ArcR Modulates Biofilm Formation in the Dental Plaque Colonizer *Streptococcus gordonii*

Jill C. Robinson[1], Nadia Rostami[1], John Casement[2], Waldemar Vollmer[3], Alexander H. Rickard[4], Nicholas S. Jakubovics[1*]

1. School of Dental Sciences, Newcastle University, Newcastle upon Tyne, UK
2. Bioinformatics Support Unit, Newcastle University, Newcastle upon Tyne, UK
3. Centre for Bacterial Cell Biology, Newcastle University, Newcastle upon Tyne, UK
4. School of Public Health, Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA

Running head: *ArcR in* Streptococcus gordonii *biofilms*

Keywords: Saliva, dental plaque, microarray, oral streptococci, ArcR regulon

[*]Correspondance: Nicholas S. Jakubovics, Centre for Oral Health Research, School of Dental Sciences, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4BW, UK. Tel: +191 208 6796; Fax: +191 208 8807; e-mail: nick.jakubovics@ncl.ac.uk

WORD COUNT = 8,579 words

SUMMARY

Biofilm formation and cell-cell sensing by the pioneer dental plaque coloniser *Streptococcus gordonii* is dependent upon arginine. This study aimed to identify genetic factors linking arginine-dependent responses and biofilm formation in *S. gordonii*. Isogenic mutants disrupted in genes required for biosynthesis or catabolism of arginine, or for arginine-dependent gene regulation, were screened for their ability to form biofilms in a static culture model. Biofilm formation by a knockout mutant of *arcR*, encoding an arginine-dependent regulator of transcription, was reduced to <50% that of the wild-type whereas other strains were unaffected. Complementation of *S. gordonii* Δ*arcR* with a plasmid-borne copy of *arcR* restored the ability to develop biofilms. By DNA microarray analysis, 25 genes were differentially regulated in *S. gordonii* Δ*arcR* compared with wild-type under arginine-replete conditions including 8 genes encoding components of phosphotransferase systems for sugar uptake. By contrast, disruption of *argR* or *ahrC* genes, which encode paralogous arginine-dependent regulators, each resulted in significant changes in the expression of more than 100 genes. Disruption of a gene encoding a putative extracellular protein that was strongly regulated in *S. gordonii* Δ*arcR* had a minor impact on biofilm formation. We hypothesise that genes regulated by ArcR form a critical pathway linking arginine sensing to biofilm formation in *S. gordonii*. Further elucidation of this pathway may provide new targets for the control of dental plaque formation by inhibiting biofilm formation by a key pioneer coloniser of tooth surfaces.

INTRODUCTION

The formation of dental plaque is initiated by the attachment of pioneer colonisers to the tooth surface.[1] Oral streptococci including *S. gordonii*, *S. sanguinis*, *S. oralis* and *S. mitis* are particularly well adapted for the initial colonisation of tooth surfaces since they produce a multitude of cell surface adhesin proteins and glycoproteins that recognise host receptors in the salivary pellicle.[2-4] Many of these adhesins also contribute to the subsequent development of dental plaque by mediating cell-cell binding, known as coaggregation, with other oral bacteria.[5-7]

Oral streptococci are considered to be opportunistic pathogens since they can enter the bloodstream and are among the leading causes of the rare, but life-threatening, disease infective endocarditis.[8] However, in mature dental plaque there is evidence that oral streptococci protect against dental caries. Thus, some species produce arginine deiminase, which generates ammonia and neutralises plaque acid leading to shifts in the microbiome towards health.[9-11] Arginine deiminase also directly influences other oral bacteria by acting as an interspecies signalling molecule. For example, the arginine deiminase of *Streptococcus cristatus* is sensed by the periodontal pathobiont *Porphyromonas gingivalis* and leads to down-regulation of virulence gene expression.[12] The uptake of arginine by oral streptococci such as *S. gordonii* occurs through an arginine-ornithine antiporter, ArcD, and ornithine released into the growth medium can be utilised by other species such as *Fusobacterium nucleatum.*[13] However, at high concentrations (≥50 mM), arginine can disrupt coaggregation between *S. gordonii* and *F. nucleatum* by inhibiting the *F. nucleatum* adhesin RadD.[14,15] High

concentrations of arginine also disrupt multispecies oral microbial biofilms or *S. gordonii* monospecies biofilms.[16-19]

There is evidence that oral streptococci such as *S. gordonii* use arginine as a key signal for growth processes and biofilm formation. *Streptococcus* and related genera including *Enterococcus*, *Lactococcus* and *Lactobacillus* possess between 2 and 4 copies of ArgR family arginine-sensing transcriptional regulators, enabling close co-ordination of arginine biosynthesis, catabolism and transport.[19] The *S. gordonii* genome, for example, encodes 3 different ArgR family regulators designated ArgR, AhrC and ArcR. Shifting *S. gordonii* cells from arginine-replete medium to medium lacking arginine results in changes in expression of >450 genes, nearly one quarter of the entire genome.[19] Arginine sensing and biofilm formation pathways are triggered by intermicrobial interactions. For example, interactions with *Candida albicans* lead to down-regulation of *S. gordonii arcA* and *arcB* encoding components of the arginine deiminase system, and up-regulation of the biofilm-associated operon *fruRBA*.[20] Coaggregation of *S. gordonii* with *A. oris* resulted in the co-ordinated down-regulation of 9 arginine biosynthesis genes and up-regulation of the biofilm-promoting *bfb* locus.[21] Coaggregation with *A. oris* also enabled the growth of *S. gordonii* in arginine-restricted medium. Therefore, it is possible that *S. gordonii* arginine-dependent regulators are employed to sense cell-cell interactions and respond by initiating growth and/or biofilm formation.

On the basis of the above observations, we hypothesised that one or more genes involved in arginine regulation and/or metabolism is linked to biofilm formation in oral streptococci. Since *S. gordonii* has been well-characterised in terms of responses to arginine, we used this species as a model to explore the genetic basis of the link between arginine and biofilm formation by oral bacteria. Initially, a molecular genetic approach was employed to screen for components of arginine pathways (regulation, biosynthesis, or catabolism) that are linked to biofilm formation. To obtain further insights into arginine-mediated gene regulation, we characterised the regulons of the three *S. gordonii* ArgR homologues, ArcR, ArgR and AhrC.

METHODS

Bacterial strains and growth conditions

Bacterial strains used in this study are listed in Table S1. *S. gordonii* was routinely cultured in THYE medium containing 30 g L^{-1} Bacto™ Todd Hewitt Broth (Becton Dickinson, Oxford, UK) and 5 g L^{-1} Yeast Extract (Melford Laboratories Ltd, Ipswich, UK) or on solidified THYE containing 15 g L^{-1} Bacto-agar (Becton Dickinson). Cells were cultured in candle jars without shaking for 24-48 h at 37°C. For biofilm assays, *S. gordonii* was cultured in TYEG medium containing 10 g L^{-1} Bacto Tryptone, 5 g L^{-1} Yeast Extract, 3 g L^{-1} K_2HPO_4 and 2 g L^{-1} D-glucose, adjusted to pH 7.5 before autoclaving. *E. coli* was cultured in Luria–Bertani (LB) medium at 37°C, 250 rpm or on LB medium solidified by the addition of 15 g L^{-1} Bacto-agar. When required antibiotics were included in growth media at the following concentrations: erythromycin 2 µg mL^{-1}, spectinomycin 100 µg mL^{-1}, kanamycin 250 µg mL^{-1}.

87

88 *Genetic manipulation of* S. gordonii

89 Routine genetic manipulations were conducted in accordance with standard protocols.[22]
90 Previously constructed gene knockout mutants are described in Table S1. Disruption of
91 SGO_RS04150 (here designated '*earA*') was performed using PCR overlap extension
92 mutagenesis as described by Jakubovics et al[19]. Briefly, flanking regions of the *earA* gene
93 were PCR amplified using primers *earA* F1 ovex and *earA* R1 kan ovex to generate an 869 bp
94 product in the 5' region of *earA* and kan F1 ovex and kan R1 ovex to generate a 903 bp
95 product in the 3' end of *earA* (Table S2). The *aphA3* cassette (910 bp) was amplified from
96 plasmid pSF151 with primers aphA3F2/R2.[19] The PCR products were stitched together in an
97 overlap extension PCR reaction. The resulting product was cleaned and used for
98 transformation of *S. gordonii* DL1. Successful disruption and replacement of *earA* gene with
99 the *aphA3* cassette was confirmed by DNA sequencing.

100

101 To produce a genetic complementation strain (*S. gordonii arcR*$_{Comp}$), plasmid p*arcR*$_{Comp}$ was
102 generated using the In-Fusion HD PCR ligation cloning kit (Clontech, Mountain View, CA). A
103 181 bp region of the synthetic CP25 promoter was amplified from plasmid pCM18,[23] using
104 primers CP25F and CP25R. Primers pPE1010F and pPE1010R were used to amplify a 5,652
105 bp fragment of vector pPE1010.[24] Primers *arcR*CompF and *arcR*CompR, containing 15bp
106 regions of overlap with pPE1010, were utilised to amplify a 494 bp fragment containing the
107 *arcR* gene from *S. gordonii* chromosomal DNA. The In-Fusion HD PCR ligation cloning kit was
108 employed to fuse the CP25 promoter and *arcR* gene into the pPE1010 vector. The integrity

of plasmid p*arcR*$_{Comp}$ was confirmed by sequencing, and p*arcR*$_{Comp}$ was used for transformation of *S. gordonii* Δ*arcR::aad9* to generate *S. gordonii arcR*$_{Comp}$.

Crystal violet biofilm assay

Biofilms for crystal violet assays were cultured on the surface of Cellstar® 96-well microtiter plates in TYEG medium without shaking, aerobically for 18 h at 37°C (Greiner Bio-one, Stonehouse, UK). The biomass was measured as described by Shields et al[25] Biofilms were submerged in 100 µL of 0.5% (w/v) crystal violet. After incubation at 20°C for 15 min, wells were washed three times in 200 µL of PBS, air-dried, and residual crystal violet was dissolved with 100 µL of 7% (v/v) acetic acid and quantified by measuring A_{562}. All experiments were performed three times independently. Statistical significance of differences between mutants and wild-type *S. gordonii* was assessed by ANOVA with Dunnett's post-hoc test, and $p<0.05$ was considered significant.

Fluorescent staining and Imaging

Biofilms for visualization experiments were cultured on sterile glass coverslips incubated in wells of 6-well tissue culture dishes. Following growth, biofilms were rinsed with PBS and incubated with Live/Dead BacLight stain (Molecular Probes) for 15 min at 20°C. For confocal laser scanning microscopy (CLSM), stained coverslips were rinsed with PBS and inverted onto a PBS-filled Gene frame (25 µL, 1.0 x 1.0 cm, Thermo Fischer Scientific) secured on a microscope slide. Imaging was performed using a Nikon A1R confocal laser scanning microscope fitted with CFI PLAN APO VC objective (Nikon 60x/1.40 Oil). Images were

captured with NIS-Elements C (v4.4, Nikon) software and processed using Imaris (v8.2, Bitplane) software. Biovolume quantification of Z-stacks was conducted using Volocity software (v6.3, PerkinElmer, UK), set to identify objects ≥1 μm^2 as *S. gordonii* cells. At least three Z-stacks (image size 1024 x 1024) from three different fields for view was analysed for each strain. The data were analysed from three independent experiments. Statistical significance of differences between biofilm biovolume was assessed using ANOVA with Tukey's post-hoc test.

Growth in chemically defined medium and RNA extraction

For gene regulation analysis, chemically-defined FMC medium [19] was used with 0.5 mM L-arginine HCl (Sigma-Aldrich, Dorset, UK) included or omitted as appropriate. Briefly, *S. gordonii* was cultured in FMC medium aerobically at 37°C for 18 h. Cells were harvested, washed with fresh FMC and resuspended in 20 mL FMC medium. Cultures were incubated at 37°C until an OD_{600} of 0.3-0.4 was reached. At this point, cultures were split into 5 mL aliquots, and cells were harvested and resuspended in 5 mL of either fresh arginine-replete FMC, or FMC without L-arginine. Cells were cultured at 37°C for a further 30 min. To extract RNA, 1 volume of RNALater was added, cultures were vortex-mixed and incubated for 5 min at 20°C. Cells were harvested and, after discarding the supernatant, pellets were stored for up to 5 days at -80°C. Cell pellets were thawed on ice and resuspended in 100 µl spheroplasting buffer [26% (w/v) raffinose, 10 mM $MgCl_2$, 20 mM Tris-HCl, pH 6.8] containing 0.1 mg/ml chloramphenicol or spectinomycin. Mutanolysin (500 U mL^{-1}) was added to cells and incubated at 37°C for 5 min before addition of 350 µl RNAWiz solution (Life Technologies). The mixture was vortexed vigorously for 15 s and RNA was extracted

using the Ambion RiboPure Bacteria RNA Purification kit (Life Technologies, Carlsbad, California, US) in accordance with the manufacturer's recommendations.

Gene expression analysis by RT-qPCR

For RT-qPCR, samples were reverse transcribed with QuantiTect Reverse Transcription kit (Qiagen, Valencia, CA). Reactions were performed according to manufacturer's instructions, with the modification that 3 µg/mL random hexamer primers (Bioline, London, UK) were used in place of the QuantiTect oligo-dT primers. The cDNA was cleaned and used as template in RT-qPCR experiments with the SensiMix SYBR No-ROX kit (Bioline) with the following reaction conditions: 1. 95°C for 10 min, 2. 95°C for 15 s, 3. 60°C for 1 min, 4. plate read, 5. repeat from step 2 a further 39 times, 6. melting curve from 55-90°C, read every 1°C, hold for 5 s. All samples were normalised against the 16S rDNA gene. Primer sets for this and for *argC, argG, pyrA*$_b$*, arcA, arcB, arcD* and *amyB* genes have previously been published.[19] Other primers are described in Table S2. Standard curves, melting curves and agarose gel electrophoresis analysis of the cDNA, were routinely included to validate the RT-qPCR experiments.

Microarray analysis

Microarray analysis was performed as described by Jakubovics et al[19] using a previously designed microarray containing 2,051 probes for *S. gordonii* genes (GEO accession GPL17786). Samples of RNA from four independent experiments per strain/growth condition were sent to the Genomics and Microarray Facility, Birmingham University, UK for

reverse transcription, labelling and hybridisation. The microarray series for *S. gordonii* DL1 in high or no arginine was previously deposited as GEO accession GSE51346. Other data series for arginine-dependent regulator mutants in high or no arginine were deposited in GEO with the following accession numbers: *S. gordonii* Δ*arcR*, GSE101509; *S. gordonii* Δ*argR*, GSE101506; *S. gordonii* Δ*ahrC*, GSE101507.

Data were analysed using Agilent GeneSpring GX software. Probe expression data was quantile normalised to enable unbiased comparisons between samples. To assess the relatedness of samples, principal component analysis (PCA) was carried out using the normalised data. Outliers identified in the PCA analysis were removed before proceeding to significance analysis. A moderated t-test was used to determine statistics for each probe. The resulting *p*-values were then adjusted using the Benjamini-Hochberg multiple testing correction procedure.[26] For each comparison, probes with a corrected *p*-value of ≤0.05 and with fold change numerically greater than 2 were considered to be differentially expressed between conditions. Rank product analysis was carried out using the RP method implemented in the R package 'RankProd'.[27,28] This method applies a non-parametric statistical test, based on rank ordering of genes according to fold changes, to detect genes that are consistently upregulated or downregulated in replicated experiments. The heatmap was produced using the R package ComplexHeatmap.[29]

RESULTS

Effects of disrupting arginine metabolism or regulation genes on S. gordonii *biofilm formation*

To investigate the genetic basis for the previously observed links between arginine sensing and biofilm formation in *S. gordonii*, a range of mutants lacking genes involved in arginine-dependent regulation, arginine catabolism or biosynthesis were screened for their ability to form biofilms in a high-throughput static 96-well microplate system. Biofilms were cultured for 24 h in TYEG medium, washed and stained with crystal violet to quantify the biofilm biomass (Figure 1). Biofilm formation was not significantly different from wild-type levels in any of the single mutants screened with the exception of *S. gordonii* Δ*arcR*, which formed approximately 50% less biofilm than *S. gordonii* DL1. This was not due to a defect in planktonic growth as the growth yield, measured as the OD_{600} of the well prior to washing and staining, was almost identical between *S. gordonii* DL1 and Δ*arcR* (data not shown). It was consistently observed that biofilms formed by *S. gordonii* Δ*arcR* appeared similar to those produced by the wild-type until they were agitated, indicating that cells were loosely attached. The *arcR* gene encodes an arginine-dependent regulator of the ArgR family. To assess whether ArcR acts in concert with other ArgR-family regulators to control biofilm formation, the *arcR* mutation was introduced into *S. gordonii* Δ*argR* and Δ*ahrC* backgrounds. In each case, biofilm formation by the double mutants containing an *arcR* knockout was similar to that of the Δ*arcR* single mutant, and approximately 50% reduced compared with wild-type. By contrast, an *S. gordonii* Δ*argR* Δ*ahrC* double mutant was not impaired in biofilm formation (Figure 1). Therefore, it appears that ArcR and not ArgR or AhrC is required for efficient biofilm formation in *S. gordonii*.

Visualisation of biofilms and genetic complementation of S. gordonii *Δ*arcR

Analysis of biofilms by CLSM with BacLight LIVE:DEAD staining revealed clear differences between biofilms formed by *S. gordonii* DL1 and *S. gordonii* Δ*arcR* (Figure 2). Biofilms formed by the wild-type were confluent and relatively smooth, and approximately 10-20 µm thick throughout. By contrast, *S. gordonii* Δ*arcR* biofilms were more heterogenous, with clumps of cells up to 40 µm thick and patches of surface that were not covered at all. To confirm that the observed biofilm defects were due to disruption of the *arcR* gene and not a second site mutation, the *arcR* gene was reintroduced into *S. gordonii* Δ*arcR* on a plasmid under regulation of the synthetic CP25 promoter to generate *S. gordonii* $arcR_{COMP}$. Biofilms formed by *S. gordonii* $arcR_{COMP}$ were very similar in structure to those of the wild-type. Quantitative assessment using image analysis software demonstrated that the biovolume of *S. gordonii* Δ*arcR* biofilms was significantly reduced compared with wild-type *S. gordonii* or the genetic complementation strain *S. gordonii* $arcR_{Comp}$, indicating that the biofilm formation defect observed in *S. gordonii* Δ*arcR* was a direct result of *arcR* gene disruption.

Analysis of the ArcR regulon

To assess the effects of disrupting the *arcR* gene on global gene expression in *S. gordonii*, microarrays were employed to analyse gene expression in *S. gordonii* DL1 or Δ*arcR* in high arginine or following a shift to no arginine. Cells were cultured to mid-exponential phase in arginine-replete chemically defined growth medium, harvested and resuspended in high (0.5 mM) or no arginine. After 30 min, RNA was extracted and gene expression was

241 monitored by microarray. Initially, the results from the microarray were validated by
242 assessing the expression levels of 7 genes under high and no arginine using RT-qPCR. The
243 genes *argC, argG, pyrAb, arcA, arcB, arcD* and *amyB* (Figure 3) were selected for this analysis. Of
244 these, *amyB* was was included as a control, and the other genes were chosen as they were
245 previously reported to be implicated in arginine metabolism and transport.[19] The comparison
246 showed a strong correlation between the microarray and RT-qPCR, with r^2 values of 0.995 and
247 0.777 for *S. gordonii* DL1 and Δ*arcR*, respectively, and slopes very close to 1 for each strain.
248 To further assess the validity of individual microarray experiments, gene regulation from all
249 *S. gordonii* DL1 and Δ*arcR* microarray samples under high and no arginine was compared by
250 principle coordinates analysis (PCoA; data not shown). Four independent experiments were
251 performed for each strain under each condition, giving a total of 16 microarray samples.
252 However, PCoA identified two outliers in the data: one sample each of *S. gordonii* DL1 under
253 high arginine and no arginine. Therefore these samples were removed from subsequent
254 data analysis.

255

256 In total, 26 genes were significantly different between *S. gordonii* DL1 and Δ*arcR* under
257 arginine-replete conditions (Table S3). This included *arcR* itself, which gave a detectable
258 signal in the knockout strain even though the gene was not present. This signal was very low
259 and was not dependent on arginine, and therefore was presumably due to background
260 fluorescence. Of the other 25 genes, the most strongly regulated was a putative
261 extracellular protein encoded by SGO_RS04150, which was up-regulated 110-fold in *S.*
262 *gordonii* Δ*arcR* compared with the wild-type. The *argGH* operon (SGO_RS00865-00870) was
263 up-regulated 5 to 6-fold. The gene *queA*, encoding a putative S-adenosylmethionine:tRNA

ribosyltransferase-isomerase, was up-regulated 12-fold in *S. gordonii* Δ*arcR*. An apparent 4-gene operon (SGO_RS07910-07925) encoding a predicted magnesium transporter, a CAAX amino protease, a methyltransferase and peptide chain release factor 3, was up-regulated 4 to 6-fold. Two operons encoding phosphotranserase system (PTS) components were also up-regulated: *manLM* encoding mannose-specific enzyme IIAB and IIC was up-regulated ~2.5-fold and *levDEFG* encoding a fructose PTS enzyme IIA, IIB, IIC and Ia was up-regulated 7 to 12-fold. In addition, 3 hypothetical proteins, a possible pseudogene, an additional CAAX amino protease, peptidase S11, a putative ATP-dependent Clp protease ATP-binding protein, and a possible carboxylate-amine ligase were each up-regulated 2 to 4-fold. Only 3 genes were down-regulated in *S. gordonii* Δ*arcR*: SGO_RS02495 and SGO_RS08100 encoding putative PTS enzyme IIABC components (both down-regulated ~2-fold) and the *coiA* gene encoding a putative gene involved in genetic competence, which was 9-fold down-regulated.

To assess the impact of the *arcR* gene knockout on the ability of *S. gordonii* to respond to arginine, the ratio of gene expression following a shift to no arginine versus maintenance in high arginine for each of the above genes was calculated for *S. gordonii* DL1 and Δ*arcR* (Figure 4). This analysis revealed several different patterns of gene regulation. The two PTS operons *manLM* and *levDEFG* along with the 4-gene operon containing a putative magnesium transporter (SGO_RS07910-07925) had little or no arginine-dependent regulation in *S. gordonii* DL1 but were more strongly down-regulated in response to arginine restriction in *S. gordonii* Δ*arcR*. The reverse was seen with SGO_RS07765 encoding peptidase S11 and *coiA*, which were more strongly down-regulated in the wild-type than *S.*

gordonii Δ*arcR* in response to the removal of arginine. The *queA* gene, SGO_RS00730 (hypothetical protein) and SGO_RS09090 (Clp protease ATP binding protein) were up-regulated in the absence of arginine to approximately the same extent in the wild-type and Δ*arcR* mutant. SGO_RS02495 glucose PTS enzyme IIABC subunit, SGO_RS01240 encoding a CAAX amino protease, SGO_RS08100 encoding trehalose PTS enzyme IIABC and SGO_RS10240 carboxylate-amine ligase were expressed at similar levels in no versus high arginine the wild-type, but were up-regulated under arginine restriction in *S. gordonii* Δ*arcR*. Finally, SGO_RS07655 and SGO_RS04295 encoding hypothetical proteins, the pseudogene SGO_RS06275 and *argGH* were strongly down-regulated under high arginine in the wild-type and less strongly down-regulated in *S. gordonii* Δ*arcR*.

Gene regulation in S. gordonii *Δ*argR *and Δ*ahrC

To compare the ArcR regulon with those of the orthologous ArgR-family regulators, ArgR and AhrC, microarray analysis was performed on the *S. gordonii* Δ*argR* and Δ*ahrC* mutants. The validity of the microarrays was checked by RT-qPCR for the genes *argC, argG, pyrA*$_b$*, arcA, arcB, arcD* and *amyB* (Figure 3). As with the previous microarrays, there were strong correlations between the microarray and RT-qPCR data (r^2 values >0.95 and slopes close to 1). There were no significant differences between gene expression in *S. gordonii* Δ*argR* or Δ*ahrC* under either high or no arginine. The microarrays for *S. gordonii* Δ*argR* and Δ*ahrC* were performed by a different operator and at a different time from the *S. gordonii* DL1 and Δ*arcR* arrays. Performing microarray analyses at different times can lead to 'batch effects' where absolute gene expression levels are slightly different across arrays due to the use of different reagents or slight changes in protocol.[30] Further, a relatively high proportion of

probes (~19%) did not give a quantifiable signal from the *S. gordonii* Δ*argR* and Δ*ahrC* samples. To assess the relationships between the microarrays, the data from each array was compared by PCoA (Figure S1). The data for *S. gordonii* Δ*argR* and Δ*ahrC* clustered together and were well separated from *S. gordonii* DL1 and Δ*arcR*. Although it is possible that these observations reflect genuine biological differences between the strains, the comparisons between *S. gordonii* Δ*argR*/Δ*ahrC* and *S. gordonii* DL1/*arcR* must be treated with caution. Therefore, it was not possible to define the ArgR and AhrC regulons by direct comparison with *S. gordonii* DL1 under high or no arginine. Instead, rank product analysis was used to identify significant differences in gene expression between *S. gordonii* Δ*argR*/Δ*ahrC* and *S. gordonii* DL1 under high arginine.

Rank product analysis does not quantify the differences in expression levels between mutants and wild-type, but instead ranks the most strongly regulated genes. In total, 109 genes were identified using this approach that were differentially regulated in *S. gordonii* Δ*argR* and/or Δ*ahrC* compared with *S. gordonii* DL1 (Table S4). The most strongly up-regulated genes were the *argCJBD* arginine biosynthesis operon, consistent with the previously identified roles of ArgR and AhrC as arginine-dependent repressors of transcription of arginine biosynthesis genes. Another arginine biosynthesis gene operon, *pyrRA*$_a$*A*$_b$, and the histidine biosynthesis operon (SGO_RS06875-SGO_RS06920) were also strongly and co-ordinately up-regulated in the mutants. Approximately as many genes were down-regulated in the mutants as up-regulated. Down-regulated genes included *purE* and *purM*, involved in purine metabolism, and *scaC* and *scaA* encoding components of an ABC-type manganese transporter.

Of the 26 genes identified as ArcR-regulated, 10 genes were not detected in the *S. gordonii* Δ*argR* and Δ*ahrC* samples (shaded grey in Figure 4). The *arcR* and *queA* genes appeared slightly more strongly up-regulated in arginine restriction in *S. gordonii* Δ*argR* and Δ*ahrC* mutants than the wild-type (Figure 4). The SGO_RS07910-07925 operon and *manLM* were weakly down-regulated in no arginine in *S. gordonii* Δ*argR* and Δ*ahrC*. This pattern of gene expression was more similar to *S. gordonii* DL1 than to *S. gordonii* Δ*arcR*. By contrast, the pattern of arginine-dependent gene regulation in *S. gordonii* Δ*argR* and Δ*ahrC* for the S11 peptidase gene, SGO_RS07655 and SGO_RS04295 hypothetical protein-encoding genes and *argGH* was more similar to *S. gordonii* Δ*arcR* than to the wild-type. SGO_RS09090 Clp protease ATP binding protein locus was up-regulated in no arginine to similar extents in all *S. gordonii* strains. Overall, these data demonstrate that disruption of *arcR* leads to several unique changes in gene expression that are not seen in *S. gordonii* Δ*argR* or Δ*ahrC*.

Role of SGO_RS04150 in biofilm formation

Of the differentially regulated genes, SGO_RS04150 (SGO_0846) was of particular interest since it was strongly regulated in *S. gordonii* Δ*arcR*. This gene is annotated as encoding a cell wall protein in the NCBI database. Using SignalP,[31] we identified a putative N-terminal secretion signal. However, using PSORTb[32] we were unable to detect an LPxTG cell wall anchor or a lipoprotein motif and it is possible that the protein is secreted from the cell surface. To confirm that this gene is regulated by ArcR, we assessed expression levels by RT-qPCR under arginine-replete or no arginine conditions (Figure Xa). Under high arginine, expression was significantly increased 17-fold in *S. gordonii* Δ*arcR* compared with wild-type.

In line with the microarray data (Figure 5a), expression was slightly elevated in no arginine in *S. gordonii* DL1 and slightly decreased in *S. gordonii* Δ*arcR*. In view of the strong regulation by ArcR, we have termed the SGO_RS04150 gene '*earA*', encoding '**E**xtracellular **A**rcR-**R**egulated protein A'. To assess the role of *earA* in *S. gordonii* biofilm formation, the gene was disrupted in the wild-type and the Δ*arcR* background. Using crystal violet staining, biofilm formation by *S. gordonii* Δ*earA* was slightly reduced compared with the wild-type (Figure 5b) but the difference was not significant. Similarly, knockout of *earA* in the Δ*arcR* or $arcR_{COMP}$ backgrounds did not significantly reduce biofilm formation. Therefore, it appears that dysregulation of *earA* expression is not the major cause of the biofilm defect observed in *S. gordonii* Δ*arcR*.

DISCUSSION

Our studies identified ArcR as a key determinant of biofilm formation in *S. gordonii*. As no arginine metabolic gene disruptions affected biofilm formation, it appears that ArcR may be acting through a mechanism that is independent of regulation of arginine metabolism. Interestingly, a different ArgR family regulator, EF0676, the orthologue of ArgR, was identified in a transposon mutant screen for biofilm-associated genes in *Enterococcus faecalis.*[33] Subsequently, the homologue of AhrC, EF0983, was also implicated in biofilm formation since strains carrying a transposon insertion were impaired in the development of biofilms *in vitro* and in biofilm infection models of infective endocarditis and catheter-associated urinary tract infection.[34] The *E. faecalis* genome also contains two genes encoding orthologues of *S. gordonii* ArcR that have been named *argR2* and *argR1.*[35] However, to the best of our knowledge neither of these, or any other orthologues of *S. gordonii arcR* have yet been associated with biofilm formation phenotypes.

ArgR-family regulators form hexameric complexes that are activated by the binding of arginine, and bind to regulatory regions termed ARG box elements in the upstream region of target genes.[36] It is not clear why streptococci and related bacteria produce multiple ArgR family homologues. In particular, the function of ArgR and AhrC is often very similar, and knocking out one or other gene is sufficient to abrogate arginine-dependent regulation of target genes.[19] It has been postulated that these may form a heterohexameric complex or that they may co-ordinate gene expression by cross-regulating one another.[37] Transcriptome analysis in *Lactococcus lactis* or *Streptococcus pneumoniae* has identified minor differences between the ArgR (termed ArgR1 in *S. pneumoniae*) and AhrC regulons, indicating that heterohexameric complexes containing both ArgR and AhrC are not essential for gene regulation.[38,39] By contrast, direct comparison between gene expression in *S. gordonii* Δ*argR* and Δ*ahrC* here did not identify any significant differences between the strains in either high or no arginine. Nearly 20% of the probes failed to make the quality control thresholds and it is possible that some differences between the ArgR and AhrC regulons were missed in our analysis. Nevertheless, it appears that ArgR and AhrC have very similar functions in *S. gordonii*.

In a previous microarray analysis of *S. gordonii* DL1, >450 genes were shown to be significantly regulated following a shift from high to no arginine.[19] *S. gordonii* has the capacity to biosynthesise arginine, but does not grow aerobically in media lacking arginine.[21] Therefore, it is possible that some of this gene regulation was an adaptation to stress rather than a specific response to the lack of arginine. In growth experiments using FMC, *S.*

gordonii Δ*argR*, Δ*ahrC* and Δ*arcR* were similarly impaired in aerobic growth in the absence of arginine (data not shown). Therefore, differences in gene regulation between *S. gordonii* DL1 and the mutant strains are likely to reflect the function of the regulators rather than simply stress imposed by the lack of arginine and the subsequent growth arrest. Many of the genes that were regulated in *S. gordonii* DL1 following a shift from high to no arginine, including arginine and histidine biosynthesis operons, were up-regulated specifically in high arginine in *S. gordonii* Δ*argR* or Δ*ahrC* compared with wild-type (Table S4). This pattern of regulation is consistent with the action of ArgR/AhrC acting as repressors in arginine-bound form. Using the transcription factor binding site prediction programme PePPER [40] to search for *L. lactis* ARG box elements throughout the *S. gordonii* genome, 6 ARG boxes were identified upstream of genes that were up-regulated in *S. gordonii* Δ*argR* or Δ*ahrC*. These were upstream of *argR*, *argC* (two elements), *pyrR*, *acnA* and SGO_RS08115. Of these, *argR* encodes the ArgR regulator, *argC* is the first gene in the *argCJBD* arginine biosynthesis operon and *pyrR* is the first gene in the *pyrRA*$_a$*A*$_b$ arginine biosynthesis operon. On the other hand, *acnA*, encoding aconitate hydratase and SGO_RS08115, encoding phosphoenolpyruvate carboxykinase, are involved in the citric acid cycle, indicating that arginine sensing is linked to central metabolism in *S. gordonii*. Although ArgR/AhrC are thought to act primarily as repressors of transcription, it is possible that they may also be activators in some conditions. The above screen identified one ARG box element upstream of *scaC*, the first gene in the *scaCBA* operon encoding an ATP-binding cassette manganese transporter,[41] that was down-regulated in *S. gordonii* Δ*argR* and Δ*ahrC* under high or low arginine. However, *scaB* was not regulated and *scaA* was only regulated in *S. gordonii* Δ*argR* under high arginine. Therefore, it is not clear what impact arginine-sensing has on the expression or function of the transporter as a whole. It is likely that there are many other

ARG box elements upstream of ArgR/AhrC-regulated genes in *S. gordonii*. In *E. coli*, 62 ArgR binding regions have been identified by chromatin immunoprecipitation (ChIP-Seq).[36,42] Measurement of ArgR/AhrC interactions with DNA in promoter regions by ChIP-Seq, electrophoretic mobility shift or DNase I footprinting would be required to determine which genes are directly regulated by ArgR/AhrC in *S. gordonii*.

The target DNA binding sequence for *S. gordonii* ArcR has been investigated by DNase I footprinting, and shown to consist of a 27 bp region with little identity to ARG box motifs [43]. ArcR has previously been shown to regulate the expression of the *arcABC* operon, *queA* and *argGH.*[19,44,45] The microarray analysis identified *queA* and *argGH* as genes that were up-regulated in *S. gordonii* Δ*arcR* compared with wild-type under high arginine. Interestingly, however, *arcA*, *arcB* and *arcC* were not differentially expressed between *S. gordonii* DL1 and Δ*arcR* under high or no arginine by microarray analysis. All three genes were up-regulated approximately 3 to 4-fold in high arginine compared with no arginine in both strains. It is well-established that disruption of the *arcR* gene abrogates arginine-dependent activation of *arcABC* expression[19,45] and it is therefore not clear why this was not seen in the microarray data. However, using RT-qPCR which has a higher dynamic range than microarrays,[46] it was found that expression of *arcA* and *arcB* was significantly lower in *S. gordonii* Δ*arcR* than the wild-type under high arginine, in line with previous observations (data not shown).

448 In addition to known genes, ArcR was shown to have a major impact on the regulation of
449 phosphoenolpyruvate:sugar phosphotransferase system (PTS) components. Oral
450 streptococci utilise many different carbohydrates for growth and PTSs are the primary
451 mechanism for sugar uptake.[47] Using proteomic analysis, it was found that PTSs are a
452 common target for gene regulation in *S. gordonii* following interspecies interactions with *P.*
453 *gingivalis* and *F. nucleatum.*[48] In addition, the PTS operon *fruRBA* was up-regulated
454 following adhesion of *S. gordonii* to *C. albicans.*[20] This operon was previously shown to have
455 a key role in biofilm formation.[49] Although *fruRBA* was not regulated by ArcR, another PTS
456 (*levDEFG*) was strongly (7 to 12-fold) up-regulated in *S. gordonii* Δ*arcR* under high arginine,
457 indicating that ArcR represses expression of this operon when arginine is available. The
458 genes *levDEFG* encode enzyme IA, IIA, IIB and IIC components of a PTS that primarily
459 transports fructose.[50] It is not yet known whether this system is required for the
460 development of biofilms by *S. gordonii*. Similarly, the PTS enzyme IIAB and IIC components
461 encoded by *manLM* have been assessed for their roles in sugar uptake and found to
462 transport glucose, galactose and mannose, but have not been assessed for biofilm-
463 associated functions.[50] Three additional PTSs have been shown to be important for biofilm
464 formation by *S. gordonii.*[51] Of these, the *bfb* locus encoding a cellobiose PTS has previously
465 been shown to be regulated by arginine and by coaggregation with *A. oris*.[19,21] Although the
466 *bfb* genes were not significantly regulated in *S. gordonii* Δ*arcR* compared with wild-type in
467 our microarray analyses, there was a trend towards up-regulation under high arginine.
468 Therefore, we assessed the expression of *bfbC* and *bfbF* in *S. gordonii* DL1 and Δ*arcR* by RT-
469 qPCR. Both genes were approximately 4-fold increased in *S. gordonii* Δ*arcR* compared with
470 wild-type under high arginine, and the difference was statistically significant (p<0.05) in the

471 case of *bfbC*. Therefore, it is possible that dysregulation of the *bfb* locus in the absence of
472 ArcR may be partly responsible for the observed biofilm defects of *S. gordonii* Δ*arcR*.

473

474 In addition to PTSs, disruption of *arcR* also led to a change in the regulation of an apparent
475 4-gene operon containing a magnesium transporter, CAAX amino protease,
476 methyltransferase and peptide chain release factor 3. These genes were only weakly
477 regulated by arginine in *S. gordonii* DL1, but were more strongly down-regulated in by a
478 shift to no arginine in *S. gordonii* Δ*arcR*. It is not clear why these genes are linked to arginine
479 sensing in *S. gordonii*. However, it is noteworthy that this operon is immediately upstream
480 of *acnA*, which contains an ARG box element in the promoter region. The gene encoding a
481 protease was one of two CAAX protease-encoding genes that were regulated by ArcR. The
482 function of these proteins is not fully understood, but some of them play a role in self-
483 immunity against bacteriocins.[52] Oral streptococci possess many paralogues of genes
484 encoding CAAX amino proteases: *S. gordonii* has 14, whereas *S. sanguinis* has 21 paralogues,
485 more than any other species that has been analysed to date.[52] One of these proteases in *S.*
486 *gordonii* (BfrH) and two in *S. sanguinis* (BfrH1 and BfrH2) are regulated by biofilm-associated
487 BfrAB two-component systems and it has been speculated that these may play important
488 roles in processing proteins that are secreted in order to promote biofilm development.[53]
489 The ClpP protease complex may also be involved in controlling the quality and integrity of
490 secreted proteins involved in biofilms. Thus, an ATP binding protein associated with the ClpP
491 complex was up-regulated in *S. gordonii* Δ*arcR* compared with wild-type under high
492 arginine. The ClpP complex consists of ClpP protease and a number of ATPases that together
493 are required for biofilm formation in *Streptococcus mutans.*[54] Disruption of the *arcR* gene

494 also led to a change in expression of the competence protein CoiA. In *S. pneumoniae* CoiA is
495 involved in genetic recombination and is only transiently expressed during competence. It is
496 possible that disruption of *arcR* led to the initiation of competence development,[55] although
497 it is important to note that other competence genes were not differentially regulated
498 between *S. gordonii* DL1 and Δ*arcR*. The gene encoding carboxylate-amine ligase is
499 immediately upstream of *purE* involved in purine biosynthesis, and may even be part of the
500 same operon. The *purE* gene was down-regulated in *S. gordonii* Δ*argR* and Δ*ahrC* compared
501 with wild-type under high and no arginine. Therefore, it appears that all three ArgR family
502 regulators may co-ordinate to control this locus, as they do for the regulation of *argGH*.
503 Finally, peptidase S11 is a D-alanyl-d-alanine carboxypeptidase, involved in peptidoglycan
504 biosynthesis. It is not clear whether this or any of the hypothetical proteins that were
505 regulated by ArcR could be involved in biofilm formation.

506

507 The strongest regulation in *S. gordonii* Δ*arcR* was a gene of unknown function, which we
508 have termed *earA*. This gene was up-regulated in *S. gordonii* following a shift to no arginine.
509 However, in *S. gordonii* Δ*arcR* expression of *earA* was strongly (>100-fold in the microarray
510 analysis) up-regulated under high arginine. The function of EarA is unknown. It is predicted
511 to have an SCP-like extracellular protein domain; the function of these domains is not well-
512 understood.[56] It was possible that the dramatic increase in expression of *earA* in *S. gordonii*
513 Δ*arcR* may be associated with the observed biofilm defect. However, disruption of the *earA*
514 gene in *S. gordonii* DL1, Δ*arcR* or $arcR_{Comp}$ did not affect biofilm formation compared with
515 the respective parent strain, indicating that EarA is not a major contributor to biofilm
516 development.

In conclusion, have identified ArcR as a key gene in biofilm formation by *S. gordonii*. The mechanism by which ArcR affects biofilm growth or stability remains unclear, but it is possible that changes in the co-ordination of PTS expression may be at least partly responsible. Further work is required to establish whether arginine is the only stimulus for ArcR, and to determine how arginine itself modulates *S. gordonii* biofilm development through activation of the ArcR regulatory pathway. In addition, assessment of the role of ArcR in the early stages of biofilm formation, or in biofilm formation under fluid flow, may help to shed light on the specific function of ArcR in biofilm development by *S. gordonii*.

ACKNOWLEDGEMENTS

We are very grateful to Lesley Old for expert technical assistance. We thank Lorraine Wallace, Functional Genomics, Proteomics and Metabolomics Facility, University of Birmingham, for conducting the microarray hybridisation experiments. We also acknowledge the support of Matthew Bashton, Bioinformatics Core Facility, Newcastle University, for preliminary analysis of the microarray data. The work was support by a Jeffcock and Luccock Newcastle University PhD studentship awarded to JCR. There are no conflicts of interest to declare.

REFERENCES

1. Jakubovics NS. Intermicrobial interactions as a driver for community composition and stratification of oral biofilms. *J Mol Biol.* 2015;427(23):3662-3675.

2. Nobbs AH, Jenkinson HF, Jakubovics NS. Stick to your gums: mechanisms of oral microbial adherence. *J Dent Res.* 2011;90(11):1271-1278.

3. Jakubovics NS, Yassin SA, Rickard AH. Community interactions of oral streptococci. *Adv Appl Microbiol.* 2014;87:43-110.

4. Kreth J, Merritt J, Qi F. Bacterial and host interactions of oral streptococci. *DNA Cell Biol.* 2009;28(8):397-403.

5. Zhou P, Liu J, Li X, Takahashi Y, Qi F. The sialic acid binding protein, Hsa, in *Streptococcus gordonii* DL1 also mediates intergeneric coaggregation with *Veillonella* species. *PLoS One.* 2015;10(11):e0143898.

6. Back CR, Douglas SK, Emerson JE, Nobbs AH, Jenkinson HF. *Streptococcus gordonii* DL1 adhesin SspB V-region mediates coaggregation via receptor polysaccharide of *Actinomyces oris* T14V. *Mol Oral Microbiol.* 2015;30(5):411-424.

7. McNab R, Forbes H, Handley PS, Loach DM, Tannock GW, Jenkinson HF. Cell wall-anchored CshA polypeptide (259 kilodaltons) in *Streptococcus gordonii* forms surface fibrils that confer hydrophobic and adhesive properties. *J Bacteriol.* 1999;181(10):3087-3095.

8. Bor DH, Woolhandler S, Nardin R, Brusch J, Himmelstein DU. Infective endocarditis in the U.S., 1998-2009: a nationwide study. *PLoS One.* 2013;8(3):e60033.

9. Huang X, Schulte RM, Burne RA, Nascimento MM. Characterization of the arginolytic microflora provides insights into pH homeostasis in human oral biofilms. *Caries Res.* 2015;49(2):165-176.

10. Nascimento MM, Browngardt C, Xiaohui X, Klepac-Ceraj V, Paster BJ, Burne RA. The effect of arginine on oral biofilm communities. *Mol Oral Microbiol.* 2014;29(1):45-54.

11. Nascimento MM, Liu Y, Kalra R, et al. Oral arginine metabolism may decrease the risk for dental caries in children. *J Dent Res.* 2013;92(7):604-608.

12. Ho MH, Lamont RJ, Xie H. Identification of *Streptococcus cristatus* peptides that repress expression of virulence genes in *Porphyromonas gingivalis*. *Sci Rep.* 2017;7(1):1413.

13. Sakanaka A, Kuboniwa M, Takeuchi H, Hashino E, Amano A. Arginine-ornithine antiporter ArcD controls arginine metabolism and interspecies biofilm development of *Streptococcus gordonii*. *J Biol Chem.* 2015;290(35):21185-21198.

14. Edwards AM, Grossman TJ, Rudney JD. Association of a high-molecular weight arginine-binding protein of *Fusobacterium nucleatum* ATCC 10953 with adhesion to

secretory immunoglobulin A and coaggregation with *Streptococcus cristatus*. *Oral Microbiol Immunol.* 2007;22(4):217-224.

15. Kaplan CW, Lux R, Haake SK, Shi W. The *Fusobacterium nucleatum* outer membrane protein RadD is an arginine-inhibitable adhesin required for inter-species adherence and the structured architecture of multispecies biofilm. *Mol Microbiol.* 2009;71(1):35-47.

16. Kolderman E, Bettampadi D, Samarian D, et al. L-arginine destabilizes oral multi-species biofilm communities developed in human saliva. *PLoS One.* 2015;10(5):e0121835.

17. Tada A, Nakayama-Imaohji H, Yamasaki H, et al. Cleansing effect of acidic L-arginine on human oral biofilm. *BMC Oral Health.* 2016;16:40.

18. Huang X, Zhang K, Deng M, et al. Effect of arginine on the growth and biofilm formation of oral bacteria. *Arch Oral Biol.* 2017;82:256-262.

19. Jakubovics NS, Robinson JC, Samarian DS, et al. Critical roles of arginine in growth and biofilm development by *Streptococcus gordonii*. *Mol Microbiol.* 2015;97(2):281-300.

20. Jesionowski AM, Mansfield JM, Brittan JL, Jenkinson HF, Vickerman MM. Transcriptome analysis of *Streptococcus gordonii* Challis DL1 indicates a role for the biofilm-associated *fruRBA* operon in response to *Candida albicans*. *Mol Oral Microbiol.* 2016;31(4):314-328.

21. Jakubovics NS, Gill SR, Iobst SE, Vickerman MM, Kolenbrander PE. Regulation of gene expression in a mixed-genus community: stabilized arginine biosynthesis in *Streptococcus gordonii* by coaggregation with *Actinomyces naeslundii*. *J Bacteriol.* 2008;190(10):3646-3657.

22. Sambrook J, Russell DW. *Molecular cloning : a laboratory manual.* 3rd ed. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2001.

23. Hansen MC, Palmer RJ, Jr., Udsen C, White DC, Molin S. Assessment of GFP fluorescence in cells of *Streptococcus gordonii* under conditions of low pH and low oxygen concentration. *Microbiology.* 2001;147(Pt 5):1383-1391.

24. Egland PG, Palmer RJ, Jr., Kolenbrander PE. Interspecies communication in *Streptococcus gordonii-Veillonella atypica* biofilms: signaling in flow conditions requires juxtaposition. *Proc Natl Acad Sci U S A.* 2004;101(48):16917-16922.

25. Shields RC, Mokhtar N, Ford M, et al. Efficacy of a marine bacterial nuclease against biofilm forming microorganisms isolated from chronic rhinosinusitis. *PLoS One.* 2013;8(2):e55339.

26. Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological.* 1995;57(1):289-300.

27. Del Carratore F, Jankevics A, Eisinga R, Heskes T, Hong F, Breitling R. RankProd 2.0: a refactored Bioconductor package for detecting differentially expressed features in molecular profiling datasets. *Bioinformatics.* 2017;33(17):2774-2775.

28. R Core Team. R: A language and environment for statistical computing. 2017; https://www.R-project.org/. Accessed 1st July, 2017.

29. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics.* 2016;32(18):2847-2849.

30. Manimaran S, Selby HM, Okrah K, et al. BatchQC: interactive software for evaluating sample and batch effects in genomic data. *Bioinformatics.* 2016;32(24):3836-3838.

31. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011;8(10):785-786.

32. Yu NY, Wagner JR, Laird MR, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics.* 2010;26(13):1608-1615.

33. Kristich CJ, Nguyen VT, Le T, Barnes AM, Grindle S, Dunny GM. Development and use of an efficient system for random mariner transposon mutagenesis to identify novel genetic determinants of biofilm formation in the core *Enterococcus faecalis* genome. *Appl Environ Microbiol.* 2008;74(11):3377-3386.

34. Frank KL, Guiton PS, Barnes AM, et al. AhrC and Eep are biofilm infection-associated virulence factors in *Enterococcus faecalis*. *Infect Immun.* 2013;81(5):1696-1708.

35. Barcelona-Andres B, Marina A, Rubio V. Gene structure, organization, expression, and potential regulatory mechanisms of arginine catabolism in *Enterococcus faecalis*. *J Bacteriol.* 2002;184(22):6289-6300.

36. Cho S, Cho YB, Kang TJ, Kim SC, Palsson B, Cho BK. The architecture of ArgR-DNA complexes at the genome-scale in *Escherichia coli*. *Nucleic Acids Res.* 2015;43(6):3079-3088.

37. Larsen R, Kok J, Kuipers OP. Interaction between ArgR and AhrC controls regulation of arginine metabolism in *Lactococcus lactis*. *J Biol Chem.* 2005;280(19):19319-19330.

38. Kloosterman TG, Kuipers OP. Regulation of arginine acquisition and virulence gene expression in the human pathogen *Streptococcus pneumoniae* by transcription regulators ArgR1 and AhrC. *J Biol Chem.* 2011;286(52):44594-44605.

39. Larsen R, van Hijum SA, Martinussen J, Kuipers OP, Kok J. Transcriptome analysis of the *Lactococcus lactis* ArgR and AhrC regulons. *Appl Environ Microbiol.* 2008;74(15):4768-4771.

40. de Jong A, Pietersma H, Cordes M, Kuipers OP, Kok J. PePPER: a webserver for prediction of prokaryote promoter elements and regulons. *BMC Genomics.* 2012;13:299.

640 41. Jakubovics NS, Smith AW, Jenkinson HF. Expression of the virulence-related Sca
641 (Mn^{2+}) permease in *Streptococcus gordonii* is regulated by a diphtheria toxin
642 metallorepressor-like protein ScaR. *Mol Microbiol.* 2000;38(1):140-153.

643 42. Cho BK, Federowicz S, Park YS, Zengler K, Palsson BO. Deciphering the transcriptional
644 regulatory logic of amino acid metabolism. *Nat Chem Biol.* 2012;8(1):65-71.

645 43. Zeng L, Dong Y, Burne RA. Characterization of cis-acting sites controlling arginine
646 deiminase gene expression in *Streptococcus gordonii*. *J Bacteriol.* 2006;188(3):941-949.

647 44. Liu Y, Dong Y, Chen YY, Burne RA. Environmental and growth phase regulation of the
648 *Streptococcus gordonii* arginine deiminase genes. *Appl Environ Microbiol.* 2008;74(16):5023-
649 5030.

650 45. Dong Y, Chen YY, Snyder JA, Burne RA. Isolation and molecular analysis of the gene
651 cluster for the arginine deiminase system from *Streptococcus gordonii* DL1. *Appl Environ*
652 *Microbiol.* 2002;68(11):5549-5553.

653 46. Jozefczuk J, Adjaye J. Quantitative real-time PCR-based analysis of gene expression.
654 *Methods Enzymol.* 2011;500:99-109.

655 47. Moye ZD, Zeng L, Burne RA. Fueling the caries process: carbohydrate metabolism
656 and gene regulation by *Streptococcus mutans*. *J Oral Microbiol.*
657 2014;6:10.3402/jom.v3406.24878.

658 48. Hendrickson EL, Wang TS, Dickinson BC, et al. Proteomics of *Streptococcus gordonii*
659 within a model developing oral microbial community. *BMC Microbiol.* 2012;12:211.

660 49. Loo CY, Mitrakul K, Voss IB, Hughes CV, Ganeshkumar N. Involvement of an inducible
661 fructose phosphotransferase operon in *Streptococcus gordonii* biofilm formation. *J Bacteriol.*
662 2003;185(21):6241-6254.

663 50. Tong H, Zeng L, Burne RA. The EIIABMan phosphotransferase system permease
664 regulates carbohydrate catabolite repression in *Streptococcus gordonii*. *Appl Environ*
665 *Microbiol.* 2011;77(6):1957-1965.

666 51. Kilic AO, Tao L, Zhang Y, Lei Y, Khammanivong A, Herzberg MC. Involvement of
667 *Streptococcus gordonii* beta-glucoside metabolism systems in adhesion, biofilm formation,
668 and in vivo gene expression. *J Bacteriol.* 2004;186(13):4246-4253.

669 52. Pei J, Mitchell DA, Dixon JE, Grishin NV. Expansion of type II CAAX proteases reveals
670 evolutionary origin of gamma-secretase subunit APH-1. *J Mol Biol.* 2011;410(1):18-26.

671 53. Zhang Y, Whiteley M, Kreth J, et al. The two-component system BfrAB regulates
672 expression of ABC transporters in *Streptococcus gordonii* and *Streptococcus sanguinis*.
673 *Microbiology.* 2009;155(Pt 1):165-173.

674 54. Zhang JQ, Hou XH, Song XY, Ma XB, Zhao YX, Zhang SY. ClpP affects biofilm formation
675 of *Streptococcus mutans* differently in the presence of cariogenic carbohydrates through
676 regulating *gtfBC* and *ftf*. *Curr Microbiol.* 2015;70(5):716-723.

677 55. Desai BV, Morrison DA. Transformation in *Streptococcus pneumoniae*: formation of
678 eclipse complex in a *coiA* mutant implicates CoiA in genetic recombination. *Mol Microbiol.*
679 2007;63(4):1107-1117.

680 56. Yeats C, Bentley S, Bateman A. New knowledge from old: *in silico* discovery of novel
681 protein domains in *Streptomyces coelicolor*. *BMC Microbiol.* 2003;3:3.

682

683

684 FIGURE LEGENDS

685 Figure 1. Biofilm formation by *S. gordonii* DL1 and isogenic mutants disrupted in arginine
686 metabolism or regulation genes. Biofilms were grown aerobically in TYEG medium for 18 h, and
687 quantified by measuring absorbance of crystal violet-stained cells at 562 nm. Biofilm formation
688 by Δ*arcR* single or double mutant strains was significantly lower than those of the other mutant
689 strains ($p<0.05$; indicated by asterisks). Bars represent arithmetic means of three independent
690 biological repeats, and standard error is shown. All *S. gordonii* Δ*arcR* mutant strains show a
691 significant biofilm defective phenotype, not displayed by the other mutants tested here.

692

693 Figure 2. Visualisation of biofilms formed by *S. gordonii* DL1, Δ*arcR* and *arcRComp*. Biofilms were
694 grown in TYEG medium aerobically for 18 h, stained with BacLight LIVE:DEAD stain and
695 visualised by confocal laser scanning microscopy. 3D renderings (left) or maximum projection
696 images (right) are shown. Live cells stained with Syto 9 appear green; red staining shows
697 compromised cells that have taken up propidium iodide. Images were quantified using Volocity
698 software. The quantification (mean ± standard errors) of biovolume is shown for 3 independent
699 experiments. Biovolumes of biofilms from all strains were significantly different from one
700 another.

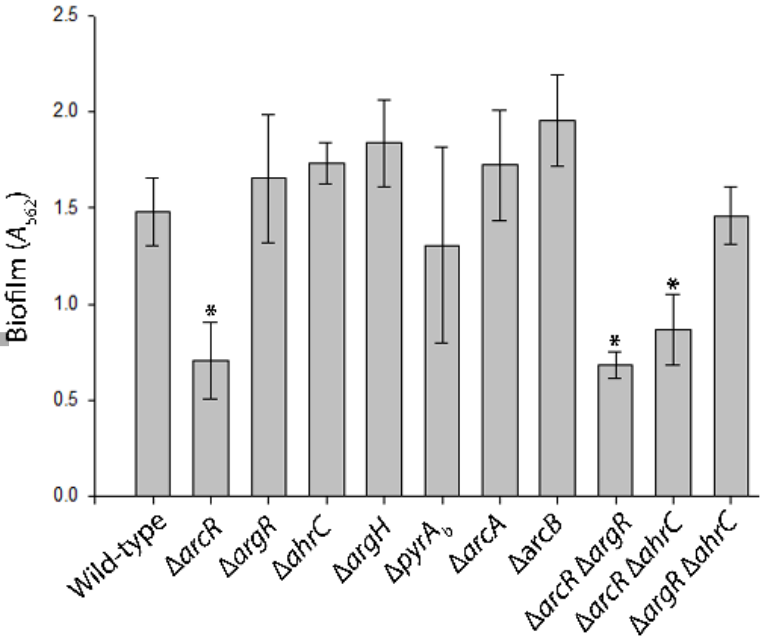
701

702 Figure 3. Validation of the *S. gordonii* DL1, Δ*arcR,* Δ*argR* and Δ*ahrC* microarrays by comparison
703 with RT-qPCR data for each strain. Gene expression was measured as a $\log_2$ fold-change in no
704 arginine compared with high arginine in *S. gordonii* DL1, Δ*arcR,* Δ*argR* and Δ*ahrC* strains. The
705 linear regression line for each correlation is shown. Values on the graphs indicate the r^2 value
706 and the slope of the line.

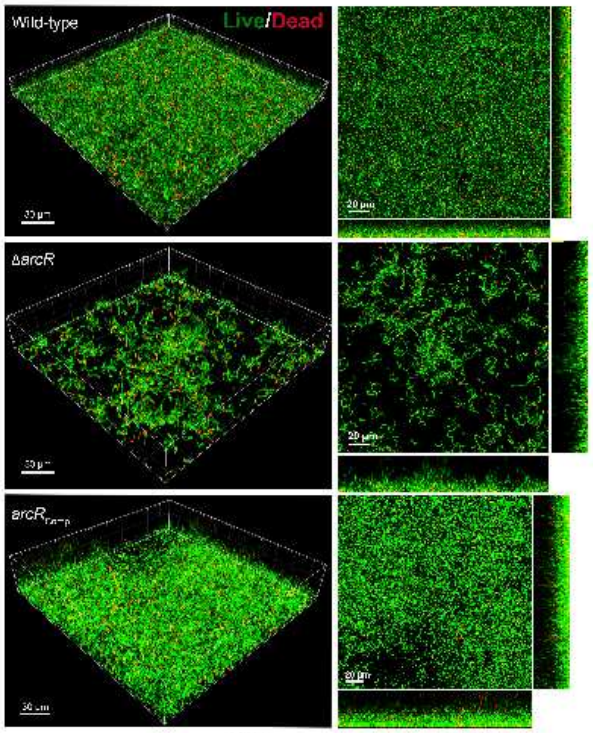
Figure 4. Heat map showing differential gene regulation in *S. gordonii* DL1, Δ*arcR*, Δ*argR* and Δ*ahrC* mutants in response to arginine. Microarray analysis was carried out on cDNA samples from *S. gordonii* strains during a shift to no arginine versus maintenance in high arginine. Colour coding represents fold change in no arginine compared with high arginine, and negative values (blue tones) indicate down-regulation. Data from *S. gordonii* DL1 and Δ*arcR* strains are not directly comparable with those from *S. gordonii* Δ*argR* and Δ*ahrC* due to concerns about batch effects between the microarray experiments, which were conducted at different times. Gray areas in *S. gordonii* Δ*argR* and Δ*ahrC* microarrays show genes that did not pass the quality threshold, and for which no reading was obtained. *QueA is a predicted S-adenosylmethionine:tRNA ribosyltransferase-isomerase.

Figure 5. Expression of *earA* and impact of *earA* gene disruption on biofilm formation by *S. gordonii*. (a) RT-qPCR analysis of *earA* expression in *S. gordonii* DL1 and Δ*arcR* under high and no arginine. Cells were grown in FMC medium, harvested and resuspended in FMC without arginine or arginine-replete FMC. Following incubation for 30 min, gene expression was measured by RT-qPCR and is expressed as $\log_2$ fold-change values. White bars represent the *earA* expression in *S. gordonii* DL1 and black bars represent *earA* expression in *S. gordonii* Δ*arcR*. Bars represent means of three independent biological replicates, and standard error is shown. (b) Impact of *earA* gene disruption on biofilm formation. Biofilms were grown in TYEG medium aerobically for 18 h, and quantified by measuring absorbance of crystal violet-stained cells at 562 nm. Bars represent means of three independent biological replicates, and error bars show standard error.

729 Significant differences between mutants and wild-type are indicated by an asterisk. Disruption
730 of *earA* did not significantly reduce biofilm formation in wild-type or Δ*arcR* genetic backgrounds.

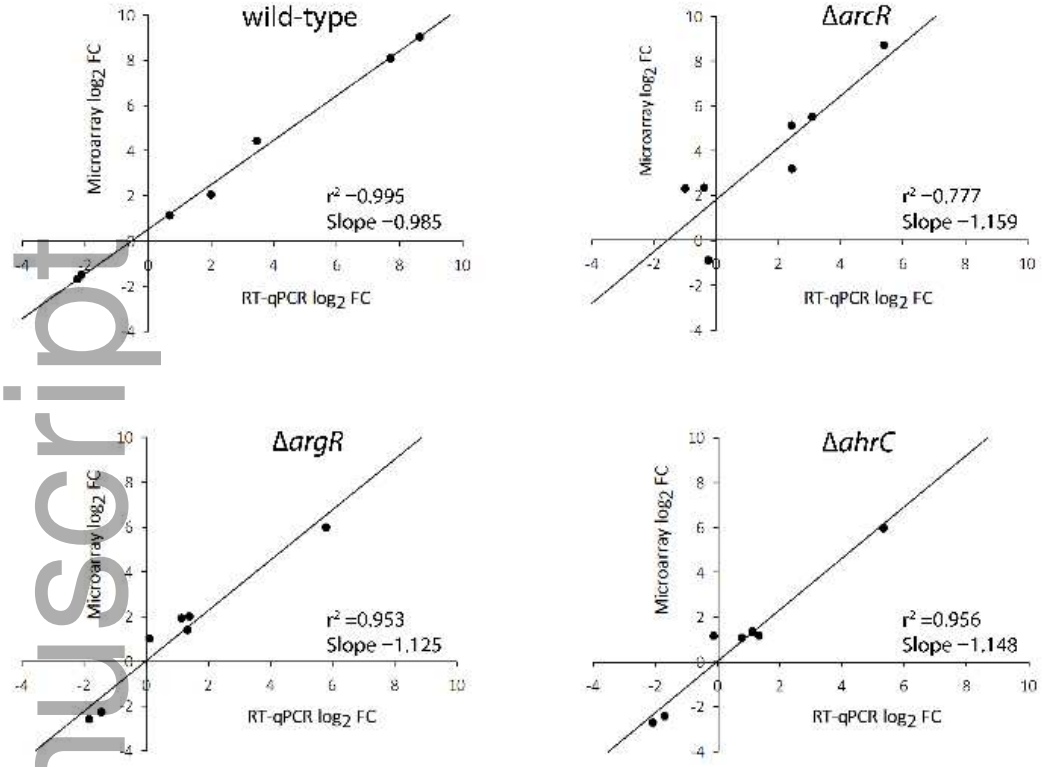


mom_12207_f1.tif

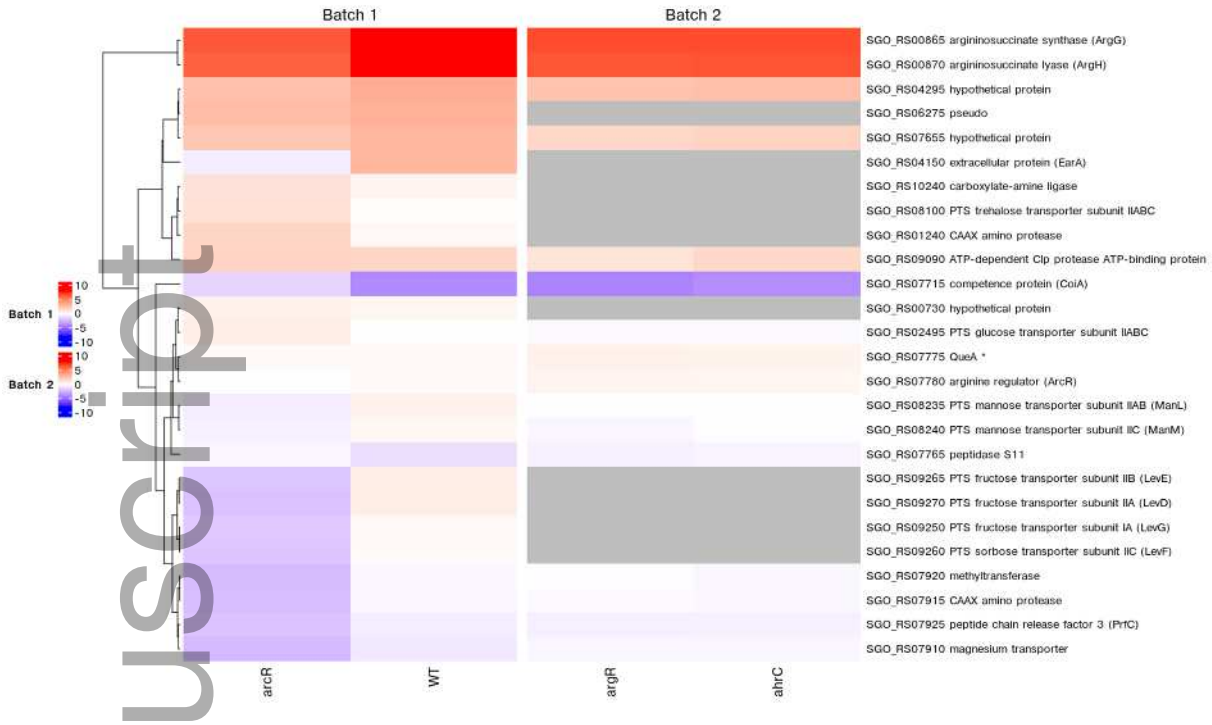


	Wild-type	Δ*arcR*	*arcR*$_{Comp}$
Biovolume (µm³/µm²)	0.80 ± 0.08	0.26 ± 0.02	1.67 ± 0.12

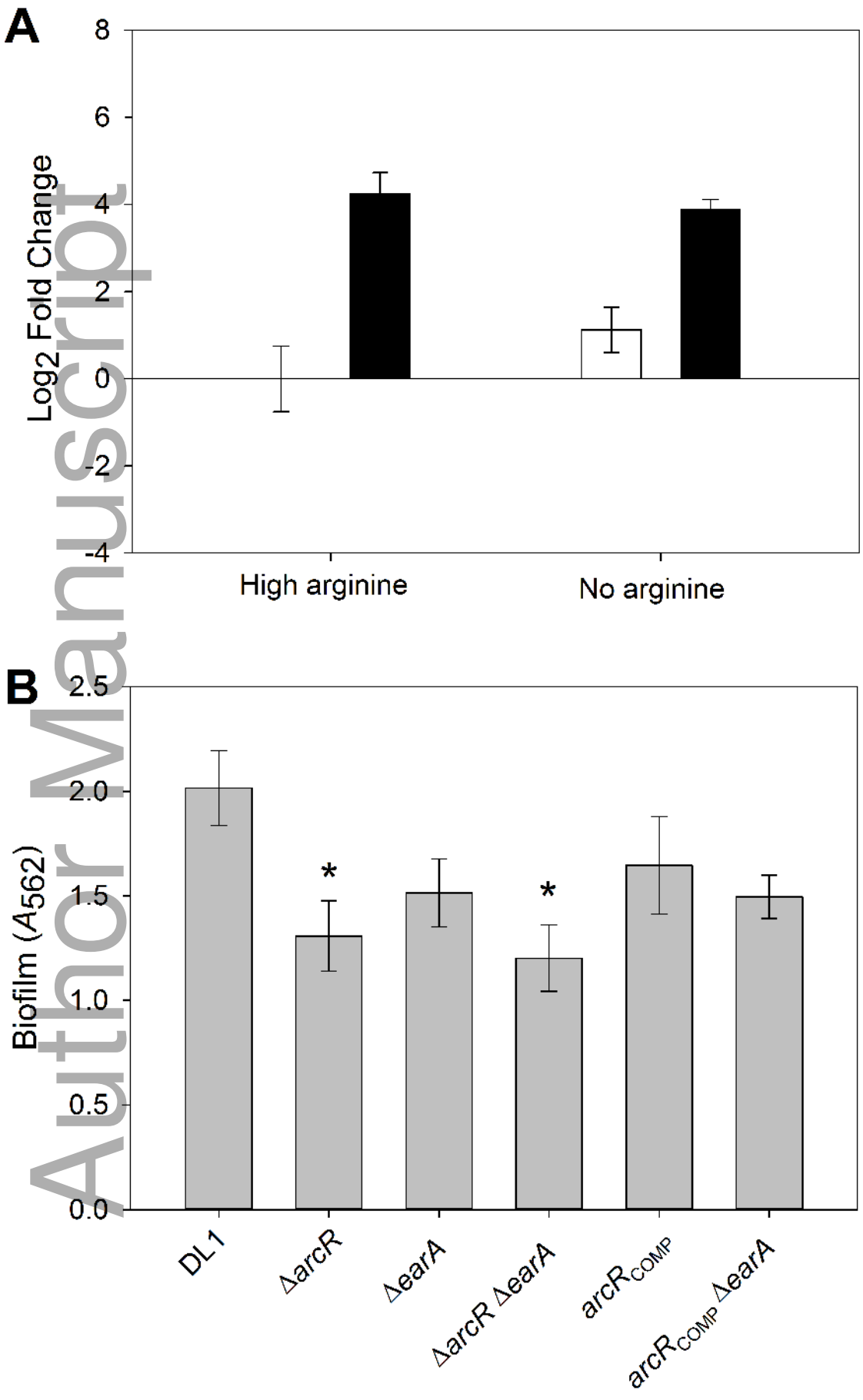
mom_12207_f2.tif



mom_12207_f3.tif



mom_12207_f4.tif



mom_12207_f5.tif