

Published in final edited form as:

FEMS Microbiol Lett. 2014 February ; 351(2): 202–208. doi:10.1111/1574-6968.12320.

Expression of *Porphyromonas gingivalis* small RNA in Response to Hemin Availability Identified Using Microarray and RNA-seq Analysis

Priscilla L. Phillips, Ann Progulske-Fox, and Nicole A. Grieshaber

Department of Oral Biology, University of Florida, Gainesville FL 32610

Abstract

There is a significant body of work suggesting that sRNA mediated post-transcriptional regulation is a conserved mechanism among pathogenic bacteria to modulate bacterial virulence and survival. *Porphyromonas gingivalis* (*Pg*) is recognized as an etiological agent of periodontitis and implicated in contributing to the development of multiple inflammatory diseases including cardiovascular disease. Using NimbleGen microarray analysis and a strand-specific method to sequence cDNA libraries of small RNA enriched *Pg* transcripts using Illumina's high-throughput sequencing technology, we identified putative sRNA and generated sRNA expression profiles in response to growth phase, hemin availability after hemin starvation, or both. We identified transcripts that mapped to intergenic sequences as well as antisense transcripts that mapped to open reading frames of the annotated genome. Overall, this approach provided a comprehensive way to survey transcriptional activity to discover functionally linked RNA transcripts, responding to specific environmental cues, that merit further investigation.

INTRODUCTION

Porphyromonas gingivalis (*Pg*) is a non-motile asaccharolytic Gram-negative obligate anaerobe that requires heme (iron and protoporphyrin IX) for growth. It is recognized as a major pathogen of severe adult periodontitis (Griffen, *et al.*, 1998) and implicated in systemic inflammatory conditions including cardiovascular disease (Desvarieux, *et al.*, 2005, Rosenfeld & Campbell, 2011, Belstrom, *et al.*, 2012), rheumatoid arthritis (Hitchon, *et al.*, 2010, Routsias, *et al.*, 2011), preeclampsia (Barak, *et al.*, 2007, Lachat, *et al.*, 2011), and preterm delivery (Lachat, *et al.*, 2011). *Pg* is able to invade multiple cell types, including human vascular and oral cell lines (Lamont, *et al.*, 1995, Deshpande, *et al.*, 1999, Progulske-Fox, *et al.*, 1999, Dorn, *et al.*, 2000). Invasion of host cells is very likely a key virulence factor for this bacterium as it provides 1) a "privileged niche" with access to host protein (nutritional) and iron substrates, 2) a sequestration from the humoral and cellular immune response, and 3) a means for persistence that is essential for a chronic pathogen. As a host associated pathogen that can cause disease pathology using various adaptive mechanisms, whether forming tissue disseminating biofilm communities in the oral cavity and other sites in the body or residing in host cells as an intracellular pathogen, *Pg* must be able to sense

and rapidly respond to changes in its environment. For example, heme limitation in subgingival plaque transitions to heme excess during active disease *in vivo*. Studies have shown that the expression of virulence factors increase after hemin starvation in *Pg* (Genco, *et al.*, 1993, Kesavalu, *et al.*, 2003, Dashper, *et al.*, 2009). Furthermore, studies assessing gene expression profiles of *Pg* strain W83 have suggested that there is a regulatory switch that initiates periodontal pathogenesis during mid-log phase under hemin limitation after hemin starvation (Kiyama-Kishikawa, *et al.*, 2005). However, the regulatory systems *Pg* uses to rapidly respond to these and other environmental cues remains unclear.

Regulation of gene expression by small non-coding regulatory RNA (sRNA) is an exciting and rapidly growing field in biology. This mechanism of gene regulation is now recognized as a common mechanism employed by bacteria to rapidly respond to dynamic environmental cues. Furthermore, there is a significant body of work suggesting that sRNA mediated post-transcriptional regulation is a conserved mechanism among pathogenic bacteria to modulate bacterial virulence and survival (Bardill & Hammer, 2012, Mann, *et al.*, 2012, Ortega, *et al.*, 2012). Most chromosomally located sRNAs currently characterized in bacteria are regulated by Hfq, a sRNA chaperone conserved across many bacterial species (Beisel & Storz, 2010, Richards & Vanderpool, 2011, Storz, *et al.*, 2011, Vanderpool, *et al.*, 2011, Bossi, *et al.*, 2012, Richter & Backofen, 2012, Shao & Bassler, 2012, Sobrero & Valverde, 2012). However, *Pg* is an Hfq negative bacterium, with as yet uncharacterized sRNA regulatory systems. Moreover, the only sRNA in the Bacteroidetes phylum currently characterized is RteR (Waters & Salyers, 2012).

MATERIALS AND METHODS

Bacterial culture

Pg strain W83 was cultured overnight in hemin rich (5ug/ml) modified TSB media (TBS with 5 ug/ml yeast extract, 0.5 ug/ml L-cysteine hydrochloride, 1ug/ml Vitamin K1) prior to hemin starvation for two days under anaerobic conditions at 37°C. The cultures were split and cultured in hemin rich modified TSB or hemin limiting (0.001 ug/ml) modified TSB media to mid-log or stationary phase. The cells were collected and processed to extract RNA.

RNA extraction

Pg W83 small RNA enriched extracts (<200 nt) were prepared using mirVana™ miRNA isolation kit (Ambion, Foster City, CA USA). As bacterial regulatory sRNAs average 100 nt in size, the resulting small RNA enriched samples greatly reduces the complexity of the sample and minimizes the “noise” of mRNA expression, thus simplifying both microarray and sequencing readout. The improved resolution of this method was verified by the ability to distinguish two *Pg* tRNAs found in the same intergenic sequence (IGS) separated by only 30 bp.

Microarray cDNA library construction and analysis

Microarray analysis was performed on libraries constructed from three of the four culture conditions: two-day hemin starved bacterial cells, subsequently cultured to mid-log or

stationary in hemin rich media or subsequently cultured to stationary in hemin limiting media. DNase treated small RNA from each condition was C-tailed at the 3' end using poly(A)polymerase (TakaRa Bio Inc, Shiga, Japan). The first strand of the cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen, Carlsbad, CA USA) and a custom oligonucleotide containing a 3' oligo G tail, primer (P1) sequence, and a NotI site at the 5' end. A double stranded DNA linker (L/P2) was ligated to the double stranded RNA/cDNA molecules then digested with NotI restriction endonuclease to remove linkers that ligated to the 3' ends. P1 and P2 primers and Phusion high-fidelity PCR Mastermix (NewEnglandBiolabs, Ipswich, MA USA) were then used to generate double stranded cDNA to probe custom designed, high density, whole-genome oligonucleotide arrays synthesized by NimbleGen Systems (Roche NimbleGen Inc, Madison, WI USA). Oligonucleotide sequences are provided in the supplement (Table S1). The arrays contained 65–75mer oligonucleotide probes, synthesized in duplicate, on glass slides using a maskless array synthesizer. Starting at bp 1 on the *Pg* W83 genome (Genebank accession: AE-015924), 180,071 probes overlapping every 10–13bp was generated to provide full coverage of the 2.34 Mb genome.

Illumina cDNA library construction and analysis

In order to obtain sequence information of the small RNA transcripts while confirming our microarray data we used a strand-specific method to sequence cDNA libraries of small RNA enriched *Pg* W83 transcripts using Illumina's high-throughput sequencing technology. DNase treated small RNA enriched samples from each condition (n=2) were used to create eight cDNA libraries using various combinations of 16 custom primers each containing a unique 6 bp barcode for paired-end Illumina sequencing. The mean size of the *Pg* RNA molecules in the DNase treated small RNA enriched samples were found to be approximately 70 nt (Bioanalyzer; Agilent Technologies, Santa Clara, CA USA). It is generally known that bacterial rRNA transcripts are processed yielding 5' monophosphate rRNAs while bacterial mRNA has a 5' triphosphate cap and a relatively short half-life. To prepare the full-length cDNA Libraries of *Pg* sRNA for paired-end Illumina sequencing, DNase treated sRNA enriched samples were treated with Terminator 5'-Phosphate-dependent exonuclease (TIP) (Epicentre Biotechnologies, Madison, WI USA) to selectively remove ribosomal RNA. TIP will also digest 5' phosphorylated ssDNA and dsDNA). The samples were then treated with Tobacco Acid Pyrophosphatase (TAP) (Epicentre Biotechnologies) to remove 5' caps and convert triphosphate RNA to 5' monophosphorylated RNA. A 46 bp custom designed RNA oligonucleotide adaptor (5'-IlluminaSeq-NNNN-BarCode-3') was then ligated to the 5' end of the sRNA samples using T4 RNA ligase (Epicentre Biotechnologies). The sRNA was then C-tailed at the 3' end using poly(A)polymerase (TakaRa). The first strand of the cDNA was synthesized using a 59 bp custom DNA oligonucleotide primer (5'-IlluminaSeq-NNNN-Barcode-oligoG(15)-3') and Superscript III reverse transcriptase (Invitrogen). *E. coli* RNase H was used to nick/ degrade the RNA strand leaving single strand cDNA. PCR was used to amplify the single strand cDNA samples to make double strand cDNA libraries using custom primers P1 and P2 and Phusion high-fidelity PCR Mastermix (NewEnglandBiolabs) for Illumina sequencing. Oligonucleotide sequences are provided in the supplement (Table S2). The majority of the final cDNA product was observed to be approximately 150–250 nt long

using denaturing 5 % Urea-PAGE electrophoresis. The quality and average length of the cDNA product was assessed using Bioanalyzer (Agilent) before the eight cDNA libraries were pooled for direct Illumina sequencing. The cDNA pool was loaded on two lanes to obtain sequencing run replicates.

Each cDNA library generated for RNA-seq analysis was designed to include unique 6 bp barcodes, specific to the 5' or 3' end of the RNA transcripts, in order to determine the strand specificity of each mapped small RNA, and to differentiate libraries during data analysis (Table S3). The raw sequencing read data were converted to fastq format files and sorted by lane, 3' or 5' end, and barcode. The sequencing reads were mapped against the genomic reference sequences using the software MAQ (Li, *et al.*, 2008). The MAQ mapped and Mapview generated text output files were converted into SQLite (language) database files and the data was mined in Base (Mac OSX application). Based on the length of the adaptors and depth of the reads (35 cycles), none of the 3' reads could be stringently mapped. However, the 5' reads could be mapped. We identified the 5' end sequence of putative sRNA and generated sRNA expression profiles in response to growth phase and/or hemin availability after hemin starvation (Table 1, Table S3, Supplement data).

Northern analysis

Biotinylated probes to sRNA 42 and 101 were prepared from PCR fragments generated from genomic *Pg* W83 using the primers described in Table S4. The PCR products were biotinylated using a BrightStar Psoralen-Biotin Nonisotopic Labeling Kit (Ambion). Small RNA enriched *Pg* RNA was fractionated in 10% TBE-urea acrylamide gels and transferred by electro-blotting onto BrightStar-Plus Nylon membranes (Ambion). Probe hybridization onto the blots was carried out in UltraHyb (Ambion) medium overnight at 42°C. Blots were washed using NorthMax Low Stringency Buffer (Ambion) at room temperature (2 × 15 min) and with NorthMax High Stringency Buffer (Ambion) at 42°C (2 × 30 min). Northern blots were developed using alkaline phosphatase and streptavidin according to manufacturer's protocols (BrightStar BioDetect, Ambion).

RESULTS AND DISCUSSION

Prokaryotic posttranscriptional regulation by and large employs mechanisms that function to stabilize or destabilize mRNA. Regulation of gene expression by small non-coding regulatory RNAs (sRNA) is an emerging paradigm in microbiology. The current understanding of sRNA regulatory systems, described in several recent reviews (Vogel & Wagner, 2007, Sharma & Vogel, 2009, Beisel & Storz, 2010, Richards & Vanderpool, 2011, Storz, *et al.*, 2011, Vanderpool, *et al.*, 2011, Bossi, *et al.*, 2012, Richter & Backofen, 2012, Shao & Bassler, 2012, Sobrero & Valverde, 2012), generalize that most bacterial sRNAs (like eukaryotic microRNA) regulate expression at the posttranscriptional level, usually by binding multiple target mRNA(s). Typically, the size of bacterial sRNAs ranges from 50 nt to 250 nt. Bacterial sRNA transcripts typically contain their own promoters and Rho-independent terminators. Cis-encoded sRNAs are fully complementary antisense regulatory RNAs encoded within the target mRNA's ORF, leader, or trailer sequence. Trans-encoded (intergenic) sRNAs bind to target mRNA(s) by imperfect base pairing, dependent on the sRNA's secondary structure. A single sRNA may repress and/or activate translation. Most

chromosomally located sRNAs currently characterized in bacteria are trans-encoded and their mRNA target(s) are not simply determined by sequence analysis or relative loci within the genome.

Pg encounters many different environmental conditions during colonization and growth within the host that require integration of sensory input and coordination of complex mechanisms of gene regulation. We hypothesized that *Pg* employ sRNA regulatory elements to rapidly respond to environmental cues. We used NimbleGen microarray analysis to identify transcripts found in cDNA libraries generated from small RNA enriched *Pg* W83 samples expressed in response to growth phase and/or hemin availability, after two days of hemin starvation. These conditions were selected based on published studies that suggest that periodontal pathogenesis is initiated during mid-log phase under hemin limitation after hemin starvation (Genco, *et al.*, 1993, Kesavalu, *et al.*, 2003, Kiyama-Kishikawa, *et al.*, 2005, Dashper, *et al.*, 2009). Following hybridization, microarrays were scanned and the median signal intensity for each probe on the array calculated by NimbleGen. A region was considered to contain a small transcript if four or more probes in succession were greater than 10-fold above background (signal intensity >10,000). This corresponds to RNA of about 75 bp or larger. Each small transcript was analyzed for the presence of a transcription terminator based on the prediction program TransTermHP (<http://transterm.cbcb.umd.edu/>). Transcripts containing a predicted terminator were analyzed to determine the location on the genome, the presence of an open reading frame (ORF), and its length estimated based on the first and last positive probe. Only those transcripts larger than 70 nt containing a predicted terminator and did not contain an ORF on the same strand were considered candidate sRNA. Approximately 8 % of the 180,071 probes covering the entire 2.34 Mbp *Pg* W83 genome were highly expressed. Thirty-seven putative sRNAs were identified whose terminators could be predicted and is summarized as follows (Table S5): Thirty mapped to IGSs. Three mapped to IGSs with slight ORF 5' overlap on the opposite strand based on terminator prediction. Four mapped to ORF on the opposite strand based on terminator prediction. All but thirteen putative sRNAs were regulated in response to limited hemin availability after hemin starvation. Those transcripts identified by microarray that also had significant counts detected by RNA-seq analysis are indicated in Table S5.

Mobile genetic elements contribute to the genetic plasticity of bacterial species (Tribble, *et al.*, 2007, Naito, *et al.*, 2011, Watanabe, *et al.*, 2013). There are numerous short and long mobile elements including complete and degenerative (truncated) insertion sequences (IS) and transposons found throughout the genome of *Pg*. Studies to date have shown that *Pg* has a high degree of genetic variation among strains. For example, conjugal transfer is proposed to be the mechanism by which *Pg* undergoes allele exchange, contributing to genetic variation (Tribble, *et al.*, 2007, Naito, *et al.*, 2011). Using RNA-seq analysis, we found that all five ISPg3 and all ten ISPg4 transposase ORFs encoded highly expressed cis-encoded antisense transcripts to the 5' untranslated region (UTR) of these genes with a 2nt overlap of coding region. We also identified cis-encoded antisense transcripts to both ORFs (PG0827, PG1446) encoding putative MatE family multidrug efflux pumps (Figure 1). These genes were located within regions flanked by integrases. Sequence analyses indicate that these regions are located within large multi-gene transposons. We also identified highly

expressed transcripts of CRISPR regulatory sRNA trans-encoded in the IGS upstream of CRISPR-associated (CAS) gene arrays. A recent study analyzing the spacer content among the genomes of 60 *Pg* isolates has indicated that IS mediated transposition may be limited by CRISPR interference (Watanabe, *et al.*, 2013). Spacer analysis showed a high degree of similarity to *Pg* genome sequence, most of which matched sequences within ORFs. We found that inverted repeats of ISPg1 type transposases appear associated with CRISPRs in W83. The identification of these highly expressed sRNA transcripts indicate another mechanism *Pg* employ to limit genetic exchange among a population with a genotype clearly predisposed to a high degree of genetic exchange. CRISPR expression was only detected under logarithmic growth and not during stationary, indicating that *Pg* uses this mechanism to limit genetic exchange only during active growth phase. In contrast, the cis-encoded antisense transcripts to ISPg3s and ISPg4s UTRs and the MatE efflux pump ORFs were expressed under all conditions assayed.

Overall, we identified 186 sites containing 5' end sequence reads, other than tRNA or rRNA, with significant counts (>5) that mapped to IGSs or were clearly antisense to identified ORFs on the *Pg* W83 genome. Transcripts from approximately 49 % of these sites (not encoding tRNA/rRNA) were highly expressed (>10 counts). Of the sequenced reads within the pooled cDNA, consisting of 6 libraries, approximately 85 % of the mapped reads were to tRNA/rRNA for all culture conditions (i.e., 107773/141115 of library containing 5' reads with barcode #1; Supplement data). Under these parameters, read depth of the sample was limited, thus greater than 10 counts of the same transcript was considered highly expressed. The expression profile was determined by normalizing the sequencing read counts of each analyzed transcript to the relative tRNA/rRNA counts within each library. Twenty-three putative sRNAs were clearly regulated in response to growth phase. Twenty-two putative sRNAs located within IGSs were clearly up-regulated in response to hemin after hemin starvation (Table S6). Northern analysis was performed on several of the small RNA transcripts that had high microarray signal intensity. For example, a transcript (sRNA 101) approximately 118nt, encoded on the negative strand of the IGS between PG2089 and PG2090 (both encoded on the positive strand), was detected by microarray. This transcript had a maximal count of 88 5'-end sequence reads (ATAAGCCGCACTGTTAGATCGGGG) as detected by RNA-seq analysis. Northern analysis confirmed expression of this IGS encoded transcript (Figure 2). In contrast, microarray analysis detected a transcript mapping to the positive strand of the IGS between PG0715 and PG0717 (sRNA 42) that had high signal intensity, while none was detected by RNA-seq nor Northern analysis (Figure 2). We believe RNA-seq analysis is an improved method to determine relative expression profiles of microbial sRNA compared to microarray. We propose that an additional alteration in the methodology to reduce tRNA in library generation would improve sRNA transcript identification by RNA-seq analysis, but may eliminate the ability to generate comparable relative expression profiles normalized to tRNA read counts.

By employing RNA-seq analysis to identify and characterize the expression profile of small RNA transcripts expressed in *Pg*, we have generated a list of possible regulatory transcripts. This technology will now allow further dissection of the factors, whether virulence related

or associated with other survival pathways, controlling gene expression in *Pg* in response to environmental signals (i.e., hemin availability or growth phase).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The data reported in this manuscript were generated from a NIH grant funded study (#R21 DE01986).

References

- Barak S, Oettinger-Barak O, Machtei EE, Sprecher H, Ohel G. Evidence of periopathogenic microorganisms in placentas of women with preeclampsia. *Journal of Periodontology*. 2007; 78:670–676. [PubMed: 17397314]
- Bardill JP, Hammer BK. Non-coding sRNAs regulate virulence in the bacterial pathogen *Vibrio cholerae*. *RNA biology*. 2012; 9:392–401. [PubMed: 22546941]
- Beisel CL, Storz G. Base pairing small RNAs and their roles in global regulatory networks. *FEMS microbiology reviews*. 2010; 34:866–882. [PubMed: 20662934]
- Belstrom D, Damgaard C, Nielsen CH, Holmstrup P. Does a causal relation between cardiovascular disease and periodontitis exist? *Microbes and infection / Institut Pasteur*. 2012; 14:411–418. [PubMed: 22200601]
- Bossi L, Schwartz A, Guillemardet B, Boudvillain M, Figueroa-Bossi N. A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. *Genes & development*. 2012; 26:1864–1873. [PubMed: 22895254]
- Dashper SG, Ang CS, Veith PD, et al. Response of *Porphyromonas gingivalis* to heme limitation in continuous culture. *J Bacteriol*. 2009; 191:1044–1055. [PubMed: 19028886]
- Deshpande RG, Khan M, Genco CA. Invasion Strategies of the Oral Pathogen *Porphyromonas gingivalis*: Implications for Cardiovascular Disease. *Invasion Metastasis*. 1999; 18:57–69. [PubMed: 10364686]
- Desvarieux M, Demmer RT, Rundek T, Boden-Albala B, Jacobs DR Jr, Sacco RL, Papapanou PN. Periodontal microbiota and carotid intima-media thickness: the Oral Infections and Vascular Disease Epidemiology Study (INVEST). *Circulation*. 2005; 111:576–582. [PubMed: 15699278]
- Dorn BR, Burks JN, Seifert KN, Progulsk-Fox A. Invasion of endothelial and epithelial cells by strains of *Porphyromonas gingivalis*. *FEMS Microbiol Lett*. 2000; 187:139–144. [PubMed: 10856647]
- Genco, CA.; Lassister, M.; Forgn, RY.; Basil, MO. Characterization of Tn4351 generated mutants of *P. gingivalis*. Vol. D123. Atlanta, Georgia: 1993.
- Griffen AL, Becker MR, Lyons SR, Moeschberger ML, Leys EJ. Prevalence of *Porphyromonas gingivalis* and Periodontal Health Status. *J Clin Microbiol*. 1998; 36:3239–3242. [PubMed: 9774572]
- Hitchon CA, Chandad F, Ferucci ED, et al. Antibodies to *Porphyromonas gingivalis* are associated with anticitrullinated protein antibodies in patients with rheumatoid arthritis and their relatives. *The Journal of rheumatology*. 2010; 37:1105–1112. [PubMed: 20436074]
- Kesavalu L, Holt SC, Ebersole JL. In vitro environmental regulation of *Porphyromonas gingivalis* growth and virulence. *Oral Microbiol Immunol*. 2003; 18:226–233. [PubMed: 12823798]
- Kiyama-Kishikawa M, Hiratsuka K, Abiko Y. Gene expression profiling and characterization under hemin limitation in *Porphyromonas gingivalis*. *Journal of oral science*. 2005; 47:191–197. [PubMed: 16415563]
- Lachat MF, Solnik AL, Nana AD, Citron TL. Periodontal disease in pregnancy: review of the evidence and prevention strategies. *The Journal of perinatal & neonatal nursing*. 2011; 25:312–319. [PubMed: 22071614]

- Lamont RJ, Chan A, Belton CM, Izutsu KT, Vasel D, Weinberg A. *Porphyromonas gingivalis* invasion of gingival epithelial cells. *Infect Immun*. 1995; 63:3878–3885. [PubMed: 7558295]
- Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome research*. 2008; 18:1851–1858. [PubMed: 18714091]
- Mann B, van Opijnen T, Wang J, et al. Control of virulence by small RNAs in *Streptococcus pneumoniae*. *PLoS pathogens*. 2012; 8:e1002788. [PubMed: 22807675]
- Naito M, Sato K, Shoji M, Yukitake H, Ogura Y, Hayashi T, Nakayama K. Characterization of the *Porphyromonas gingivalis* conjugative transposon CTnPg1: determination of the integration site and the genes essential for conjugal transfer. *Microbiology*. 2011; 157:2022–2032. [PubMed: 21527470]
- Ortega AD, Gonzalo-Asensio J, Garcia-del Portillo F. Dynamics of Salmonella small RNA expression in non-growing bacteria located inside eukaryotic cells. *RNA biology*. 2012; 9:469–488. [PubMed: 22336761]
- Progulske-Fox A, Kozarov E, Dorn B, Dunn W Jr, Burks J, Wu Y. *Porphyromonas gingivalis* virulence factors and invasion of cells of the cardiovascular system. *J Periodontal Res*. 1999; 34:393–399. [PubMed: 10685367]
- Richards GR, Vanderpool CK. Molecular call and response: the physiology of bacterial small RNAs. *Biochimica et biophysica acta*. 2011; 1809:525–531. [PubMed: 21843668]
- Richter AS, Backofen R. Accessibility and conservation: general features of bacterial small RNA-mRNA interactions? *RNA biology*. 2012; 9:954–965. [PubMed: 22767260]
- Rosenfeld ME, Campbell LA. Pathogens and atherosclerosis: update on the potential contribution of multiple infectious organisms to the pathogenesis of atherosclerosis. *Thrombosis and Haemostasis*. 2011; 106:858–867. [PubMed: 22012133]
- Routsias JG, Goules JD, Goules A, Charalampakis G, Pikazis D. Autopathogenic correlation of periodontitis and rheumatoid arthritis. *Rheumatology*. 2011; 50:1189–1193. [PubMed: 21343168]
- Shao Y, Bassler BL. Quorum-sensing non-coding small RNAs use unique pairing regions to differentially control mRNA targets. *Molecular Microbiology*. 2012; 83:599–611. [PubMed: 22229925]
- Sharma CM, Vogel J. Experimental approaches for the discovery and characterization of regulatory small RNA. *Current opinion in microbiology*. 2009; 12:536–546. [PubMed: 19758836]
- Sobrero P, Valverde C. The bacterial protein Hfq: much more than a mere RNA-binding factor. *Critical reviews in microbiology*. 2012; 38:276–299. [PubMed: 22435753]
- Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. *Molecular cell*. 2011; 43:880–891. [PubMed: 21925377]
- Tribble GD, Lamont GJ, Progulske-Fox A, Lamont RJ. Conjugal transfer of chromosomal DNA contributes to genetic variation in the oral pathogen *Porphyromonas gingivalis*. *Journal of bacteriology*. 2007; 189:6382–6388. [PubMed: 17573478]
- Vanderpool CK, Balasubramanian D, Lloyd CR. Dual-function RNA regulators in bacteria. *Biochimie*. 2011; 93:1943–1949. [PubMed: 21816203]
- Vogel J, Wagner EG. Target identification of small noncoding RNAs in bacteria. *Current opinion in microbiology*. 2007; 10:262–270. [PubMed: 17574901]
- Watanabe T, Nozawa T, Aikawa C, Amano A, Maruyama F, Nakagawa I. CRISPR regulation of intraspecies diversification by limiting IS transposition and intercellular recombination. *Genome biology and evolution*. 2013; 5:1099–1114. [PubMed: 23661565]
- Waters JL, Salyers AA. The small RNA RteR inhibits transfer of the *Bacteroides* conjugative transposon CTnDOT. *Journal of bacteriology*. 2012; 194:5228–5236. [PubMed: 22821972]

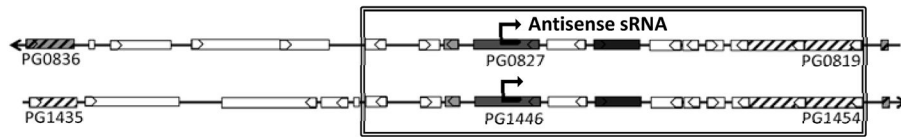


Figure 1. Antisense sRNA to ORFs PG0827 and PG1446 are located within putative conjugative transposons

The 5' end (GTCCAAGTACATTCTTATCGTTGG) of the identified antisense transcripts is centrally located within their respective *Pg* ORFs encoding MatE family multi-drug efflux pumps. PG0827 and PG1446 are indicated by red shading. ORFs encoding ISPg1 transposase are indicated by blue shading. Truncated ORFs are indicated by grey shading. ORFs encoding integrases are indicated by crosshatching. The homologous regions between these loci were directionally aligned and box outlined in this graphic representation.

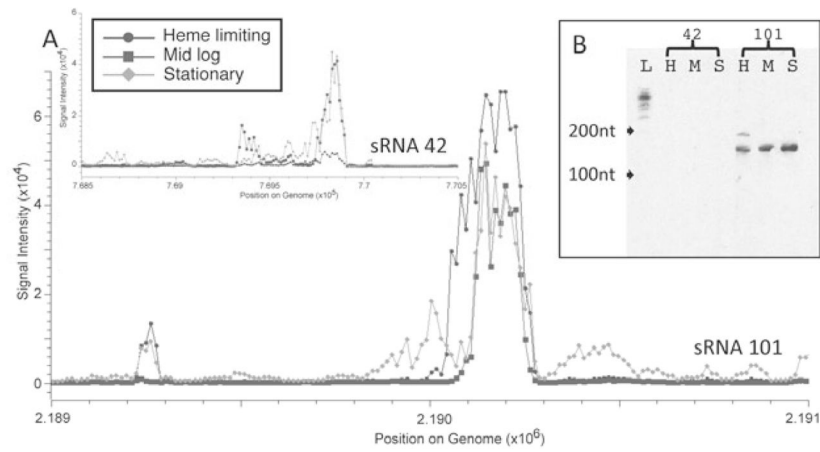


Figure 2. *Pg* small RNA transcripts 42 and 101

Microarray signal intensity graphs (A) and Northern blot autoradiographic image (10min exposure) (B). The lanes were loaded with molecular weight marker (L) and small RNA enriched, DNase digested, RNA isolated from hemin-limiting stationary (H), hemin-excess mid-log (M), or hemin-excess stationary (S), *Pg* culture condition (after two day hemin pre-starvation).

Table 1

List of expressed small transcripts mapped to IGS (>100 reads)

Position	Sequence*	Strand	Expression/ Response Condition	Illumina counts (Max) [†]
1181291	GTACAGATAGCAGCCCACGCCTAT	Negative	All	22224
2279237	TCCTTAGCCTCCAAATCGTGCAGG	Negative	All	972
1299082	GTGTTTTTCATGGTATTAGATTTT	Positive	All	803
615301	GTGTTTTTCATGGTATTAGATTTT	Positive	All	791
284667	GTTACGGATGGATGGAGGAGGCGA	Positive	All	261
185224	AGGCGATTCTTTAAATGAGACTT	Negative	Growth Phase	161
232810	TTTGAAGGAGGCACCGCTCGGGAG	Positive	Heme	136
2325144	TCGTTAGTTACATGCTCTGAAAAC	Negative	All (Heme)	110

* Sequence relative to the *Pg* W83 reference genome. If encoded on the negative strand, the 5' end sequence of the sRNA transcript is the complement of the sequence listed.

[†] Extremely high reads may possibly result from processed product of a single transcript of a highly expressed multi-gene operon.