Probing Bacterial Metabolism during Infection Using High-Resolution Transcriptomics

Peter Jorth,a Urvish Trivedi,b Kendra Rumbaugh,b and Marvin Whiteley*a

Section of Molecular Genetics and Microbiology, Institute of Cell and Molecular Biology, The University of Texas at Austin, Austin, Texas, USA; Department of Surgery, Texas Tech University Health Sciences Center, Lubbock, Texas, USA

A fundamental aspect of most infectious diseases is the need for the invading microbe to proliferate in the host. However, little is known about the metabolic pathways required for pathogenic microbes to colonize and persist in their hosts. In this study, we used RNA sequencing (RNA-seq) to generate a high-resolution transcriptome of the opportunistic pathogen Aggregatibacter actinomycetemcomitans in vivo. We identified 691 A. actinomycetemcomitans transcriptional start sites and 210 noncoding RNAs during growth in vivo and as a biofilm in vitro. Compared to in vitro biofilm growth on a defined medium, ~14% of the A. actinomycetemcomitans genes were differentially regulated in vivo. A disproportionate number of genes coding for proteins involved in metabolic pathways were differentially regulated in vivo, suggesting that A. actinomycetemcomitans in vivo metabolism is distinct from in vitro growth. Mutational analyses of differentially regulated genes revealed that formate dehydrogenase H and fumarate reductase are important A. actinomycetemcomitans fitness determinants in vivo. These results not only provide a high-resolution genomic analysis of a bacterial pathogen during in vivo growth but also provide new insight into metabolic pathways required for A. actinomycetemcomitans in vivo fitness.

Over 100 years ago, Louis Pasteur et al. emphasized the importance of understanding the metabolic processes involved in bacterial colonization and persistence during infection (1). However, progress in this area has proven challenging, in part because of the difficulty in characterizing the nutritional content of infection sites. One approach that has been successful in probing the growth environment of the infection site is transcriptomics using microarrays and RNA sequencing (RNA-seq). Microarrays have yielded insights into the carbon sources and metabolic pathways used by pathogens during infection (2–5), although these studies are technically challenging because of the large amounts of host RNA often present in disease samples. In contrast, high-throughput RNA-seq provides a robust tool for global gene expression analyses in vivo since it can be performed with small amounts of RNA and bacterial and host gene expression can be separated computationally (6). However, to date, very few in vivo RNA-seq studies of mammalian pathogens have been performed. One study revealed significant insight into the in vivo physiology of Vibrio cholerae during infection (7); however, the undefined medium used as an in vitro control limited the conclusions that could be drawn about the metabolic pathways active in vivo.

Our laboratory has studied nutrition in numerous pathogenic bacteria, including the opportunistic human pathogen Aggregatibacter actinomycetemcomitans (8–14). A. actinomycetemcomitans is a Gram-negative, facultative anaerobe that resides in the oral cavities of humans and old-world primates. Specifically, A. actinomycetemcomitans colonizes the subgingival crevice, defined as the area around the tooth bounded by the gingival epithelium on one side and the tooth surface on the other. A. actinomycetemcomitans is the proposed causative agent of localized aggressive periodontitis (15, 16), an acute disease characterized by massive tissue destruction and tooth loss. A. actinomycetemcomitans also causes extraoral infections, including abscess infections (17–19), and our laboratory has used a murine abscess infection model to characterize A. actinomycetemcomitans genes required for growth in vivo (14). In this study, we used high-resolution transcriptomics to examine the physiology of A. actinomycetemcomitans during growth in the murine abscess. These analyses revealed significant insights into the transcriptional organization of the A. actinomycetemcomitans genome, including transcription start sites (TSS) and noncoding RNAs (ncRNAs). In addition, numerous metabolic genes displaying enhanced transcription during in vivo growth were identified. Targeted mutagenesis of these differentially regulated genes revealed roles for fermentative metabolism and anaerobic respiration for in vivo persistence.

MATERIALS AND METHODS

Bacterial growth conditions. A. actinomycetemcomitans strain 624, a rough serotype A clinical isolate, was used in this study. A. actinomycetemcomitans 624 was routinely cultured in brain heart infusion (BHI) broth or tryptic soy broth supplemented with 0.5% yeast extract (TSBYE) while shaking at 150 rpm in a 5% CO2 atmosphere at 37°C. For mutant A. actinomycetemcomitans, BHI and TSBYE were supplemented with 50 μg/ml spectinomycin. For RNA-seq, A. actinomycetemcomitans was grown in chemically defined, modified Socransky medium supplemented with 20 mM glucose and 50 mM morpholinepropanesulfonic acid (MOPS), pH 7.2 (CDM) (8, 20), while shaking at 150 rpm in a 5% CO2 atmosphere at 37°C. Prior to biofilm growth, overnight liquid cultures of A. actinomycetemcomitans were diluted 1:2 with fresh CDM, grown for 2 h, centrifuged at >16,000 × g, resuspended in CDM, and inoculated onto CDM agar plates as colony biofilms (21). Colony biofilms were grown for 4 h on CDM agar, transferred to fresh CDM agar for 4 more h of growth, and collected in RNAlater solution.

Received 24 July 2013. Accepted 14 August 2013

Address correspondence to Marvin Whiteley, mwhiteley@austin.utexas.edu.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JB.00875-13.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.
Determination of generation times. Wild-type A. actinomycetemcomitans 624 and the ΔfdhF1F2, ΔfrdABC, and Δpfl mutants were grown while shaking at 150 rpm in liquid CDM aerobically at 37°C in a 5% CO3 atmosphere and statically in an anaerobic chamber (Coy). Because A. actinomycetemcomitans 624 grows in large aggregates in vitro that cannot be adequately dispersed, cellular protein concentrations were determined with a Bradford assay (Bio-Rad) as a proxy for cellular growth and for calculation of generation times. To prepare lysates for a Bradford assay, cells were removed throughout the exponential growth phase, pelleted in a microcentrifuge, resuspended in 6 M urea, and boiled for 30 min at 100°C to lyse the cells. Generation times were calculated by plotting cellular protein concentrations over time.

A. actinomycetemcomitans murine abscess infections. Three-day murine abscess infections were established with wild-type and mutant A. actinomycetemcomitans strains as described previously (14). Severity of infections was determined by plate counting on BHI agar and on BHI agar supplemented with 50 μg/ml spectinomycin for the wild-type and mutant A. actinomycetemcomitans strains, respectively. The experiments described here were conducted according to the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The Texas Tech University Health Sciences Center Institutional Animal Care and Use Committee approved the protocol [99039].

RNA isolation and high-throughput sequencing library preparation. A. actinomycetemcomitans biofilm cells stored in RNALater solution were centrifuged at 5,000 × g, resuspended in 100 μl of 10 mg/ml lysozyme in TE buffer, and incubated for 10 min at 25°C. After lysozyme treatment, 200 μl of phosphate-buffered saline (PBS) was added to each lysed cell solution and further lysis was carried out by mechanical disruption of each sample four times for 30 s at maximum speed in a Mini-Beadbeater (Biospec Products) and lysed cell solutions were stored in an ice bath for 2 min between bead beatings. Total RNA was isolated from the resulting lysed cells with 1 ml RNA Bee solution (Tel-Test) according to the manufacturer’s protocol. Murine abscesses were resuspended immediately in 1 ml RNA Bee and bead beaten four times as described above, and RNA was purified from pooled abscesses (see Table S1 in the supplemental material). DNA contamination was removed from RNA samples by treating 5 μg of total RNA with 2.5 U of RQ1 DNase (Promega) for 30 min at 37°C, and RNA was purified with 500 μl of RNA Bee solution. DNA removal was verified by PCR amplification of the A. actinomycetemcomitans clpX protease gene from DNaSe-treated RNA (see Table S7). Bacterial and host RNAs were depleted by commercially available capture methods (Ambion MICROBExpress and MICROBEnrich kits) or enzymatic degradation (Epipcentre Terminator 5’-monophosphate-dependent RNase) according to the manufacturers’ protocols. To produce RNA between 20 and 500 nucleotides (nt) that is optimal for sequencing, 1 μg of an RNA-depleted RNA sample was treated with NEBNext RNA fragmentation buffer (NEB). Strand-specific cDNA libraries were prepared with commercially available T4 RNA ligase-based kits and sequenced on the SOLiD V4 and Illumina HiSeq2000 platforms. Both SOLiD and Illumina cDNA libraries were subjected to polyacrylamide gel extraction to purify cDNA between 130 and 500 nt, removing library adapters and primers from the samples. For a summary of the sequencing library preparation kits and sequencing outputs used, see Table S1 in the supplemental material.

Computational methods. The A. actinomycetemcomitans D7S-1 reference genome (GenBank accession no. ADCCF01000001.1) was used for RNA-seq read alignment (22). The 50-bp single-end SOLiD sequence reads were aligned in colospace to the reference genome by using SHRIMP version 2.2.1 to produce sam format read alignment files (23). The longer −100-bp Illumina sequence reads were prefiltered with FLEXBAR version 2.0 (24) to remove library adapter sequences from the sequence reads (Illumina index sequencing primer sequence, 5’-AGATCGGAAAG GCAACAGTCTGAACTCAGTACGC-3’; Illumina 3’ adapter sequence, 5’-TCGATATGCCGTCTCCTGCTGTCG-3’) and improve alignment with the reference genome. Filtered Illumina reads were aligned with the reference using Bowtie version 2.0.0 to produce read alignment files in sam format (25). The read alignment files were converted from sam to bam format, sorted, and indexed with Samtools (26). Individual read alignments in bam file format from each condition were visualized with the Integrative Genomics Viewer (IGV) (27, 28). By using the read alignments visualized with the IGV, TSS were manually recorded in GFF format by identifying the position where reads aligned upstream of genes (see Table S2 in the supplemental material). By an analogous approach, ncRNAs visualized in the IGV were manually annotated by the identification of reads mapping to intergenic regions and antisense to coding sequences (see Table S3). Whole-genome annotations for TSS and ncRNAs are provided in GFF and GenBank formats with instructions for opening the files on our laboratory’s website at http://web.biosci.utexas.edu/whiteley_lab/pages/resources.html (DataSets1.gff and DataSets2.gbk). Prior to the determination of differential gene expression, read alignment files produced by SHRIMP were reformatted for HTSeq read counting with a custom Perl script. The number of reads aligned to each gene and ncRNA in the newly annotated reference genome was calculated with HTSeq (http://www.huber.embld.de/users/anders/HTSeq), not including tRNAs and rRNAs. Prior to the calculation of differential gene expression, read counts per gene were summed for technical replicates in accordance with the DESeq recommendation. Differences in gene and ncRNA expression between A. actinomycetemcomitans biofilms and abscesses were determined on the basis of a negative binomial distribution with R package DESeq version 1.6.1 (29). Clusters of orthologous groups (COG) gene enrichment analysis was conducted with COGs determined previously (30). Enrichment of differentially regulated genes in a given COG category was determined by comparing the prevalence of up- or downregulated genes assigned to a specific COG category to the prevalence of genes in the entire genome assigned to that COG category. Enrichment of a COG category in either the up- or downregulated gene set relative to the genome was calculated with a Fisher’s exact test custom macro in Microsoft Excel. Whole-genome diagrams were produced with Circos version 0.55 (31).

Construction of A. actinomycetemcomitans deletion mutants. A. actinomycetemcomitans mutants were constructed for the formate dehydrogenase (FDH) H operon (fdhF1F2; D7S_2219-D7S_2220), the fumarate reductase operon (frdABC; D7S_1533-D7S_1536), and pyruvate formate lyase (pf; D7S_2028) by double homologous recombination. Wild-type genes of interest were replaced with aad9 (encodes spectinomycin resistance) from pVT1461 by overlap extension PCR and natural transformation of A. actinomycetemcomitans 624 (32, 33). The aad9 spectinomycin resistance gene was amplified from pVT1461 with Spec-F and Spec-R to generate Spec'. The PCR overlap extension constructs were amplified for the fdhF1F2 operon with primers fdhKO-P1F-PUSS and fdhKO-P1R-PUSS to generate the fdhKO-P1 PCR product and with primers fdhKO-P2F-Sp and fdhKO-P2R-PUSS to generate the fdhKO-P2 PCR product. Overlap extension PCR was done by mixing 500 ng each of fdhKO-P1, Spec', and fdhKO-P2 with primers fdhKO-P1F-PUSS and fdhKO-P2R-PUSS, making the fdhKO PCR product containing the −1-kb region upstream of the fdhF1F2 operon, aad9, and the −1-kb region downstream of the fdhF1F2 operon and A. actinomycetemcomitans-specific DNA uptake signal sequences on each end to enhance natural transformation. Prior to natural transformation, the −3-kb fdhKO overlap extension PCR product was gel extracted with a Fermentas gel extraction kit. The fdhF1F2 operon was replaced by natural transformation on tryptic soy agar supplemented with 0.5% yeast extract (TSAYE) with 5% heat-inactivated horse serum and 1 mM cyclic AMP in an anaerobic chamber via double homologous recombination with 1 μg of the fdhKO overlap extension PCR product. The resulting ΔfdhF1F2 mutants were selected on TSAYE with 50 μg/ml spectinomycin. The ΔfdhF1F2 mutant was verified by growing the mutant in TSBYE with 50 μg/ml spectinomycin, purifying genomic DNA with a Qiagen DNeasy kit, and PCR amplifying the aad9 gene with the Spec-R primer specific to the aad9 gene and the fdhKO-verify primer, which anneals upstream of the fdhKO-P1F-PUSS primer. The verification product was observed for ΔfdhF1F2 mutant genomic DNA but not wild-type genomic DNA. Analogous methods were used to...
generate the ΔfrdABCD mutant with primers frdKO-P1F- USS, frdKO-P1R-Sp, frdKO-P2F-Sp, frdKO-P2R- USS, and frdKO-verify, as well as the Δpfl mutant with primers pflKO-P1F- USS, pflKO-P1R-Sp, pflKO-P2F-Sp, pflKO-P2R- USS, and pflKO-verify. For the sequences of the primers used, see Table S6 in the supplemental material.

Sequence read accession number. Sequence reads from this study have been deposited in the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) under accession number SRP022893.

RESULTS

In vitro and in vivo RNA-seq. The goal of this study was to use RNA-seq to provide a high-resolution analysis of the Aggregatibacter actinomycetemcomitans genetic elements expressed during in vivo growth, with a specific interest in metabolic genes. While Aggregatibacter actinomycetemcomitans is most noted for its ability to cause oral infections, our laboratory has used an extraoral abscess model to study Aggregatibacter actinomycetemcomitans virulence (14). We prefer this model for the following reasons. (i) Aggregatibacter actinomycetemcomitans causes extraoral infections, including abscess infections; thus, the model has clinical relevance (17–19). (ii) The infectious dose is easily controlled and results in a contained infection that has been used to study the pathogenesis of oral bacteria (14, 34, 35). (iii) Unlike periodontal models of infection, infected tissue can be easily removed to assess disease severity (14) or for RNA isolation. In addition to performing transcriptome analysis in a relevant animal model of infection, it is also critical to examine Aggregatibacter actinomycetemcomitans gene expression during in vitro growth under defined conditions. Growth under known nutritional conditions is essential to provide a well-defined control transcriptome for comparison with the undefined in vivo transcriptome. Therefore, we grew Aggregatibacter actinomycetemcomitans as a colony biofilm on a solid defined medium with glucose as the sole energy source. Since glucose metabolism is well characterized and Aggregatibacter actinomycetemcomitans forms robust biofilms, this growth condition provided a well-defined, relevant control for the in vivo experiments.

Total RNA was harvested from biofilms and abscesses, and strand-specific cDNA libraries were subjected to RNA-seq analysis. Because murine abscesses contain small populations of bacteria (approximately 10^6 bacteria per abscess), it was essential to develop a method by which to efficiently purify RNA from infected tissue, deplete the abundant murine rRNA, and enrich for bacterial transcripts in order to obtain sufficient bacterial sequence reads for analysis. Since sequencing technologies are constantly evolving, we used two sequencing platforms to quantify transcript levels. However, it is important to note that all sequencing libraries were prepared by using similar RNA ligation-based protocols and each technology produces reads of sufficient length to map specifically to reference genomes. For both technologies, gene expression data were normalized among all replicates for differential expression analyses. The resulting sequence reads were processed and aligned to the Aggregatibacter actinomycetemcomitans genome to determine the Aggregatibacter actinomycetemcomitans transcriptome during in vitro and in vivo growth. In total, 7.1 million biofilm and 63.3 million in vivo sequence reads were aligned to the Aggregatibacter actinomycetemcomitans genome (see Table S1 in the supplemental material), resulting in averages of ~190 and 870 reads/gene for in vitro- and in vivo-grown Aggregatibacter actinomycetemcomitans, respectively. By using these data, we manually annotated Aggregatibacter actinomycetemcomitans TSS and ncRNAs. Differential expression of protein-coding genes and ncRNAs was determined after normalization for read number to identify in vivo-regulated genes.

Transcription start site mapping. To begin characterizing the primary transcriptome of Aggregatibacter actinomycetemcomitans, TSS were manually annotated by using the aligned RNA-seq data (as detailed in the “computational methods” portion of Materials and Methods). TSS were identified by manually recording the position where reads aligned upstream of annotated genes with the IGV (see Table S2 in the supplemental material) (27, 28). In total, 691 mRNA TSS were identified in Aggregatibacter actinomycetemcomitans (Fig. 1). Among these TSS, several previously characterized TSS, including katA, apiA, and lysT were verified by our RNA-seq analyses, thus validating our approach (21, 36). Because of the methods used to generate the RNA-seq libraries, it is important to note that the TSS identified represent authentic TSS, as well as the 5′ ends of processed RNAs. However, most bacterial transcripts are not specifically processed at a defined nucleotide but are progressively degraded (37). Moreover, since several of the TSS identified here...
correspond to TSS identified by other approaches (21, 36), many of these TSS likely represent true start sites.

Secondary TSS have been observed among genes in Helicobacter pylori and in a cyanobacterium, but the roles of alternate TSS are relatively obscure (38, 39). When TSS for genes differ from one growth condition to another, it suggests that the gene is transcribed from multiple promoters or the mRNA is differentially processed. We observed 11 genes with different TSS described from multiple promoters or the mRNA is differentially regulated 2-fold or more (Fig. 3A; see Tables S4 and S5 in the supplemental material). The ncRNA regions and antisense to protein-coding genes observed in the IGV regions and antisense to coding genes. Like TSS and operons, ncRNAs were identified on the basis of contiguous reads aligning with intergenic regions and antisense to protein-coding genes observed in the IGV (Fig. 2; see Table S3 in the supplemental material). The ncRNA sequences were collected and compared to the Rfam database to determine homology to ncRNAs characterized in other bacteria (40, 41). In biofilms and in vivo, A. actinomycetemcomitans expressed a number of ncRNAs, including riboswitches, cis-antisense RNAs (asRNAs), small ncRNAs (sRNAs), and clustered regularly interspaced short palindromic repeat RNAs (crRNAs) (Fig. 2). In total, A. actinomycetemcomitans expressed 210 ncRNAs, including 127 asRNAs, 3 pre-crRNAs, 3 riboswitches, 75 sRNAs, and 1 unidentified tRNA. Among these ncRNAs were housekeeping ncRNAs involved in transcription, translation, and protein secretion (Table 1) and several previously discovered and predicted ncRNAs (21, 42–47). In addition, many ncRNAs were expressed differently in vivo than during in vitro biofilm growth. Indeed, 80 out of 210 ncRNAs were differentially expressed in vivo, including 39 upregulated and 41 downregulated ncRNAs (Fig. 2). This result is significant but perhaps expected, since the roles of ncRNAs in in vivo growth and persistence have been described for other bacteria (48–50).

Fermentative metabolism and anaerobic respiration promote in vivo survival. Our main goal was to use RNA-seq to identify metabolic pathways that impact A. actinomycetemcomitans fitness in vivo. The hypothesis was that genes differentially regulated in vivo will be enriched for those important for growth in the murine abscess. To identify differentially regulated genes, in vitro and in vivo transcriptomes were compared. This analysis yielded 337 genes (~14% of all predicted genes) that were differentially regulated 2-fold or more (Fig. 3A; see Tables S4 and S5 in the supplemental material) in the abscess infection than in the in vitro biofilm. Of these genes, 107 were differentially regulated 5-fold or more. Notably, genes that encode the A. actinomycetemcomitans

![FIG 2 A. actinomycetemcomitans ncRNAs are differentially regulated in vivo. A. actinomycetemcomitans ncRNAs mapped along the genome to the positive strand (outer circle) and negative strand (inner circle). ncRNAs are colored according to type as follows: sRNA, red; asRNA, blue; riboswitch, green; pre-crRNA, light blue. The circular histogram shows log2 fold changes in ncRNA expression in vivo (~10 to 10, inner line to outer line). Bars are colored relative to the changes in ncRNA expression as follows: >2-fold upregulated, yellow; >5-fold upregulated, red; >2-fold downregulated, light blue; >5-fold downregulated, purple; no change, gray.](http://jb.asm.org/)

### TABLE 1 Rfam predictions for A. actinomycetemcomitans ncRNAs

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Start</th>
<th>Stop</th>
<th>Strand</th>
<th>RNA family(ies)</th>
<th>Rfam prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>D7S_0043.1</td>
<td>41437</td>
<td>41836</td>
<td>+</td>
<td>sRNA, RNase P class A</td>
<td>2 390 4.91E-40</td>
</tr>
<tr>
<td>D7S_0235.1</td>
<td>240380</td>
<td>240648</td>
<td>+</td>
<td>sRNA, gcvR</td>
<td>18 174 8.90E-23</td>
</tr>
<tr>
<td>D7S_0343.1</td>
<td>340697</td>
<td>340887</td>
<td>+</td>
<td>sRNA, tfoR</td>
<td>91 181 5.10E-03</td>
</tr>
<tr>
<td>D7S_0661.1</td>
<td>640001</td>
<td>640182</td>
<td>+</td>
<td>sRNA, His leader</td>
<td>42 171 4.28E-17</td>
</tr>
<tr>
<td>D7S_0743.1</td>
<td>704845</td>
<td>704936</td>
<td>+</td>
<td>tRNA</td>
<td>2 74 5.29E-14</td>
</tr>
<tr>
<td>D7S_0849.1</td>
<td>794824</td>
<td>795016</td>
<td>+</td>
<td>sRNA, 6S RNA</td>
<td>1 184 1.52E-23</td>
</tr>
<tr>
<td>D7S_1442.2</td>
<td>1323184</td>
<td>1323273</td>
<td>+</td>
<td>sRNA, C4</td>
<td>1 74 4.90E-12</td>
</tr>
<tr>
<td>D7S_1454.1</td>
<td>1330343</td>
<td>1330779</td>
<td>-</td>
<td>sRNA, tmRNA</td>
<td>1 366 7.92E-92</td>
</tr>
<tr>
<td>D7S_1607.1</td>
<td>1477807</td>
<td>1477998</td>
<td>-</td>
<td>Riboswitch, FMN&lt;sup&gt;a&lt;/sup&gt; riboswitch</td>
<td>1 182 5.17E-29</td>
</tr>
<tr>
<td>D7S_1716.1</td>
<td>1572027</td>
<td>1572259</td>
<td>-</td>
<td>Riboswitch, glycine riboswitch</td>
<td>87 231 1.00E-10</td>
</tr>
<tr>
<td>D7S_2270.1</td>
<td>2096164</td>
<td>2096351</td>
<td>+</td>
<td>sRNA bacterial SRP</td>
<td>47 145 2.08E-18</td>
</tr>
</tbody>
</table>

<sup>a</sup> ncRNA nucleotide sequences were collected and used to search Rfam database families (RNA family). Locus tags are in Table S3 in the supplemental material. Start and stop refer to the location of the ncRNA on the A. actinomycetemcomitans reference genome. The portions of the A. actinomycetemcomitans ncRNA sequences that possessed sequence homology to RNAs in the Rfam database are indicated (Start and End), as are the expectation (E) values.

<sup>a</sup> FMN, flavin mononucleotide.
comitans virulence factors leukotoxin (locus tag D7S_0615) and cytolethal distending toxin (locus tag D7S_2348) (51–55) were upregulated in vivo. While leukotoxin is upregulated under low-oxygen conditions (51), prior to this study, neither toxin was known to be upregulated during infection.

To gain a broader perspective on genes differentially regulated in vivo, we performed a COG enrichment analysis to define (56, 57) differentially regulated groups of genes. This analysis first clusters genes on the basis of their putative function and then examines whether a particular category is enriched for differentially regulated genes compared to what would be expected by chance. Using the COGs defined by Kittichotirat et al. (30), only one category, COG C, was significantly enriched in both up- and down-regulated genes. COG C includes genes involved in energy production and conversion (Fig. 3B), suggesting that A. actinomycetemcomitans undergoes substantial changes in metabolism during in vivo growth.

Included within the differentially regulated metabolic genes were operons that encode FDH H (fdhF1F2) and fumarate reductase (frdABCD), which are involved in fermentative metabolism and anaerobic respiration, respectively (Fig. 4). Thus, we hypothesized that these operons would be important for A. actinomycetemcomitans growth in the murine abscess. To test this hypothesis, individual mutants containing deletions of fdhF1F2 or frdABCD were constructed via allelic exchange. In addition, a strain containing a deletion of the gene that encodes pyruvate formate lyase (pfl) was also constructed. Since pfl was not differentially regulated in vivo, we hypothesized that it would not be critical for A. actinomycetemcomitans growth in the abscess; thus, this mutant served as a control for subsequent in vivo experiments.

Abscesses produced by both the ΔfdhF1F2 and ΔfrdABCD mutants contained ~10-fold less bacteria than wild-type infections (Fig. 5), indicating that these pathways are critical for A. actinomycetemcomitans fitness in the murine abscess. In contrast, the Δpfl mutant showed bacterial numbers similar to those of wild-type A. actinomycetemcomitans (Fig. 5). Importantly, deletion of the fdhF1F2 and frdABCD operons did not impact aerobic or anaerobic in vitro growth in a glucose-defined medium (data not shown). Collectively, these data suggest that the fdhF1F2 and frdABCD pathways are critical for A. actinomycetemcomitans growth in the murine abscess.
frdABCD operons are not only highly upregulated during in vivo growth but also impact A. actinomycetemcomitans fitness in the murine abscess.

**DISCUSSION**

In this study, a high-resolution transcriptome analysis of in vitro- and in vivo-grown A. actinomycetemcomitans was performed by RNA-seq. Protocols were developed to isolate total RNA from abscess infections and enrich for bacterial RNA before RNA-seq analysis. From these data, the terminal 5’ sequences of ~700 RNAs were mapped, most of which likely represent authentic transcriptional start sites. In addition, over 300 A. actinomycetemcomitans genes were shown to be differentially regulated during abscess infection in comparison with in vitro-grown bacteria. This study provides the first high-resolution transcriptome analysis of a bacterium during growth in an abscess and establishes a method for performing RNA-seq with samples containing predominantly host RNA.

The primary goal of this work was to identify metabolic pathways used by A. actinomycetemcomitans in vivo by focusing on metabolic genes differentially regulated during growth in an abscess. The use of defined in vitro growth conditions provided a robust control to probe A. actinomycetemcomitans metabolism during in vivo growth, allowing the identification of both fermentative and respiratory pathways important for A. actinomycetemcomitans growth in the murine abscess (Fig. 4). The enrichment of pathways including fdhF1F2 and frdABCD, which are involved primarily in the regeneration of NAD⁺ during anaerobic growth, suggests that A. actinomycetemcomitans experiences low oxygen levels in an abscess. This observation is curious in regard to a previous study by our group that demonstrated that during coculture infection with the peroxigenic oral bacterium Streptococcus gordontii, A. actinomycetemcomitans grows aerobically in an abscess (14). Why the A. actinomycetemcomitans-S. gordontii coculture abscess infection is aerobic is unknown, but these results emphasize that a pathogen may require distinct metabolic pathways to persist in an infection site dependent on whether it is alone or in the presence of other microbes. Of course, the infection site may also impact gene expression, and it will be interesting to examine A. actinomycetemcomitans gene expression in the oral cavity during mono- and coculture.

One of the primary advantages of RNA-seq is that it provides unparalleled insight into ncRNAs expressed under a given growth condition. Here we reported the differential regulation of 80 A. actinomycetemcomitans ncRNAs and increased expression of hsf in vivo. The upregulation of hsf is particularly striking, since it encodes an sRNA chaperone that promotes ncRNA interactions with mRNA targets (58, 59). Similar to other studies, this suggests that A. actinomycetemcomitans ncRNAs may play important regulatory roles in vivo (48, 49, 60). Differential regulation of ncRNAs also provided insight into the nutritional content of the infection site. For example, the A. actinomycetemcomitans lysine riboswitch and the sRNA GcvB (21, 61, 62) showed reduced abundance during infection. These ncRNAs regulate the expression of lysine and glycine transport proteins, respectively, and their decreased levels suggest that these amino acids are found at reduced levels in an abscess.

Transcriptome analysis by RNA-seq is a powerful and sensitive tool for studying infectious disease processes in vivo. The development of methods for performing RNA-seq with in vivo samples dominated by host RNA allowed tremendous insight into the physiology and metabolism of A. actinomycetemcomitans during in vivo growth. As evidenced by the abscess CFU counts, the method used here can be used to perform gene expression analyses of as few as 10⁵ in vivo bacteria. Ultimately, the application of this technology to defined polymicrobial infections, as well as undefined complex human infections, will provide a window into the physiology of bacterial pathogens in human infections.

**ACKNOWLEDGMENTS**

This work was supported by grants from the NIH (1R01DE020100 to M.W. and 5F31DE021633-02 to P.J. and M.W.). M.W. is a Burroughs Wellcome Investigator in the Pathogenesis of Infectious Disease.

We thank the Whiteley lab members for critical discussions of the manuscript and the University of Texas Genome Sequencing and Analysis Facility for discussions about computational analyses.

**REFERENCES**


