS5005 was represented on the Affymetrix gene chip used in that
306 study. Second, it is possible that some of the genes were expressed, but at
307 levels too low to detect with the techniques available at that time. Third, it may be
308 that some of the genes in MGAS2221 are expressed at different time points than
309 we used in the two studies, for example farther into the asymptomatic carriage
310 phase (that is, later than day 7), the end point used in our current study. Finally,
311 human saliva *ex vivo* is a different environment than the primate oropharynx, a
312 niche that also contains, for example, host innate and acquired immune cells and
313 epithelial cells. Therefore, it was not unexpected to find potential differences in
314 evidence of gene expression between the two data sets. Despite this, the
315 remarkable 74% gene overlap unambiguously shows that many similar
316 mechanisms are at work on GAS regardless of whether it is exposed to human
317 saliva *ex vivo* or inoculated into the primate oropharynx.

318 For decades, the analysis of GAS-mediated processes contributing to
319 pharyngitis was restricted predominantly to inferences obtained by evaluating
320 serologic responses to relatively few extracellular molecules that participate in

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

321 pathogen-host molecular interactions, such as M protein, DNase B, and
322 streptolysin O (51-53). Although important information has been obtained from
323 these descriptive studies, the inability to directly identify large numbers of GAS
324 genes contributing to pharyngitis means that we have a very imprecise
325 understanding of molecular processes contributing to this important human
326 infection. Given the 74% overlap in genes between our TraDIS screen and NHP
327 pharyngitis expression data, it is reasonable to conclude that our study advances
328 understanding of molecular events occurring in the oropharynx, and thereby
329 provides a critical foundation for subsequent molecular pathogenesis studies.

330 Growth and persistence in human saliva contributes to the ability of GAS
331 to be successfully transmitted by respiratory droplets. The work of Hamburger
332 (9-11) demonstrated that individuals with higher CFUs of GAS in saliva were
333 more likely to transmit the organism to others. This observation implies that any
334 process detrimentally affecting fitness in saliva (such as gene inactivation as
335 done herein) is likely to decrease the probability of successful transmission. It
336 also suggests that mutants with substantial fitness defects are more highly likely
337 to have decreased ability to disseminate and/or cause clinical pharyngitis. The
338 isogenic mutants constructed by inactivation of *spy0644* (herein denoted *sptA*,
339 for streptococcal persistence) and *spy0646* (*sptC*) had the most pronounced
340 growth phenotype of the six isogenic mutant strains during saliva growth *ex vivo*.
341 Of note, the transposon mutants of these two genes had the lowest fitness
342 scores at the early time point (12 hr) of any of the six mutants we tested (Table
343 S2). In contrast, according to the screen, *carB* and *lacR.1* transposon mutants

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

344 have significantly decreased fitness only at the latest time point (48 hr) (Table
345 S2). Isogenic mutants Δ*carB* and Δ*lacR.1* had a moderate saliva growth
346 phenotype compared to the Δ*sptA* and Δ*sptC* strains (Fig. 4B and C). These
347 findings suggest that a relationship exists between magnitude of the fold-change
348 at an early time point and growth of the resulting isogenic mutant strain in saliva
349 *ex vivo*. However, more data generated with additional isogenic mutant strains is
350 required to rigorously test this idea. If true, use of these data could be an
351 important characteristic used to help triage GAS genes for more in depth
352 analysis, including translation research activities.

353 Our results show that mutants with insertions in *pst* operon genes are
354 significantly attenuated for growth in human saliva (Fig. 4E, Table S2). The *pst*
355 operon encodes a high-affinity phosphate transporter and has been reported to
356 contribute to bacterial virulence and fitness in a wide variety of pathogens (41,
357 43-45, 54-56). For example, genome-wide transposon mutagenesis screens
358 found that *pst* operon is required for the fitness of *S. pneumoniae* in human
359 saliva *ex vivo* (40) and mouse lung infections (57). In the oral pathogen *S.*
360 *mutans*, deleting *pstS* results in decreased production of extracellular
361 polysaccharide, and reduced ability to adhere to saliva-coated surfaces (56).
362 The *pst* operon genes in gram-negative pathogens, have been reported to
363 contribute to the virulence of *Proteus mirabilis* and *Escherichia coli* (41-45, 54).
364 Together, these results suggest efficient phosphate uptake is important for the
365 ability of multiple human and animal pathogens to survive and thrive in their
366 specific host niches, including *S. pyogenes* in human saliva.

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

367 The transposon mutagenesis screen study we performed has several
368 limitations. In the human oropharynx, saliva is constantly replenished, whereas
369 in our experimental system a single aliquot of pooled saliva was used for the
370 entire incubation period. A second important limitation is the lack of intact
371 immune cells contributing to innate and acquired immunity in the saliva
372 preparation used. These and other cells (e.g., epithelial) are lacking because the
373 saliva is filter-sterilized prior to use. It is also possible that very small genes were
374 missed in the screen because we excluded genes with fewer than 10 inserts from
375 the analysis, a common practice in transposon mutagenesis studies (33, 58).
376 Overcoming some of these limitations will require experimental infection of an
377 intact animal that faithfully recapitulates all phases of human pharyngitis, such as
378 NHPs. A third limitation is the six isogenic mutants we used to validate the
379 TraDIS results were generated by insertional inactivation. Although the isogenic
380 mutants have no spurious mutations, and their phenotype recapitulated the
381 TraDIS findings, this does not rule out potential polar effects on neighboring
382 genes, especially genes located in the same operon. To overcome this limitation
383 in future follow-up in-depth functional studies on key genes and operons
384 identified in this initial screen, non-polar deletion mutants and complemented
385 mutant strains can be used.

386 To summarize, by identifying genes contributing to fitness in human saliva
387 *ex vivo*, our work complements important information obtained in other
388 transposon mutant screens that identified GAS genes contributing to fitness in
389 human blood *ex vivo* and mouse soft tissue disease after subcutaneous infection,

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

390 and for growth *in vitro* (39, 59, 60). It is reasonable to suggest that the
391 information we obtained from this genome-wide screen for genes contributing to
392 fitness in human saliva can be successfully exploited for future pathogenesis
393 investigations.
394
395
396 **MATERIALS AND METHODS**

397 **Bacterial strains and growth conditions.** Strain MGAS2221 is
398 genetically representative of a pandemic clone of serotype M1 that arose in the
399 1980s and has spread worldwide (30-32). Isogenic mutant strains Δ*carB*,
400 Δ*lacR.1*, Δ*sptA*, Δ*sptC*, Δ*nifS1*, Δ*pstS* were derived from parental strain
401 MGAS2221, the organism used for construction of the transposon mutant library.
402 All GAS strains were grown in Todd-Hewitt broth supplemented with 0.2% yeast
403 extract (THY broth) at 37°C with 5% CO_2.

404 **Generation of GAS transposon mutant library.** A transposon mutant
405 library was generated in strain MGAS2221 using transposon plasmid pGh9:IS*S1*
406 based on a recently described protocol (26). Briefly, pGh9:IS*S1* was transformed
407 into strain MGAS2221 by electroporation (26). A single colony of the
408 transformants was picked and grown overnight at 28°C (permissive temperature)
409 in THY broth supplemented with 0.5 µg/ml erythromycin. The resulting overnight
410 culture was heat shocked at 40°C (nonpermissive temperature) for 3 hrs to
411 permit random transposition and integration of pGh9:IS*S1* into the GAS genome.
412 The GAS cells in the culture were harvested by centrifugation, plated on THY
413 agar supplemented with 0.5 µg/ml erythromycin, and grown overnight at 37°C.

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

414 The transposon mutant library (i.e., pooled transposon mutants) was collected by
415 washing the colonies off the agar plates with THY broth containing 25% glycerol.
416 The bacterial cell suspension (transposon mutant library) was stored at -80°C.
417 This process was repeated on three separate occasions and the three libraries
418 were pooled.

419 **Human saliva collection and processing.** Human saliva was collected
420 and processed as described previously (16, 17). Briefly, saliva was collected
421 from five healthy donors, pooled, clarified by centrifugation at 45,000 x g for 15
422 min and sterilized with a 0.20 μm filter (Corning Inc.). The resulting sterile saliva
423 was used for subsequent transposon mutant library screening and individual
424 strain growth.

425 **Exposure of the transposon mutant library to human saliva *ex vivo*.**
426 The transposon mutant library was inoculated into 20 ml of THY and grown at
427 37°C to mid-exponential phase (OD = 0.5). The bacteria were harvested by
428 centrifugation and washed three times with an equal volume of PBS to remove
429 trace THY broth. 50 μl of the cell suspension in PBS was inoculated into four
430 tubes, each containing 40 ml of filtered saliva (Fig. S1). The four inoculated tubes
431 constitute four biological replicates of the saliva persistence assay (Fig. S1).
432 Immediately after inoculation, 200 μl of the inoculated saliva from each of the four
433 replicate cultures was plated onto four THY plates, and incubated at 37°C for 12
434 hrs. GAS cells growing on the plates were harvested and represent the
435 composition of the input mutant pools (0 hr mutant pools, n = 4). Mutant pools
436 present at 12 hrs, 24 hrs and 48 hrs post-inoculation were recovered by plating

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

437 200 µl of the inoculated saliva onto THY agar plates at the aforementioned time
438 points and incubated for 12 hrs at 37°C. The collected mutant pools (4 replicates,
439 4 time points, *n* = 16) were stored at -80°C for the subsequent TraDIS analysis.

440 **DNA preparation and massive parallel sequencing.** Genomic DNA
441 preparation and DNA sequencing was performed according to procedures
442 described previously for TraDIS analysis, with minor modifications (26). Briefly,
443 genomic DNA of mutant pools collected at the various time points was isolated
444 using DNeasy blood & tissue kit (Qiagen). 2 µg purified genomic DNA was
445 fragmented by incubating with NEBNext dsDNA Fragmentase (New England
446 Biolabs) for 25 min at 37 °C to obtain DNA fragments in the range of 200-1,000
447 bp. A Y-adaptor (26) (Table S7) was ligated to 1 µg of fragmented DNA using the
448 NEBNext Ultra II DNA library prep kit for Illumina (New England Biolabs). The
449 adaptor-ligated DNA fragments were then purified using AMPure XP beads
450 (Agencourt, Beckman Coulter) and digested with restriction enzyme BamHI for 3
451 hrs at 37 °C to minimize the mapping of TraDIS reads to the transposon plasmid
452 backbone (26). The resulting DNA was purified, and 100 ng of the purified library
453 DNA was subjected to PCR using the specific IS*S1* primer and one of the 8
454 indexing PCR primers per DNA library (Table S7), to amplify regions that span
455 the 5' end of IS*S1* and the GAS genomic regions adjacent to the chromosomal
456 location of the transposon. The PCR amplified libraries were sequenced using a
457 single end 76-cycle protocol on a NextSeq550 instrument (Illumina) using a
458 custom Read 1 primer (Table S7) and a custom Index Read sequencing primer
459 (Table S7).

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

460 **Processing of DNA sequencing reads and data analysis.** The raw
461 Illumina reads obtained from the input and output pools were parsed with FASTX
462 Barcode Splitter (http://hannonlab.cshl.edu/fastx_toolkit/commandline.html). After
463 removing adaptor, low quality reads, and index sequences, PRINSEQ lite version
464 0.20.4 (http://prinseq.sourceforge.net/) was used to eliminate reads shorter than
465 nucleotides. The resulting sequencing reads were analyzed with the TraDIS
466 toolkit (25) according to previously described methods (26, 61). Briefly,
467 bacteria_tradis was used to trim transposon tag sequence and map the
468 remaining reads to the serotype M1 strain MGAS5005 reference genome. The
469 plot files generated by bacteria_tradis were analyzed by tradis_gene_insert_sites
470 to generate spreadsheets listing read count, insertion count, and insertion index
471 for each gene. The output files from the tradis_gene_insert_sites analysis were
472 transferred to tradis_comparison.R to compare the reads mapped per gene
473 between the input pools (T0 pools) and the output pools (T12, T24, and T48
474 pools). Essential genes were determined by analyzing the input library using the
475 tradis_essentiality TraDIS toolkit script (25).
476 **Construction of isogenic mutant strains.** Construction of isogenic
477 mutant strains was performed by previously described methods (31, 34, 35). An
478 internal fragment from six different genes identified as important for fitness in
479 saliva (*spy0644*, *spy0646*, *carB*, *lacR1*, *nifS1*, and *pstS*) was amplified by PCR
480 from genomic DNA from wild-type parental M1 strain MGAS2221 with relevant
481 primers (Table S7), digested with *Bam*HI, and cloned into suicide vector
482 pBBL740 (62). The resulting constructs were transformed into strain MGAS2221

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

483 to inactivate each of the six genes. pBBL740 has a *cat* gene which confers
484 chloramphenicol resistance. The plasmid integrant strains (mutants) were
485 selected using THY agar plates supplemented with 5 μg/ml chloramphenicol. The
486 genome of each isogenic strain was sequenced before use to confirm the mutant
487 construct and ensure that no spurious mutations were introduced during mutant
488 construction.

489 **Saliva growth assay.** To compare the ability of GAS strains to grow in
490 human saliva, we first grew GAS strains to mid-exponential phase in THY broth
491 (OD = 0.5). GAS cells were washed three times with PBS, and suspended with
492 the equivalent volume of PBS. 100ul of the GAS-PBS suspension was inoculated
493 into 10 ml aliquots of filter-sterilized human saliva. GAS strains in saliva were
494 incubated at 37° C with 5 % CO_2. Samples (100 ml) of the GAS-saliva
495 suspension were recovered over time, and GAS CFU was determined by serial
496 dilution and growing on blood agar plates. Four biological replicates were
497 included for each strain. Statistical significance were assessed by repeated-
498 measures 2-way ANOVA.
499
500

501 **ACKNOWLEDGMENTS**

502 This work was supported by funds from the Fondren Foundation (to
503 James M. Musser), the PetPlan Charitable Trust (ref: S14-51) and the Horse
504 Trust (ref: G4104) (to Andrew S. Waller). Amelia R. L. Charbonneau is supported
505 by the University of Cambridge Doctoral Training Partnership scheme, which is

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

506 funded by the Biotechnology and Biological Sciences Research Council, UK (ref:
507 1503883). We thank Kathryn Stockbauer for critical reading of the manuscript.

508

509

510 **REFERENCES**

511

512 1. Baumler A, Fang FC. 2013. Host specificity of bacterial pathogens. Cold Spring
513 Harb Perspect Med 3:a010041.
514 2. Monack DM, Mueller A, Falkow S. 2004. Persistent bacterial infections: the
515 interface of the pathogen and the host immune system. Nat Rev Microbiol
516 2:747-65.
517 3. Carapetis JR, Steer AC, Mulholland EK, Weber M. 2005. The global burden of
518 group A streptococcal diseases. Lancet Infect Dis 5:685-94.
519 4. Peter G, Smith AL. 1977. Group A streptococcal infections of the skin and
520 pharynx (second of two parts). N Engl J Med 297:365-70.
521 5. Bisno AL. 2001. Acute pharyngitis. N Engl J Med 344:205-11.
522 6. Eisenberg MJ. 1993. Rheumatic heart disease in the developing world:
523 prevalence, prevention, and control. Eur Heart J 14:122-8.
524 7. Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A,
525 Sriprakash KS, Sanderson-Smith ML, Nizet V. 2014. Disease manifestations
526 and pathogenic mechanisms of Group A Streptococcus. Clin Microbiol Rev
527 27:264-301.
528 8. Wilkening RV, Federle MJ. 2017. Evolutionary Constraints Shaping
529 Streptococcus pyogenes-Host Interactions. Trends Microbiol 25:562-572.
530 9. Hamburger M. 1944. Studies on the transmission of hemolytic streptococcus
531 infections: I. Cross infections in army hospital wards. J Infect Dis 75:58-70.
532 10. Hamburger M. 1944. Studies on the transmission of hemolytic streptococcus
533 infections: II. Beta hemolytic streptococci in the saliva of persons with
534 positive throat cultures. J Infect Dis 75:71-78.
535 11. Hamburger M, Robertson O. 1948. Expulsion of group A hemolytic
536 streptococci in droplets and droplet nuclei by sneezing, coughing, and
537 talking. Am J Med 4:690-701.
538 12. Kaplan EL, Couser R, Huwe BB, McKay C, Wannamaker LW. 1979.
539 Significance of quantitative salivary cultures for group A and non-group A
540 and non-group A beta-hemolytic streptococci in patients with pharyngitis
541 and in their family contacts. Pediatrics 64:904-12.
542 13. Sitkiewicz I, Musser JM. 2006. Expression microarray and mouse virulence
543 analysis of four conserved two-component gene regulatory systems in group
544 a streptococcus. Infect Immun 74:1339-51.
545 14. Virtaneva K, Graham MR, Porcella SF, Hoe NP, Su H, Graviss EA, Gardner TJ,
546 Allison JE, Lemon WJ, Bailey JR, Parnell MJ, Musser JM. 2003. Group A

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

547 Streptococcus gene expression in humans and cynomolgus macaques with
548 acute pharyngitis. Infect Immun 71:2199-207.
549 15. Virtaneva K, Porcella SF, Graham MR, Ireland RM, Johnson CA, Ricklefs SM,
550 Babar I, Parkins LD, Romero RA, Corn GJ, Gardner DJ, Bailey JR, Parnell MJ,
551 Musser JM. 2005. Longitudinal analysis of the group A Streptococcus
552 transcriptome in experimental pharyngitis in cynomolgus macaques. Proc
553 Natl Acad Sci U S A 102:9014-9.
554 16. Shelburne SA, 3rd, Granville C, Tokuyama M, Sitkiewicz I, Patel P, Musser JM.
555 2005. Growth characteristics of and virulence factor production by group A
556 Streptococcus during cultivation in human saliva. Infect Immun 73:4723-31.
557 17. Shelburne SA, 3rd, Sumby P, Sitkiewicz I, Granville C, DeLeo FR, Musser JM.
558 2005. Central role of a bacterial two-component gene regulatory system of
559 previously unknown function in pathogen persistence in human saliva. Proc
560 Natl Acad Sci U S A 102:16037-42.
561 18. Shea PR, Virtaneva K, Kupko JJ, 3rd, Porcella SF, Barry WT, Wright FA,
562 Kobayashi SD, Carmody A, Ireland RM, Sturdevant DE, Ricklefs SM, Babar I,
563 Johnson CA, Graham MR, Gardner DJ, Bailey JR, Parnell MJ, Deleo FR, Musser
564 JM. 2010. Interactome analysis of longitudinal pharyngeal infection of
565 cynomolgus macaques by group A Streptococcus. Proc Natl Acad Sci U S A
566 107:4693-8.
567 19. Langridge GC, Phan MD, Turner DJ, Perkins TT, Parts L, Haase J, Charles I,
568 Maskell DJ, Peters SE, Dougan G, Wain J, Parkhill J, Turner AK. 2009.
569 Simultaneous assay of every Salmonella Typhi gene using one million
570 transposon mutants. Genome Res 19:2308-16.
571 20. Chaudhuri RR, Morgan E, Peters SE, Pleasance SJ, Hudson DL, Davies HM,
572 Wang J, van Diemen PM, Buckley AM, Bowen AJ, Pullinger GD, Turner DJ,
573 Langridge GC, Turner AK, Parkhill J, Charles IG, Maskell DJ, Stevens MP. 2013.
574 Comprehensive assignment of roles for Salmonella typhimurium genes in
575 intestinal colonization of food-producing animals. PLoS Genet 9:e1003456.
576 21. Weerdenburg EM, Abdallah AM, Rangkuti F, Abd El Ghany M, Otto TD,
577 Adroub SA, Molenaar D, Ummels R, Ter Veen K, van Stempvoort G, van der
578 Sar AM, Ali S, Langridge GC, Thomson NR, Pain A, Bitter W. 2015. Genome-
579 wide transposon mutagenesis indicates that Mycobacterium marinum
580 customizes its virulence mechanisms for survival and replication in different
581 hosts. Infect Immun 83:1778-88.
582 22. Moule MG, Spink N, Willcocks S, Lim J, Guerra-Assuncao JA, Cia F, Champion
583 OL, Senior NJ, Atkins HS, Clark T, Bancroft GJ, Cuccui J, Wren BW. 2015.
584 Characterization of New Virulence Factors Involved in the Intracellular
585 Growth and Survival of Burkholderia pseudomallei. Infect Immun 84:701-10.
586 23. Grant AJ, Oshota O, Chaudhuri RR, Mayho M, Peters SE, Clare S, Maskell DJ,
587 Mastroeni P. 2016. Genes Required for the Fitness of Salmonella enterica
588 Serovar Typhimurium during Infection of Immunodeficient gp91-/- phox
589 Mice. Infect Immun 84:989-97.
590 24. Subashchandrabose S, Smith S, DeOrnellas V, Crepin S, Kole M, Zahdeh C,
591 Mobley HL. 2016. Acinetobacter baumannii Genes Required for Bacterial
592 Survival during Bloodstream Infection. mSphere 1.

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

593 25. Barquist L, Mayho M, Cummins C, Cain AK, Boinett CJ, Page AJ, Langridge GC,
594 Quail MA, Keane JA, Parkhill J. 2016. The TraDIS toolkit: sequencing and
595 analysis for dense transposon mutant libraries. Bioinformatics 32:1109-11.
596 26. Charbonneau ARL, Forman OP, Cain AK, Newland G, Robinson C, Boursnell M,
597 Parkhill J, Leigh JA, Maskell DJ, Waller AS. 2017. Defining the ABC of gene
598 essentiality in streptococci. BMC Genomics 18:426.
599 27. Ruiz L, Bottacini F, Boinett CJ, Cain AK, O'Connell-Motherway M, Lawley TD,
600 van Sinderen D. 2017. The essential genomic landscape of the commensal
601 Bifidobacterium breve UCC2003. Sci Rep 7:5648.
602 28. Senior NJ, Sasidharan K, Saint RJ, Scott AE, Sarkar-Tyson M, Ireland PM,
603 Bullifent HL, Rong Yang Z, Moore K, Oyston PCF, Atkins TP, Atkins HS, Soyer
604 OS, Titball RW. 2017. An integrated computational-experimental approach
605 reveals Yersinia pestis genes essential across a narrow or a broad range of
606 environmental conditions. BMC Microbiol 17:163.
607 29. Musser JM, Kapur V, Szeto J, Pan X, Swanson DS, Martin DR. 1995. Genetic
608 diversity and relationships among Streptococcus pyogenes strains
609 expressing serotype M1 protein: recent intercontinental spread of a subclone
610 causing episodes of invasive disease. Infect Immun 63:994-1003.
611 30. Sumby P, Porcella SF, Madrigal AG, Barbian KD, Virtaneva K, Ricklefs SM,
612 Sturdevant DE, Graham MR, Vuopio-Varkila J, Hoe NP, Musser JM. 2005.
613 Evolutionary origin and emergence of a highly successful clone of serotype
614 M1 group a Streptococcus involved multiple horizontal gene transfer events.
615 J Infect Dis 192:771-82.
616 31. Zhu L, Olsen RJ, Nasser W, Beres SB, Vuopio J, Kristinsson KG, Gottfredsson
617 M, Porter AR, DeLeo FR, Musser JM. 2015. A molecular trigger for
618 intercontinental epidemics of group A Streptococcus. J Clin Invest 125:3545-
619 59.
620 32. Nasser W, Beres SB, Olsen RJ, Dean MA, Rice KA, Long SW, Kristinsson KG,
621 Gottfredsson M, Vuopio J, Raisanen K, Caugant DA, Steinbakk M, Low DE,
622 McGeer A, Darenberg J, Henriques-Normark B, Van Beneden CA, Hoffmann S,
623 Musser JM. 2014. Evolutionary pathway to increased virulence and epidemic
624 group A Streptococcus disease derived from 3,615 genome sequences. Proc
625 Natl Acad Sci U S A 111:E1768-76.
626 33. van Opijnen T, Bodi KL, Camilli A. 2009. Tn-seq: high-throughput parallel
627 sequencing for fitness and genetic interaction studies in microorganisms. Nat
628 Methods 6:767-72.
629 34. Zhu L, Olsen RJ, Lee JD, Porter AR, DeLeo FR, Musser JM. 2016. Contribution
630 of Secreted NADase and Streptolysin O to the Pathogenesis of Epidemic
631 Serotype M1 Streptococcus pyogenes Infections. Am J Pathol
632 doi:10.1016/j.ajpath.2016.11.003.
633 35. Zhu L, Olsen RJ, Horstmann N, Shelburne SA, Fan J, Hu Y, Musser JM. 2016.
634 Intergenic Variable-Number Tandem-Repeat Polymorphism Upstream of
635 rocA Alters Toxin Production and Enhances Virulence in Streptococcus
636 pyogenes. Infect Immun 84:2086-93.
637 36. Beres SB, Kachroo P, Nasser W, Olsen RJ, Zhu L, Flores AR, de la Riva I, Paez-
638 Mayorga J, Jimenez FE, Cantu C, Vuopio J, Jalava J, Kristinsson KG,

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

639 Gottfredsson M, Corander J, Fittipaldi N, Di Luca MC, Petrelli D, Vitali LA,
640 Raiford A, Jenkins L, Musser JM. 2016. Transcriptome Remodeling
641 Contributes to Epidemic Disease Caused by the Human Pathogen
642 Streptococcus pyogenes. MBio 7.
643 37. Nyunoya H, Lusty CJ. 1983. The carB gene of Escherichia coli: a duplicated
644 gene coding for the large subunit of carbamoyl-phosphate synthetase. Proc
645 Natl Acad Sci U S A 80:4629-33.
646 38. Arioli S, Monnet C, Guglielmetti S, Mora D. 2009. Carbamoylphosphate
647 synthetase activity is essential for the optimal growth of Streptococcus
648 thermophilus in milk. J Appl Microbiol 107:348-54.
649 39. Le Breton Y, Mistry P, Valdes KM, Quigley J, Kumar N, Tettelin H, McIver KS.
650 2013. Genome-wide identification of genes required for fitness of group A
651 Streptococcus in human blood. Infect Immun 81:862-75.
652 40. Verhagen LM, de Jonge MI, Burghout P, Schraa K, Spagnuolo L, Mennens S,
653 Eleveld MJ, van der Gaast-de Jongh CE, Zomer A, Hermans PW, Bootsma HJ.
654 2014. Genome-wide identification of genes essential for the survival of
655 Streptococcus pneumoniae in human saliva. PLoS One 9:e89541.
656 41. Jacobsen SM, Lane MC, Harro JM, Shirtliff ME, Mobley HL. 2008. The high-
657 affinity phosphate transporter Pst is a virulence factor for Proteus mirabilis
658 during complicated urinary tract infection. FEMS Immunol Med Microbiol
659 52:180-93.
660 42. Armbruster CE, Forsyth-DeOrnellas V, Johnson AO, Smith SN, Zhao L, Wu W,
661 Mobley HLT. 2017. Genome-wide transposon mutagenesis of Proteus
662 mirabilis: Essential genes, fitness factors for catheter-associated urinary tract
663 infection, and the impact of polymicrobial infection on fitness requirements.
664 PLoS Pathog 13:e1006434.
665 43. Daigle F, Fairbrother JM, Harel J. 1995. Identification of a mutation in the pst-
666 phoU operon that reduces pathogenicity of an Escherichia coli strain causing
667 septicemia in pigs. Infect Immun 63:4924-7.
668 44. Cheng C, Tennant SM, Azzopardi KI, Bennett-Wood V, Hartland EL, Robins-
669 Browne RM, Tauschek M. 2009. Contribution of the pst-phoU operon to cell
670 adherence by atypical enteropathogenic Escherichia coli and virulence of
671 Citrobacter rodentium. Infect Immun 77:1936-44.
672 45. Lamarche MG, Dozois CM, Daigle F, Caza M, Curtiss R, 3rd, Dubreuil JD, Harel
673 J. 2005. Inactivation of the pst system reduces the virulence of an avian
674 pathogenic Escherichia coli O78 strain. Infect Immun 73:4138-45.
675 46. Shelburne SA, 3rd, Sumby P, Sitkiewicz I, Okorafor N, Granville C, Patel P,
676 Voyich J, Hull R, DeLeo FR, Musser JM. 2006. Maltodextrin utilization plays a
677 key role in the ability of group A Streptococcus to colonize the oropharynx.
678 Infect Immun 74:4605-14.
679 47. Shelburne SA, 3rd, Sahasrobhajane P, Suber B, Keith DB, Davenport MT,
680 Horstmann N, Kumaraswami M, Olsen RJ, Brennan RG, Musser JM. 2011.
681 Niche-specific contribution to streptococcal virulence of a MalR-regulated
682 carbohydrate binding protein. Mol Microbiol 81:500-14.
683 48. Wiesner J, Vilcinskas A. 2010. Antimicrobial peptides: the ancient arm of the
684 human immune system. Virulence 1:440-64.

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

685 49. Marcotte H, Lavoie MC. 1998. Oral microbial ecology and the role of salivary
686 immunoglobulin A. Microbiol Mol Biol Rev 62:71-109.
687 50. Humphrey SP, Williamson RT. 2001. A review of saliva: normal composition,
688 flow, and function. J Prosthet Dent 85:162-9.
689 51. Mc CM. 1949. The inhibition of streptococcal desoxyribonuclease by rabbit
690 and human antisera. J Exp Med 90:543-53.
691 52. Kaplan EL, Rothermel CD, Johnson DR. 1998. Antistreptolysin O and anti-
692 deoxyribonuclease B titers: normal values for children ages 2 to 12 in the
693 United States. Pediatrics 101:86-8.
694 53. Hysmith ND, Kaplan EL, Cleary PP, Johnson DR, Penfound TA, Dale JB. 2017.
695 Prospective Longitudinal Analysis of Immune Responses in Pediatric
696 Subjects After Pharyngeal Acquisition of Group A Streptococci. J Pediatric
697 Infect Dis Soc 6:187-196.
698 54. O'May GA, Jacobsen SM, Longwell M, Stoodley P, Mobley HL, Shirtliff ME.
699 2009. The high-affinity phosphate transporter Pst in Proteus mirabilis
700 HI4320 and its importance in biofilm formation. Microbiology 155:1523-35.
701 55. Lamarche MG, Wanner BL, Crepin S, Harel J. 2008. The phosphate regulon
702 and bacterial virulence: a regulatory network connecting phosphate
703 homeostasis and pathogenesis. FEMS Microbiol Rev 32:461-73.
704 56. Luz DE, Nepomuceno RS, Spira B, Ferreira RC. 2012. The Pst system of
705 Streptococcus mutans is important for phosphate transport and adhesion to
706 abiotic surfaces. Mol Oral Microbiol 27:172-81.
707 57. Hava DL, Camilli A. 2002. Large-scale identification of serotype 4
708 Streptococcus pneumoniae virulence factors. Mol Microbiol 45:1389-406.
709 58. van Opijnen T, Camilli A. 2013. Transposon insertion sequencing: a new tool
710 for systems-level analysis of microorganisms. Nat Rev Microbiol 11:435-42.
711 59. Le Breton Y, Belew AT, Valdes KM, Islam E, Curry P, Tettelin H, Shirtliff ME,
712 El-Sayed NM, McIver KS. 2015. Essential Genes in the Core Genome of the
713 Human Pathogen Streptococcus pyogenes. Sci Rep 5:9838.
714 60. Le Breton Y, Belew AT, Freiberg JA, Sundar GS, Islam E, Lieberman J, Shirtliff
715 ME, Tettelin H, El-Sayed NM, McIver KS. 2017. Genome-wide discovery of
716 novel M1T1 group A streptococcal determinants important for fitness and
717 virulence during soft-tissue infection. PLoS Pathog 13:e1006584.
718 61. Dembek M, Barquist L, Boinett CJ, Cain AK, Mayho M, Lawley TD, Fairweather
719 NF, Fagan RP. 2015. High-throughput analysis of gene essentiality and
720 sporulation in Clostridium difficile. MBio 6:e02383.
721 62. Ramirez-Pena E, Trevino J, Liu Z, Perez N, Sumby P. 2010. The group A
722 Streptococcus small regulatory RNA FasX enhances streptokinase activity by
723 increasing the stability of the ska mRNA transcript. Mol Microbiol 78:1332-
724 47.
725

726

727

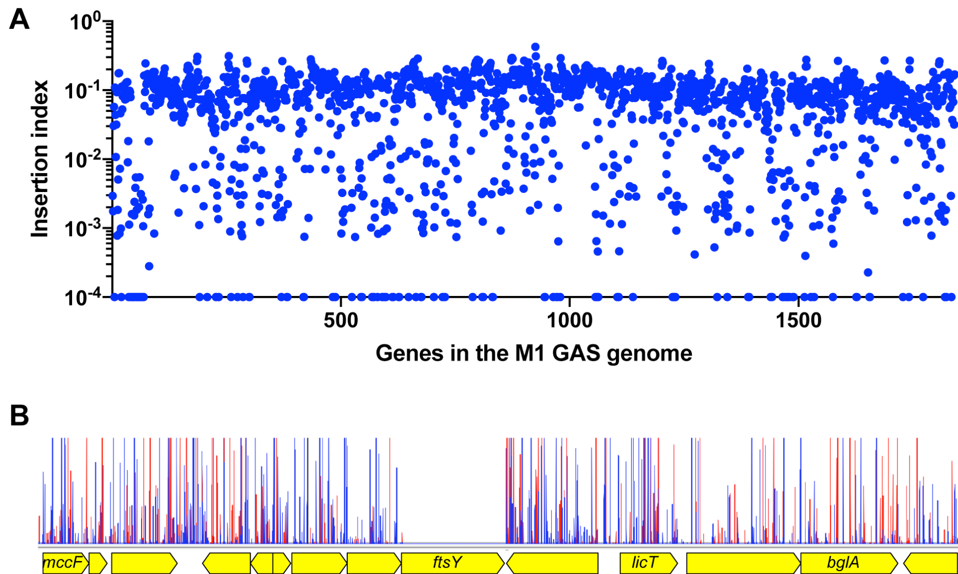
bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

728

729

730

731 **Figures and figure legends**

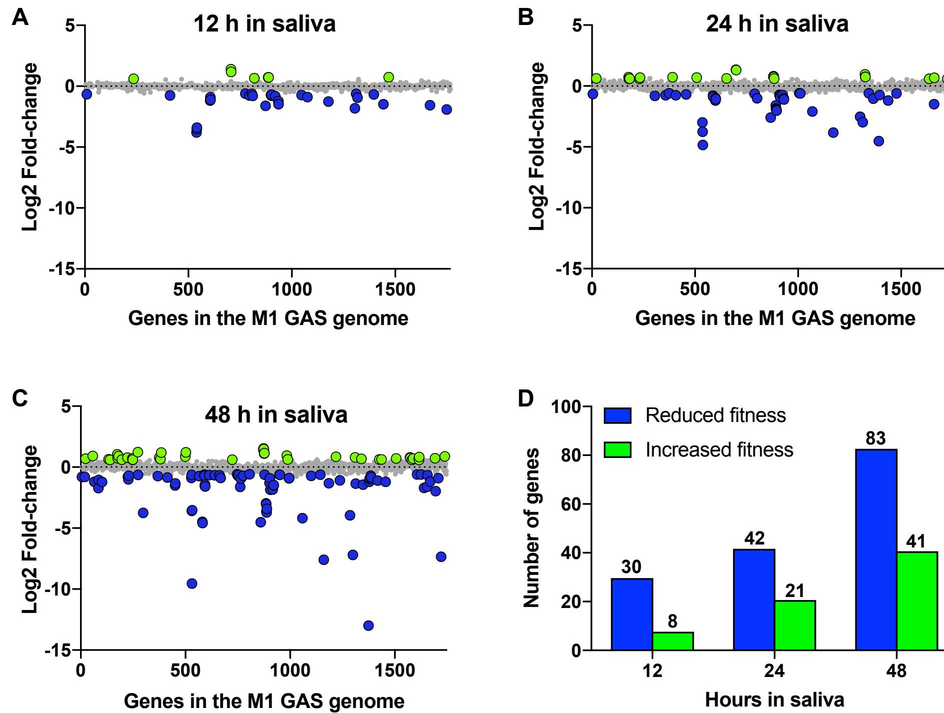


732

733 **FIG 1 Near-random distribution (A) and high density of transposon insertions (B)**
734 **of the M1 GAS input mutant library. (A) Insertion index (number of insertion sites**
735 **divided by gene length) (Y-axis) of each gene (X-axis) in the M1 GAS reference**
736 **genome. (B) A representative section of the transposon insertion map. As**
737 **expected, essential gene *ftsY* has no insertion because it is not represented in the**
738 **library. Red vertical spikes are forward reads; blue vertical spikes are reverse**
739 **reads. Read orientation indicates the direction of the transposon insertion.**

740

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



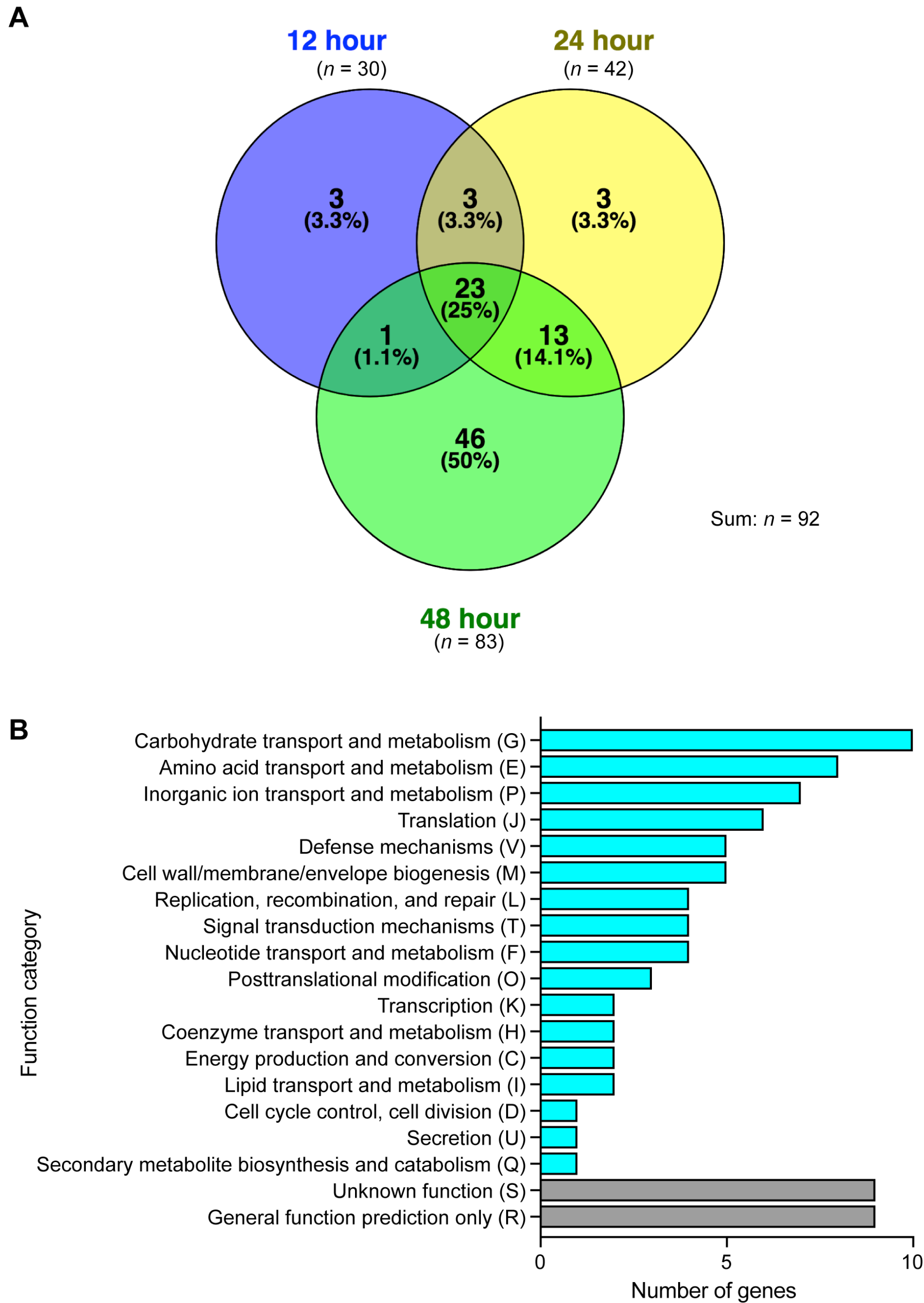
741

742 **FIG 2 TraDIS analyses of GAS fitness genes during exposure to human saliva *ex***
743 ***vivo*. (A-C) Genome-scale summary of the change in mutant abundance (Y-axis)**
744 **for each gene (X-axis) in the output mutant pools recovered after 12 hrs, 24 hrs**
745 **and 48 hrs of incubation in human saliva *ex vivo*. Gene mutations (insertions)**
746 **conferring significantly decreased (blue circles) or increased (green circles)**
747 **fitness are highlighted. Insertion mutations that lack a significantly altered fitness**
748 **phenotype (grey circles) are also indicated. (D) Summary of GAS genes identified**
749 **to be important for fitness in saliva after indicated period of saliva incubation**.

750

751

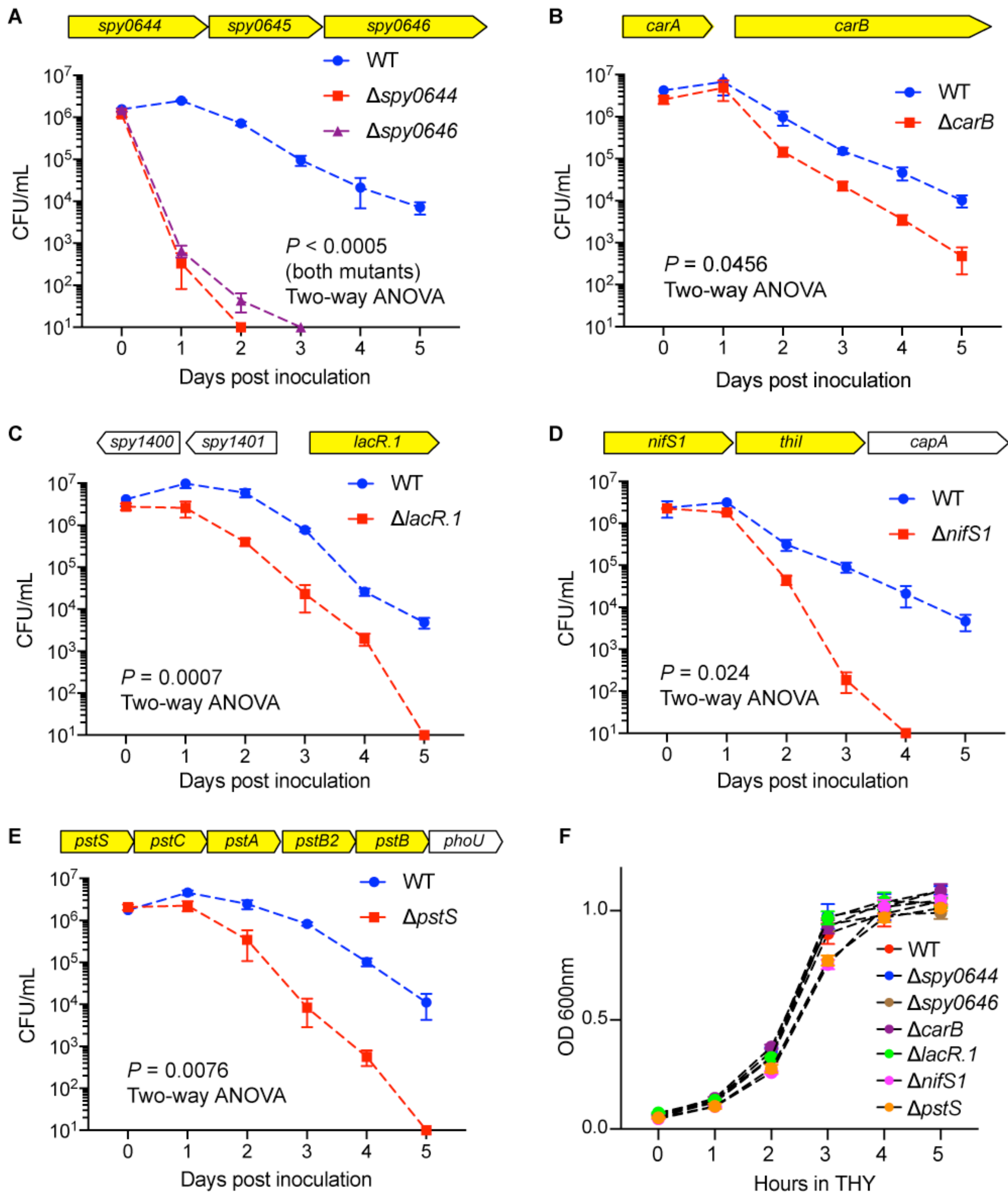
bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



752

753 **FIG 3 Venn diagram showing number of GAS genes identified to be important for**
754 **fitness in saliva after indicated period of incubation (A), and functional**
755 **categorization of the 92 identified GAS saliva-fitness genes (B). Note that in panel**
756 **A the circles sizes are not proportional to the numbers of genes identified simply**
757 **to improve visual presentation and clarity.**

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



758

759 **FIG 4 Validation of the findings from the TraDIS saliva screen. The saliva**

760 **persistence phenotype was determined for each of six GAS isogenic mutant**

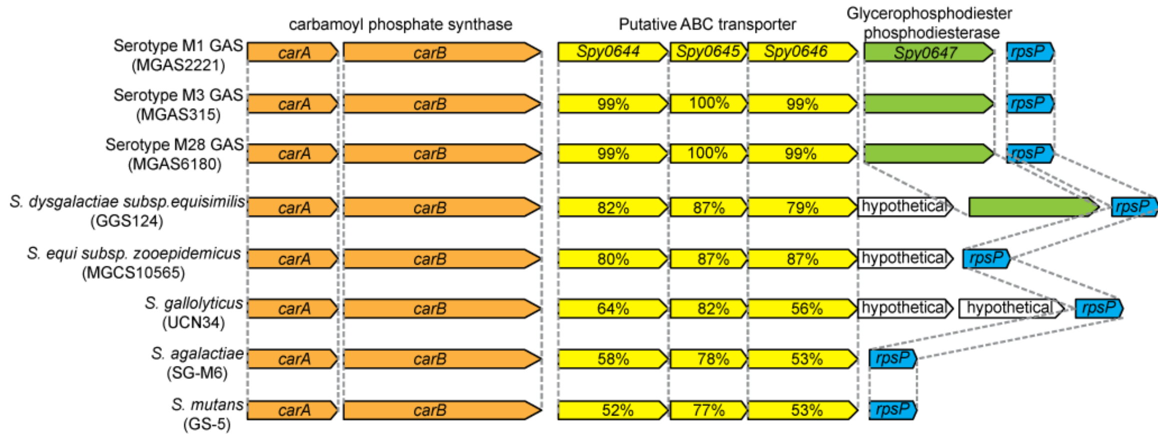
761 **strains (A-E). Highlighted genes (yellow) indicate the putative saliva-fitness genes**

762 **identified by TraDIS. *P*-values were assessed by repeated-measures 2-way**

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

763 **ANOVA. The growth phenotype in rich medium (THY) was also determined for**
764 **each of six GAS isogenic mutant strains (F).**
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



786

787

788 **FIG 5 Homologous regions encoding genes *Spy0644-Spy0646* in GAS and other**

789 **bacteria. Yellow arrows represent genes *Spy0644*, *Spy0645*, *Spy0646* and their**

790 **homologs. Percentages denote amino acid identity compared to serotype M1 GAS**

791 **strain MGAS2221.**

792

793

794

795

796

797

798

799

800

801

802

803

804

805

bioRxiv preprint doi: https://doi.org/10.1101/196931; this version posted October 2, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

806 **FIG S1** Experimental strategy of TraDIS mutant screens in human saliva.

807

808 **TABLE S1** Essential genes identified in the genome of serotype M1 GAS strain

809 MGAS2221 under the conditions tested.

810

811 **TABLE S2** Gene mutations conferring decreased fitness after 12, 24 and 48 hour

812 incubation with human saliva.

813

814 **TABLE S3** Gene mutations conferring increased fitness after 12, 24 and 48 hour

815 incubation with human saliva.

816

817 **TABLE S4** Fitness score of each gene in the M1 GAS genome after 12 hour saliva

818 incubation.

819

820 **TABLE S5** Fitness score of each gene in the M1 GAS genome after 24 hour saliva

821 incubation.

822

823 **TABLE S6** Fitness score of each gene in the M1 GAS genome after 48 hour saliva

824 incubation.

825

826 **TABLE S7** Primers used in this study.

827

828

829

830

831